GLOBAL EDITION

Manufacturing Engineering and Technology

Eighth Edition in SI Units

Serope Kalpakjian | Steven R. Schmid

Manufacturing Engineering and Technology

EIGHTH EDITION IN SI UNITS

Serope Kalpakjian
Illinois Institute of Technology

Steven R. Schmid
The University of Notre Dame

SI Contributions by
K. S. Vijay Sekar
Sri Sivasubramaniya Nadar College of Engineering

Pearson

Please contact https://support.pearson.com/getsupport/s/ with any queries on this content.

Cover Photo: *Shutterstock/Praphan Jampala*

Pearson Education Limited

KAO Two
KAO Park
Hockham Way
Harlow
CM17 9SR
United Kingdom

and Associated Companies throughout the world

Visit us on the World Wide Web at: www.pearsonglobaleditions.com

Authorized adaptation from the United States edition, entitled *Manufacturing Engineering and Technology*, 8th Edition, ISBN 978-0-13-668165-6 by Serope Kalpakjian and Steven R. Schmid published by Pearson Education © 2020.

ISBN 10: 1-292-42224-6 (print)
ISBN 13: 978-1-292-42224-4 (print)
ISBN 13: 978-1-292-42229-9 (uPDF eBook)

British Library Cataloguing-in-Publication Data
A catalogue record for this book is available from the British Library

eBook formatted by B2R Technologies Pvt. Ltd.

Dedicated to our families, whose patience and
support made this book possible.

Contents

Part IV: Machining Processes and Machine Tools 633

Case Studies

Preface

Since the mid-2000s, manufacturing has undergone a rebirth in its development and research activities. With the recognition that manufacturing adds value to products, and wealth to national economies, governments around the world have been investing in their infrastructure and are now partnering with industry to bring new manufacturing capabilities to the global marketplace.

Manufacturing continues to be a dynamic activity, providing all-encompassing opportunities for contributions from several traditional disciplines. The proliferation of powerful software and Internet communication tools, especially their wireless capabilities, has made new approaches possible. The engineering terminology now includes such terms as Big Data, mass customization, cobots, and cybersecurity. Traditional manufacturing approaches and the materials involved are constantly being adjusted for ever-increasing efficiency and continuous improvements in performance.

In view of the advances being made in all aspects of manufacturing, the authors have continued their efforts to present a comprehensive, balanced, and, most importantly, an up-to-date coverage of the science, engineering, and technology of manufacturing. As in its previous editions, this text maintains the same number of chapters, while continuing to emphasize the interdisciplinary nature of all manufacturing activities, including complex interactions among materials, design, and manufacturing processes and operations.

Every attempt has been made to motivate and challenge students to understand and develop an appreciation of the vital importance of manufacturing in the modern global economy. The extensive questions and problems, at the end of each chapter, are designed to encourage students to explore viable solutions to a wide variety of challenges, giving them an opportunity to describe and assess the capabilities as well as limitations of all manufacturing processes and operations. These challenges include economic considerations and the competitive aspects in a global marketplace. The numerous examples and case studies throughout the book also help give students a perspective on real-world applications of the topics described throughout the book.

What's New in This Edition

- The eighth edition has been thoroughly updated, with numerous new topics and illustrations relevant to all aspects of manufacturing. See the table on page 23 for specifics.
- Wherever appropriate, illustrations and examples have been replaced, indicating recent advances in manufacturing.
- The text contains more cross references to other relevant sections, tables, and illustrations in the book.
- The Questions, Qualitative Problems, Quantitative Problems, and Design/Projects at the end of each chapter have been expanded.
- The Bibliographies at the end of each chapter have been thoroughly updated.
- *Manufacturing Engineering and Technology* is also available as an eText. Pearson eText offers a simple-to-use, mobile-optimized, personalized reading experience. It lets students add bookmarks, highlight, and take notes all in one place, even when offline. Seamlessly integrated videos engage students and give them access to the help they need, when they need it. Educators can easily schedule readings and share their own notes with students so they see the connection between their eText and what they learn in class—motivating them to keep reading, and keep learning. And, reading analytics offer insight into how students use the eText, helping educators tailor their instruction.

- The Solutions Manual, available for use by instructors, has been expanded; it now provides MATLAB code for numerous problems, allowing instructors to easily be able change relevant parameters.
- Reflecting the rapid advances in additive manufacturing, Chapter 20 has been thoroughly revised to include the latest technologies.

New or expanded topics in this edition are:

Chapter	Topics
Introduction	Complexity of products; definition of Technology Readiness Level and Manufacturing Readiness Level, to show the stages in product development and manufacturing at scale; case study on three-dimensional printing of guitars; expansion and update of the section on Trends in Manufacturing.
1	ISO for grain size number.
5	Second- and third-generation high-strength steels; nano-structured steels; and new case study on high-strength steels in automobiles.
6	Addition of the Hall-Héroult process for aluminum manufacture; new sections on lithium and rare earth metals; a case study on Tesla automobile design and manufacture; metamaterials.
7	Electrically conductive and semi-conductive polymers; gels and aerogels.
8	Porous ceramics; graphene; carbon and graphite foam.
10	Freeze casting.
11	Integrated computational materials engineering (ICME); machining of sand molds; new case study on a die-cast magnesium liftgate.
13	Tailor-rolled blanks and tailored coils.
15	Friction stir extrusion.
16	Expansion of hot stamping of sheet metal; camera-based forming-limit diagrams; electrically assisted forming; new case study on single-point incremental forming.
17	Expansion of powder morphology effects.
20	Additive Manufacturing: mass customization; distributed manufacturing; composite AM; projection stereolithography; continuous liquid interface production (CLIPS); new case study on AM of athletic shoes; powder bed processes; JetFusion; wire and arc AM; bioprinting; architectural applications of AM; conformal cooling; expanded section on design for AM; topology optimization; economic considerations; new case study on the implications of powder reuse.
23-24	Design considerations in machining; new case study on machining aerospace parts from monolithic aluminum.
25	Stability lobes in chatter.
26	Engineered abrasives; new case study on gear grinding with engineered abrasives.
27	Electrolytic trepanning; shaped-tube electrolytic machining.
28	Roll-to-roll printing; flexible electronics; conductive and semi-conductive inks; rotogravure, flexography, flat and rotary screen printing; self-aligned imprint lithography; flexible hybrid electronics; new case study on a flexible Arduino.
29	Photonic integrated circuits; mesoscale manufacturing.
31	Friction stir spot welding; expansion of linear friction welding.
37	Microcontrollers; cloud computing; cybersecurity; gain scheduling; cobot; design considerations.
38	Cloud storage; expansion of ERP and MES.
39	Mass customization; Internet of Things; cloud computing; MTConnect; Big Data; digital twin.
40	Life-cycle engineering; energy use in manufacturing.

Acknowledgments

We gratefully acknowledge individuals for their contributions to various sections in the book, as well as for their reviews, comments, and constructive suggestions for this revision.
We are happy to present below as list of those individuals, in academic institutions, as well as in industrial and research organizations, who, in one way or another, have made various contributions to this and the recent editions of this book. Kent M. Kalpakjian, Micron Technology, Inc., was the original author of the sections on fabrication of microelectronic devices.

D. Adams
W. Bair
N. Banu
G. Boothroyd
D. Bourell
J. Cesarone
A. Cinar
D. Dods
D.A. Dornfeld
M. Dugger
D.R. Durham
M. Giordano
M. Hawkins
T. Jaramillo
K. Jones
R. Kassing
K. Kozlovsky
K.M. Kulkarni
M. Madou
H. Malkani
J. Marsh
M. Molnar
S. Mostovoy
C. Nair
P.G. Nash
J. Neidig
C. Petronis
M. Prygoski
P. Saha
K. Savala
J. Schiltz
T. Schmitz
R. Shivpuri
K.S. Smith
A. Taub
B.S. Thakkar
J.E. Wang
K.R. Williams
Robert Ivester
John Vickers
P.K. Wright

We would also like to acknowledge the dedication and continued help and cooperation of Holly Stark, Executive Portfolio Manager at Pearson, and Amanda Brands, Content Producer at Pearson. We are grateful to numerous organizations that supplied us with illustrations, videos, and various materials for the text. These contributions have specifically been acknowledged throughout the text.

SEROPE KALPAKJIAN
STEVEN R. SCHMID

Acknowledgments for the Global Edition

Pearson would like to acknowledge and thank the following for the Global Edition:

Contributor

K. S. Vijay Sekar graduated from the University of Madras, Chennai, with a Bachelor's in Production Engineering and a Master's in Computer-aided Design. He completed his PhD in Mechanical Engineering from Anna University, Chennai, specializing in finite element analysis in machining.

He has over twenty-five years of teaching experience and more than fifteen years of research experience. His areas of research interest include materials, materials processing, and application of finite elements in manufacturing.

At present, he is a professor at the Department of Mechanical Engineering, Sri Sivasubramaniya Nadar College of Engineering (SSNCE), Kalavakkam, Tamil Nadu, India.

Reviewers

K. Baxevanakis, Loughborough University
C. Chatwin, University of Sussex
X. Song, The Chinese University of Hong Kong

About the Authors

Serope Kalpakjian is professor emeritus of Mechanical and Materials Engineering, the Illinois Institute of Technology. He is the author of *Mechanical Processing of Materials* and co-author of *Lubricants and Lubrication in Metalworking Operations* (with E.S. Nachtman). The first editions of his textbooks *Manufacturing Processes for Engineering Materials* and *Manufacturing Engineering and Technology* have received the M. Eugene Merchant Manufacturing Textbook Award. He has conducted research in various areas of manufacturing, is the author of numerous technical papers and articles in handbooks and encyclopedias, and has edited a number of conference proceedings. He also has been editor and co-editor of several technical journals and has served on various editorial boards, including the *Encyclopedia Americana*.

Among various awards, Professor Kalpakjian has received the Forging Industry Educational and Research Foundation Best Paper Award, the Excellence in Teaching Award from lIT, the ASME Centennial Medallion, the International Education Award from SME, A Person of the Millennium Award from lIT, the Albert Easton White Outstanding Teacher Award from ASM International, and the 2016 SME Gold Medal Award. The Outstanding Young Manufacturing Engineer Award of SME, for 2001, was named after him. Professor Kalpakjian is a Life Fellow ASME, Fellow SME, Fellow and Life Member ASM International, Fellow Emeritus International Academy for Production Engineering (CIRP), and is a founding member and past president of NAMRI. He is a graduate of Robert College (High Honor, Istanbul), Harvard University, and the Massachusetts Institute Technology.

Steven R. Schmid is professor of Aerospace and Mechanical Engineering, the University of Notre Dame, where he teaches and conducts research in the general areas of manufacturing, machine design, and tribology. From 2016 to 2018, he served as a Program Director of the Advanced Manufacturing Program at the National Science Foundation. He received his B.S. degree from Illinois Institute of Technology (with Honors) and Master's and Ph.D. degrees from Northwestern University, all in mechanical engineering. He has received numerous awards, including the John T. Parsons Award from SME, the Newkirk Award from ASME, the Kaneb Center Teaching Award (three times), and the Ruth and Joel Spira Award for Excellence in Teaching.

Professor Schmid served as President of the North American Manufacturing Research Institution (NAMRI, 2015–2016) and was appointed the first Academic Fellow at the Advanced Manufacturing National Program Office, U.S. Department of Commerce, where he helped design the Manufacturing USA program. Dr. Schmid is the author of over 140 technical papers, and has co-authored *Fundamentals of Machine Elements*, *Fundamentals of Fluid Film Lubrication*, *Manufacturing Processes for Engineering Materials*, and *Manufacturing Engineering and Technology*. He is a Fellow of the American Society of Mechanical Engineers and the Society of Manufacturing Engineers. He was elected to the Board of the Society of Manufacturing Engineers in 2018.

Chapter I

General Introduction

Case Studies:

I.1 What Is Manufacturing?

As you begin to read this chapter, take a few moments to inspect various objects around you: mechanical pencil, light fixture, chair, cell phone, and computer. You soon will note that all these objects, and their numerous individual components, are made from a variety of materials and have been produced and assembled into the products you now see. You also will note that some objects, such as a paper clip, nail, spoon, and door key, are made of a single component. However, as shown in Table I.1 and Fig. I.1, the vast majority of objects around us consist of numerous individual parts that are built and assembled by a combination of processes called **manufacturing** (Fig. I.2).

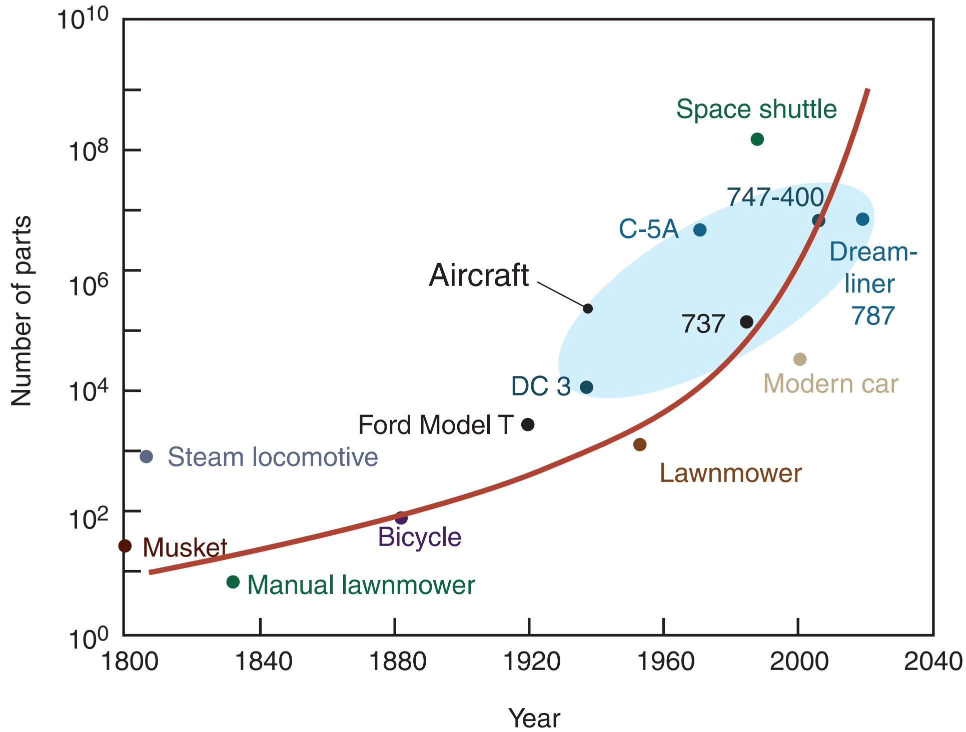


Figure I.1: Increasing complexity of products since the Industrial Revolution. *Source:* After J. Jeswiet.

The word *manufacture* first appeared in English in 1567, and is derived from the Latin *manu factus*, meaning made by hand. The word *manufacturing* first appeared in 1683, and the word *production*, which is often used interchangeably with the word *manufacturing*, first appeared sometime during the 15th century.

Manufacturing is making products. A manufactured product may, in turn, itself be used to make other products, such as (a) a large press, to shape flat sheet metal into appliances, (b) a drill, for producing holes, and (c) machinery, to produce a countless variety of individual items.

Nails, bolts, screws, nuts, paper clips, tires are *discrete products*, meaning individual items. By contrast, a spool of wire, metal or plastic tubing, and a roll of aluminum foil are *continuous products*, which are then cut into individual pieces of various lengths for specific purposes.

Because a manufactured item typically begins with raw materials, which are then subjected to a sequence of processes to make individual products, it has a certain *value*. Clay, for example, has some value as mined, but when made into pottery, electrical insulator, or cutting tool, it has *added value*. Similarly, a nail has a value over and above the cost of a short piece of wire from which it is made. Products such as computer chips, electric motors, medical implants, machine tools, and aircraft are known as *high-value-added* products.

Table I.1: Approximate Number of Parts in Products.

Common pencil	4
Rotary lawn mower	300
Grand piano	12,000
Automobile	15,000
Boeing 747–400	6,000,000

Figure I.2: John Deere tractor showing the variety of materials and processes incorporated. *Source:* Shutterstock/Nils Versemann.

A Brief History of Manufacturing. Manufacturing dates back to the period 5000 to 4000 B.C. (Table I.2). Thus, it is older than recorded history, which dates back to the Sumerians, around 3500 B.C. Primitive cave drawings, as well as markings on clay tablets and stones, needed (a) some form of a brush and some sort of pigment, as in the prehistoric cave paintings in Lascaux, France, estimated to be 16,000 years old; (b) a means of first scratching the clay tablets and then baking them, as in cuneiform scripts and pictograms of 3000 B.C.; and (c) simple tools for making incisions and carvings on the surfaces of stone, as in the hieroglyphs in ancient Egypt.

Manufacturing items for specific uses began with the production of household artifacts, typically made of wood, stone, or metal. The materials first used in making utensils and ornamental objects included gold, copper, and iron, followed by silver, lead, tin, bronze, and brass. The processing methods first employed involved mostly *casting* and *hammering*, because they were relatively easy to perform. Over the centuries, these simple processes gradually began to be developed into more and more complex operations, at increasing rates of production, and at higher levels of product quality. Note from Table I.2 that, for example, lathes for cutting screw threads already were available during the period from 1600 to 1700, but it was not until three centuries later that automatic screw machines were developed.

Although ironmaking began in about 1100 B.C. in the Middle East, a major milestone was the production of steel, in Asia, during the period 600 to 800 A.D. A wide variety of materials then began to be developed. Today, countless metallic and nonmetallic materials with unique properties are available, including *engineered materials* and other advanced materials. Among the available materials now are industrial ceramics, composite materials (often in the form of fiber reinforced plastics), and nanomaterials

Table I.2: Historical Development of Materials and Manufacturing Processes.

Period	Dates	Metals and casting	Various materials and composites	Forming and shaping	Joining	Tools, machining and manufacturing systems
	Before 4000 B.C.	Gold, copper, meteoric iron	Earthenware, glazing, natural fibers	Hammering		Tools of stone, flint, wood, bone, ivory, composite tools
Egypt: ~ 3100 B.C. to ~ 300 B.C. Greece: ~ 1100 B.C. to ~ 146 B.C. Roman Empire: ~ 500 B.C. to 476 A.D.	4000-3000 B.C.	Copper casting, stone and metal molds, lost-wax process, silver, lead, tin, bronze		Stamping, jewelry	Soldering (Cu-Au, Cu-Pb, Pb-Sn)	Corundum (alumina, emery)
	3000-2000 B.C.	Bronze casting and drawing, gold leaf	Glass beads, potter's wheel, glass vessels	Wire by slitting sheet metal	Riveting, brazing	Hoe making, hammered axes, tools for ironmaking and carpentry
	2000-1000 B.C.	Wrought iron, brass				
	1000-1 B.C.	Cast iron, cast steel	Glass pressing and blowing	Stamping of coins	Forge welding of iron and steel, gluing	Improved chisels, saws, files, woodworking lathes
	1-1000 A.D.	Zinc, steel	Venetian glass	Armor, coining, forging, steel swords		Etching of armor
Middle Ages: ~ 476 to 1492. Renaissance: 14th to 16th centuries.	1000-1500	Blast furnace, type metals, casting of bells, pewter	Crystal glass	Wire drawing, gold- and silver-smith work		Sandpaper, windmill-driven saw
	1500-1600	Cast-iron cannon, tinplate	Cast plate glass, flint glass	Water power for metalworking, rolling mill for coinage strips		Hand lathe for wood
	1600-1700	Permanent-mold casting, brass from copper and metallic zinc	Porcelain	Rolling (lead, gold, silver), shape rolling (lead)		Boring, turning, screw-cutting lathe, drill press

(continues on next page)

Table I.2: Historical Development of Materials and Manufacturing Processes (continued).

Period	Dates	Metals and casting	Various materials and composites	Forming and shaping	Joining	Tools, machining and manufacturing systems
Industrial Revolution: ~ 1750 to 1850	1700-1800	Malleable cast iron, crucible steel (iron bars and rods)		Extrusion (lead pipe), deep drawing, rolling		
	1800-1900	Centrifugal casting, Bessemer process, electrolytic aluminum, nickel steel, babbitt, galvanized steel, powder metallurgy, open-hearth steel	Window glass from slit cylinder, light bulb, vulcanization, rubber processing, polyester, styrene, celluloid, rubber extrusion, molding	Steam hammer, steel rolling, seamless tube, steel-rail rolling, continuous rolling, electroplating		Shaping, milling, copying lathe for gunstocks, turret lathe, universal milling machine, vitrified grinding wheel
WWI	1900-1920		Automatic bottle making, bakelite, borosilicate glass	Tube rolling, hot extrusion	Oxyacetylene; arc, electrical-resistance, and thermit welding	Geared lathe, automatic screw machine, hobbing, high-speed steel tools, aluminum oxide and silicon carbide (synthetic)
	1920-1940	Die casting	Development of plastics, casting, molding, polyvinyl chloride, cellulose acetate, polyethylene, glass fibers	Tungsten wire from metal powder	Coated electrodes	Tungsten carbide, mass production, transfer machines
WWII	1940-1950	Lost-wax process for engineering parts	Acrylics, synthetic rubber, epoxies, photosensitive glass	Extrusion (steel), swaging, powder metals for engineering parts	Submerged arc welding	Phosphate conversion coatings, total quality control
	1950-1960	Ceramic mold, nodular iron, semiconductors, continuous casting	Acrylonitrile-butadiene-styrene, silicones, fluorocarbons, polyurethane, float glass, tempered glass, glass ceramics	Cold extrusion (steel), explosive forming, thermochemical processing	Gas metal arc, gas tungsten arc, and electroslag welding; explosion welding	Electrical and chemical machining, automatic control.

(concludes on next page)

Table I.2: Historical Development of Materials and Manufacturing Processes (concluded).

Period	Dates	Metals and casting	Various materials and composites	Forming and shaping	Joining	Tools, machining and manufacturing systems
	1960-1970	Squeeze casting, single-crystal turbine blades	Acetals, polycarbonate, cold forming of plastics, reinforced plastics, filament winding	Hydroforming, hydrostatic extrusion, electroforming	Plasma-arc and electron -beam welding, adhesive bonding	Titanium carbide, synthetic diamond, numerical control, integrated circuit chip
Space Age	1970-1990	Compacted graphite, vacuum casting, organically-bonded sand, automation of molding and pouring, rapid solidification, metal-matrix composites, semi- solid metalworking, amorphous metals, shape-memory alloys	Adhesives, composite materials, semiconductors, optical fibers, structural ceramics, ceramic-matrix composites, biodegradable plastics, electrically-conducting polymers	Precision forging, isothermal forging, superplasticforming, dies made by computer-aided design and manufacturing, net-shape forging and forming, computersimulation	Laser beam, diffusion bonding (also combined with superplastic forming), surface-mount soldering	Cubic boron nitride, coated tools, diamond turning, ultraprecision machining, computer-integrated manufacturing, industrial robots, machining and turning centers, flexible manufacturing systems, sensor technology, automated inspection, computer simulation and optimization
Information Age	1990-2000	Rheocasting, computer-aided design of molds and dies, rapid tooling	Nanophase materials, metal foams, high-temperature superconductors. machinableceramics, diamond-like carbon	Additive manufacturing, rapid tooling, environmentally-friendly metalworking fluids	Friction stir welding, lead-free solders, laser butt-welded (tailored) sheet-metal blanks	Micro- and nanofabrication, LIGA, dry etching, linear motor drives, artificial neural networks, six sigma
	2000-2010s	TRIP and TWIP steels	Carbon nanotubes, graphene	Single point incremental forming, hot stamping, electrically assisted forming	linear friction welding	Digital manufacturing, three-dimensional computer chips, blue-arc machining, soft lithography, flexible electronics

that are now used in an extensive variety of products, ranging from prosthetic devices and computers to supersonic aircraft.

Until the **Industrial Revolution**, which began in England in the 1750s (also called the *First Industrial Revolution*), goods had been produced in batches, which required high reliance on manual labor in all phases of production. The *Second Industrial Revolution* is regarded, by some, as having begun in the mid-1900s, with the development of solid-state electronic devices and computers (Table I.2). **Mechanization** began in England and other countries of Europe with the development of textile machinery and machine tools for cutting metal. Mechanization soon moved to the United States, where it continued to be further developed.

A major advance in manufacturing began in the early 1800s, with the design, production, and use of **interchangeable parts**, conceived by the American manufacturer and inventor E. Whitney (1765–1825). Prior to the introduction of interchangeable parts, much hand fitting was necessary, because no two parts could be made exactly alike. By contrast, it is now taken for granted that a broken bolt can easily be replaced with an identical one produced decades after the original was made. Further developments soon followed, resulting in countless consumer and industrial products which we now cannot imagine being without.

Beginning in the early 1940s, several milestones were reached in all aspects of manufacturing, as can be observed by a review of Table I.2. Note particularly the progress that has been made during the 20th century, as compared with those achieved during the 40-century long period from 4000 B.C. to 1 B.C.

For example, in the Roman Empire (around 500 B.C. to 476 A.D.), factories were available for mass production of glassware; the methods used were generally very slow, and much manpower was required in handling the parts and operating the machinery. Today, production methods have advanced to such an extent that (a) aluminum beverage cans are made at rates of more than 500 per minute, with each can costing about four cents to make; (b) holes in sheet metal can be punched at rates of 800 holes per minute; and (c) light bulbs are made at rates of more than 2000 bulbs per minute, each costing less than one dollar.

The period from the 1940s to the 1990s was characterized by *mass production* and expanding *global* markets. Initially, the United States had a dominant position, as it was the only developed nation with an intact infrastructure following World War II; however, this advantage dissipated by the 1960s. The *quality* revolution began to change manufacturing in the 1960s and 1970s, and in the 1980s, programmable computers became widely used.

The **digital manufacturing** era began around 1990. As a fundamental change in manufacturing operations, powerful computers and software are now fully integrated across the design and manufacturing enterprise. Advances in communications, some Internet-based, have led to further improvements in organizations and their capabilities. The effects are most striking when considering the origin and proliferation of **additive manufacturing**, described in Chapter 20.

Prior to 1990, the prototype of a part could be produced only through intensive effort and costly manufacturing approaches, requiring significant operator skill. Today, a part can first be drafted in a CAD program, then produced generally in a matter of minutes or hours (depending on size and part complexity) without the need for hard tools or skilled labor. Prototyping systems have become more economical, faster, and with improved raw materials. The term *digital manufacturing* has been applied to reflect the notion that manufacturing parts and components can take place completely through such computer-driven CAD and production machinery.

Recent innovations are the proliferation of communications protocols, sensors and controls throughout the manufacturing enterprise. Referred to as **Industry 4.0** (Section 37.2.2) or *Digital Manufacturing*, some of the key developments are the following:

1. Sensors and smart device designs can be printed from conductive and insulate inks, or use very thin silicon integrated circuits in flexible hybrid designs (Fig. I.3). These low-cost devices include communication ability and are central to the **Internet of Things** (Section 39.8.1).

2. Machines can be monitored at all times, using Internet-based communications protocols such as MTConnect, so that precise information is available at all times for every machine in an organization.

3. The wide application of sensors to all aspects of manufacturing has led to the development of **Big Data**, where trends and conditions of manufacturing systems can be accurately measured at all times.

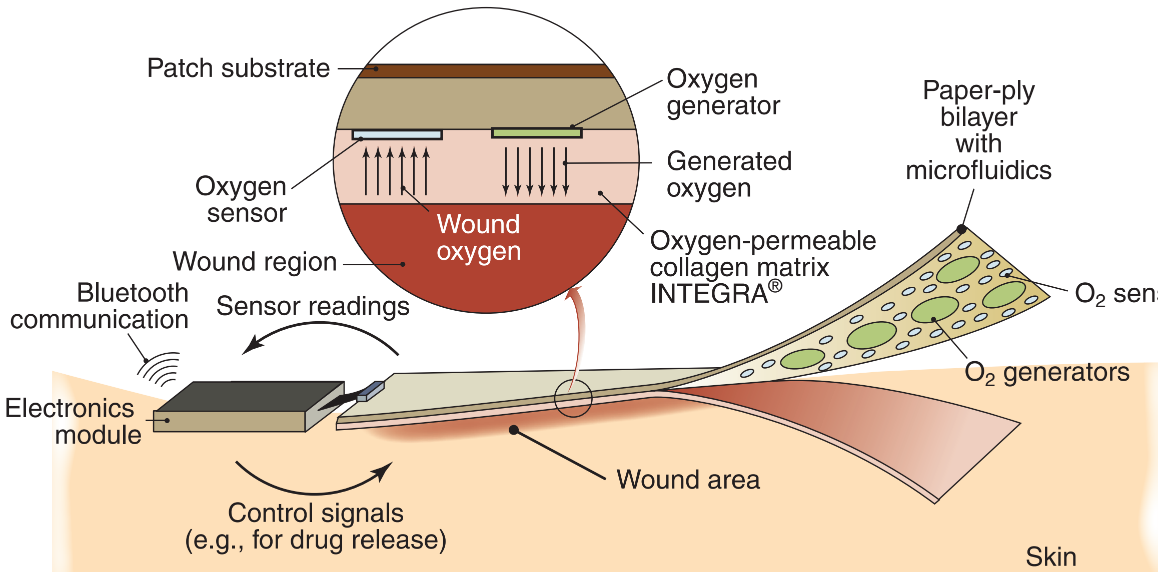


Figure I.3: A flexible hybrid electronic wound care system (see Section 28.15). The device integrates communication ability (in this case a Bluetooth ability), sensors to detect the oxygen levels at the wound, and oxygen generating devices to increase oxygen level to optimize healing. These devices use combinations of printed electronic devices and silicon-based integrated circuits. *Source:* Courtesy of NextFlex.

4. Machine learning algorithms, and the incorporation of physics-based mathematical models of manufacturing processes and systems, has led to the ability to apply advanced controls to the manufacturing enterprise.

5. Advanced models of manufacturing processes, combined with the detailed measurement of the manufacturing and service environment of a product (Big Data), lead to the computer-based representation of the product, referred to as a *digital twin*. The digital twin represents a virtual model of the part, and accurate performance models applied to the virtual twin can predict failure or required service of the actual part.

These developments are a natural extension of the computer revolution that started in the 1990s, and developments are certain to continue.

I.2 Product Design and Concurrent Engineering

Product design involves the creative and systematic prescription of the shape and characteristics of a product to achieve specified objectives, while simultaneously satisfying several important constraints. Design is a critical activity, because it has been estimated that as much as 80% of the cost of product development and manufacturing is determined by the decisions made in the *initial* stages of design. The product design process has been studied extensively; it is briefly introduced here because of the strong interactions among manufacturing and design activities.

Innovative approaches are essential in successful product design, as are clearly specified functions and a clear statement of the performance expected of the product. The market for a product, which may be new or a modified version of an existing product, and its anticipated use or uses, also must be clearly defined at this stage. This aspect also involves the assistance of market analysts and sales personnel who will bring valuable and timely input to the manufacturer, especially regarding market needs and trends.

Table I.3: Definitions of Technology Readiness Level (TRL) and Manufacturing Readiness Level (MRL).

TRL	Description	MRL	Description
1	Basic principles observed and reported	1	Manufacturing feasibility assessed
2	Technology concept and/or application formulated	2	Manufacturing concepts defined
3	Analytical and experimental critical function and/or characteristic proof of concept	3	Manufacturing concepts developed
4	Component and/or breadboard validation in a laboratory environment	4	Capability to produce the technology in a laboratory environment
5	Component or breadboard validation in a relevant environment	5	Capability to produce prototype components in a production relevant environment
6	System/subsystem model or prototype demonstration in a relevant environment	6	Capability to produce a prototype system or subsystem in a production relevant environment
7	System prototype demonstration in an operational environment	7	Capability to produce systems, subsystems or components in a production representative environment
8	Actual system completed and qualified through test and demonstration	8	Pilot line capability demonstrated; Ready to begin low rate initial production
9	Actual system proven through successful mission operations	9	Low rate production demonstrated; capability in place to begin full rate production

Technology Readiness Level and Manufacturing Readiness Level. Product development generally follows the flow outlined in Table I.3. *Technology readiness level* (TRL) and *Manufacturing readiness level* (MRL) are measures of a products ability to be produced, marketed, and sold. In practice, all technologies must progress from some starting point up to a TRL and MRL of 9. A new scientific discovery or a product idea begins at a TRL of 1, and it may or may not ever be suitable for commercial application. New versions of existing products may start at some higher TRL or MRL level, but the flow of its development is always the same.

Note that each stage of a products development typically requires different skills and resources. Demonstrating a new concept in a laboratory environment (TRL 3) and demonstrating it in a new system in a real environment (TRL 7) are very different tasks. Similarly, producing a laboratory prototype (MRL 4) is very different from demonstrating manufacturing strategies for producing a product at scale (MRL 7), which is also very different from having a production facility in place.

The Design Process. Traditionally, design and manufacturing activities took place *sequentially*, as shown in Fig. I.4a. This methodology may, at first, appear to be straightforward and logical; in practice, however, it is wasteful of resources. Consider the case of a manufacturing engineer who, for example, determines that, for a variety of reasons, it would be more desirable to (a) use a different material, such as a polymer or a ceramic instead of metal; (b) use the same material but in a different condition, such as a softer instead of a harder or one with a smoother surface finish; or (c) modify the design of a component in order to make it easier, faster, and less costly to manufacture. Note that these decisions must take place at the material-specification stage (the sixth box from the top in Fig. I.4a).

Each of the modifications just described will necessitate a repeat of the design analysis stage (the third box from the top in Fig. I.4a) and the subsequent stages. This approach is to ensure that the product will still meet all specified requirements and will function satisfactorily. A later change from, say, a forged, cast, or machined component will, likewise, necessitate a repeat analysis. Such iterations obviously waste both time and the resources of a company.

Concurrent Engineering. Driven primarily by the consumer electronics industry, a continuing trend has been to bring products to the marketplace as rapidly as possible, so as to gain a higher percentage share of the market and thus higher profits. An important methodology aimed at achieving this end is *concurrent engineering*, which involves the product-development approach shown in Fig. I.4b.

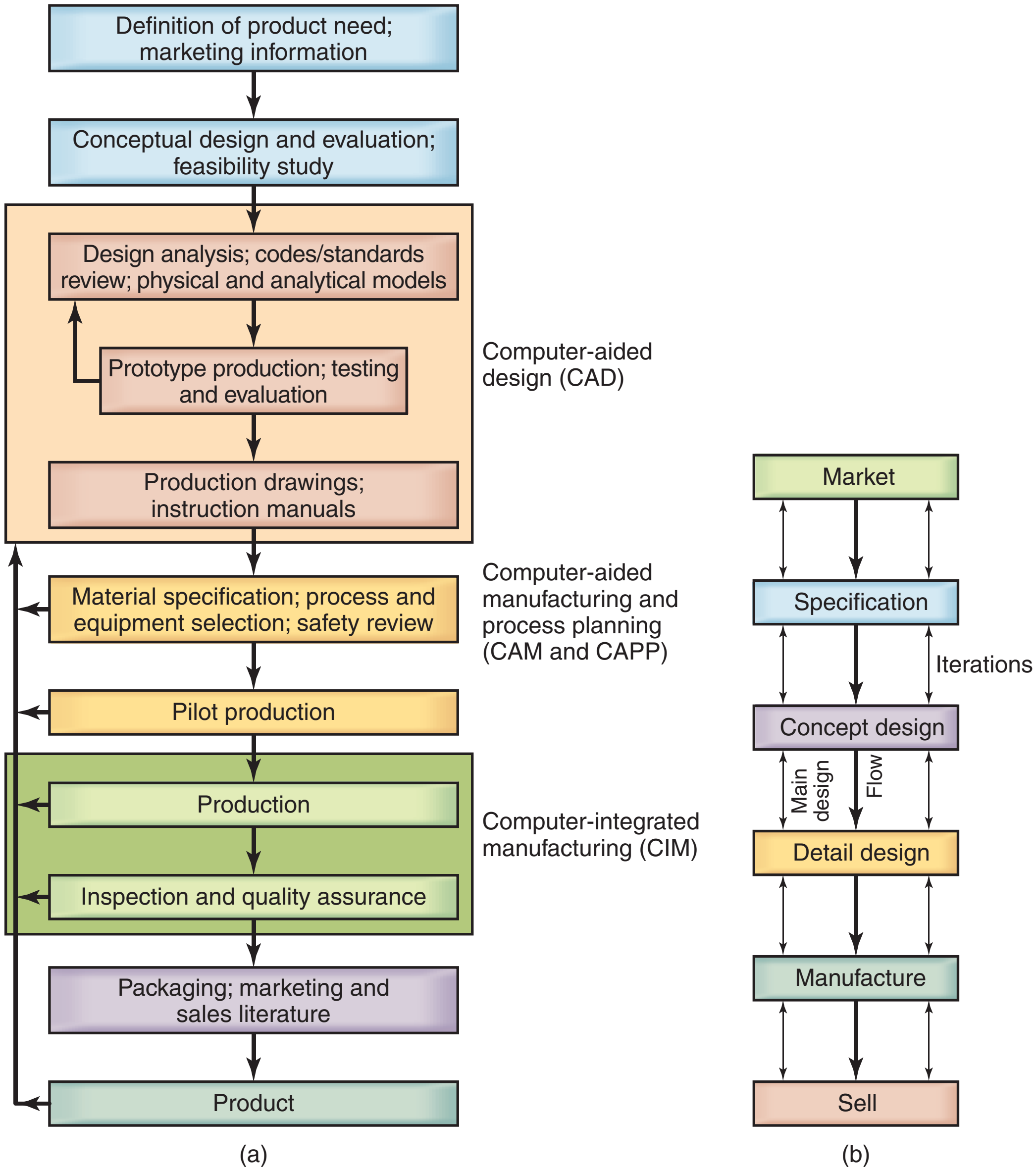


Figure I.4: (a) Chart showing various steps involved in *traditional* design and manufacture of a product. Depending on the complexity of the product and the type of materials used, the time span between the original concept and the marketing of the product may range from a few months to several years. (b) Chart showing general product flow in *concurrent engineering*, from market analysis to marketing the product. *Source:* After S. Pugh.

Although this concept still has the same general product-flow sequence as in the traditional approach, shown in Fig. I.4a, it now includes several deliberate modifications. From the earliest stages of product design and engineering, all relevant disciplines are now *simultaneously considered*. As a result, any iterations that may have to be made will require a smaller effort, resulting in much less wasted time than occurs in the traditional approach to design. It should be apparent that a critical feature of this approach is the recognition of the importance of *communication* among and within all disciplines.

Concurrent engineering can be implemented in companies large or small, which is particularly significant because 98% of all U.S. manufacturing companies have fewer than 500 employees; the companies are generally referred to as *small businesses* or *small manufacturing enterprises* (SMEs). As an example of the benefits of concurrent engineering, one automotive company reduced the number of components in one of its engines by 30%, decreased the engine weight by 25%, and reduced its manufacturing time by 50%.

Life Cycle. In concurrent engineering, the design and manufacture of products are integrated, with a view toward optimizing all elements involved in the *life cycle* of the product (see Section I.4). The life cycle of a new product generally consists of four stages:

1. Product start-up
2. Rapid growth of the product in the marketplace
3. Product maturity
4. Decline.

Consequently, **life-cycle engineering** requires that the *entire life* of a product be considered, beginning with the design stage and on through production, distribution, use, and, finally, recycling or the disposal of the product.

Role of Computers in Product Design. Typically, product design first requires the preparation of *analytical* and *physical models* of the product, for the purposes of visualization and engineering analysis. Although the need for such models depends on product complexity, constructing and studying these models are now done using **computer-aided design** (CAD) and **computer-aided engineering** (CAE) techniques.

CAD systems are capable of rapid and complete analyses of designs, whether it is a simple part in large and complex structures. The Boeing 777 passenger airplane, for example, was designed completely by computers, in a process called **paperless design**, with 2000 workstations linked to eight design servers. Unlike previous mock-ups of aircraft, no prototypes or mock-ups were built and the 777 was built and assembled *directly* from the CAD/CAM software that had been developed.

Through computer-aided engineering, the performance of structures subjected, for example, to static or fluctuating loads or to temperature gradients also can be simulated, analyzed, and tested, rapidly and accurately. The information gathered is stored, and it can be retrieved, displayed, printed, and transferred anytime and anywhere within an organization. Design modifications can be made and optimized directly, easily, and at any time.

Computer-aided manufacturing involves all phases of manufacturing, by utilizing and processing large amounts of information on materials and processes gathered and stored in the organization's database. Computers greatly assist in such tasks as (a) programming for numerical-control machines and for robots for material-handling and assembly operations (Chapter 37), (b) designing tools, dies, molds, fixtures, and work-holding devices (Parts II, III, and IV), and (c) maintaining quality control throughout the total operation (Chapter 36).

On the basis of the models developed and analyzed, product designers finalize the geometric features of each of the product's components, including specifying their dimensional tolerances and surface characteristics. Because all components, regardless of their size, eventually have to be *assembled* into the final product, dimensional tolerances are a major consideration in manufacturing (Chapter 35). The models developed also allow the specification of the mechanical and physical properties required, which in turn affect the selection of materials. (Section I.5).

Prototypes. A *prototype* is a physical model of an individual component or product. The prototypes developed are carefully reviewed for possible modifications to the original design, materials, or production methods. An important and continuously evolving technology is **additive manufacturing** (Chapter 20). Using CAD/CAM and various specialized technologies, designers make prototypes rapidly and at low cost, from a variety of metallic or nonmetallic materials.

Additive manufacturing significantly reduces costs and associated product-development times. The technology has now advanced to such a level that it is used for low-volume economical production of a variety of *actual and functional parts*.

(a) (b)

Figure I.5: Guitars produced through additive manufacturing. (a) Spider design being removed from a powder bed. Note that the support material, or *cake*, has some strength and needs to be carefully removed. (b) Finished Spider guitars. *Source:* Courtesy of O. Diegel.

Case Study I.1 Three-dimensional Printing of Guitars

The design flexibility of additive manufacturing is illustrated by the custom guitars produced by ODD, Inc. These guitars are designed in CAD programs, with full artistic freedom to pursue innovative designs; those in Fig. I.5 are only a selection of the many available. The CAD file is then sent to a three-dimensional printer, using the selective laser sintering process and produced from nylon (Duraform PA). As printed, the guitars are white; they are first dyed to a new base color, then hand-painted and sprayed with a clear satin lacquer. The customer-specified hardware (pickups, bridges, necks, tuning heads, etc.) are then mounted to produce the electric guitar.

Virtual Prototyping. This is a software-based method that uses advanced graphics and virtual-reality environments to allow designers to view and examine a part in detail. This technology, also known as **simulation-based design**, uses CAD packages to render a part such that, in a 3-D interactive virtual environment, designers can observe and evaluate the part as it is being developed. Virtual prototyping has been gaining importance, especially because of the availability of low-cost computers and simulation and analysis tools.

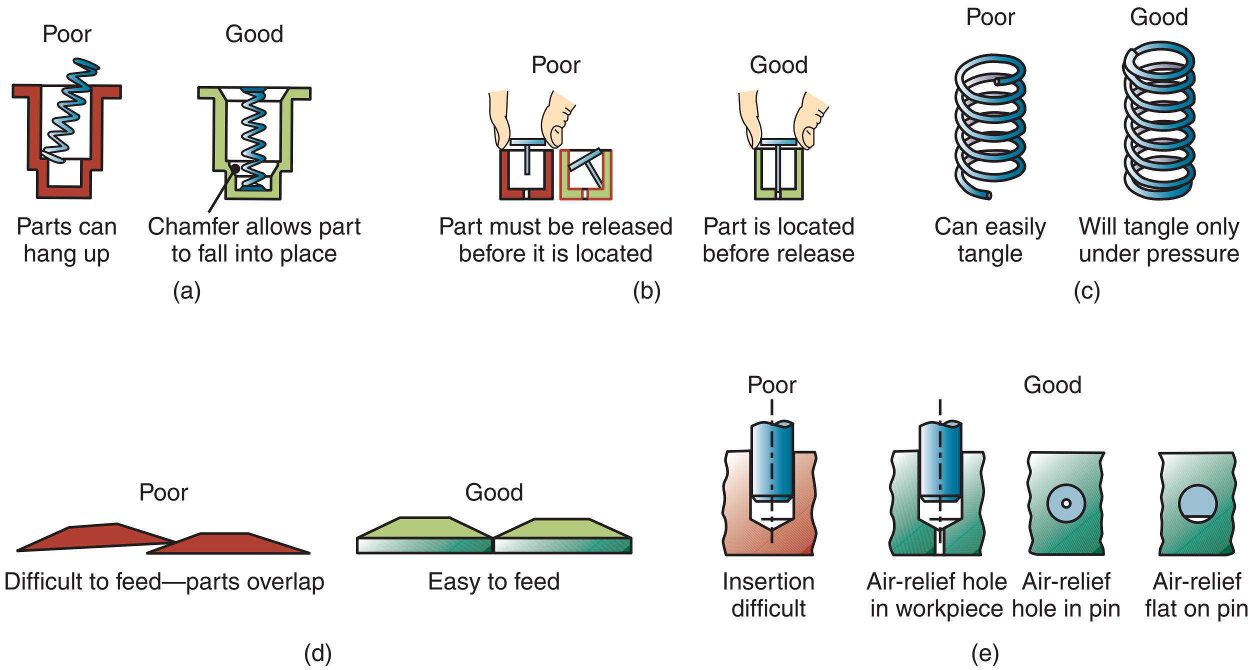


Figure I.6: Redesign of parts to facilitate assembly. *Source:* After G. Boothroyd and P. Dewhurst.

I.3 Design for Manufacture, Assembly, Disassembly, and Service

Design for manufacture (DFM) is a *comprehensive* approach to integrating the design process with production methods, materials, process planning, assembly, testing, and quality assurance. DFM requires a fundamental understanding of (a) the characteristics and capabilities of materials, manufacturing processes, machinery, equipment, and tooling, and (b) variability in machine performance, dimensional accuracy and surface finish of the workpiece, processing time, and the effect of processing methods employed on product quality. Establishing *quantitative relationships* is essential in order to be able to analyze and optimize a design, for ease of manufacturing and assembly at the lowest cost.

The concepts of **design for assembly** (DFA), **design for manufacture and assembly** (DFMA), and **design for disassembly** (DFD) are all important considerations in manufacturing. Methodologies and computer software are available for design for assembly, utilizing 3-D conceptual designs and solid models. Subassembly, assembly, and disassembly times and costs are minimized, while maintaining product integrity and performance.A product which is easy to assemble is usually also easy to disassemble.

Assembly is an important phase of manufacturing, requiring considerations of the ease, speed, and cost of putting together the numerous individual components of a product (Fig. I.6). Assembly costs in manufacturing can be substantial, typically ranging from 20% to 60% of the total product cost. *Disassembly* of a product, for such activities as maintenance, servicing, and eventual recycling of its individual components, is an equally important consideration.

As described in Part VI, there are several methods of assembly of components, including the use of a wide variety of fasteners, adhesives, or joining techniques, such as welding, brazing, or soldering. As is the case in all types of manufacturing, each of these assembly operations has its own specific characteristics, times, advantages, limitations, associated costs, and design considerations. Individual parts may be assembled by hand or by a variety of automatic equipment and industrial robots. The choice depends on several factors, such as product complexity, number of components to be assembled, care and protection required to prevent damage to the parts, and relative cost of labor compared with the cost of machinery required for automated assembly.

Design for Service. In addition to design for assembly and for disassembly, *design for service* is an important aspect of product design. Products often have to be disassembled, to varying degrees, in order to service them and, if necessary, repair them. The design should take into account the concept that, for ease of access, components that are most likely to be in need of servicing be placed, as much as possible, at the *outer layers* of the product. This methodology can be appreciated by anyone who has had the experience of servicing machinery.

I.4 Environmentally Conscious Design, Sustainable Manufacturing, and Product Life Cycle

In the United States alone, more than 30 billion kg of plastic products are discarded each year, and 62 billion kg of paper products. Every three months, U.S. industries and consumers discard enough aluminum to rebuild the country's commercial air fleet.

Globally, countless metric tons of automobiles, televisions, appliances, and computers are discarded each year. Metalworking fluids, such as lubricants and coolants, and fluids and solvents, such as those used in cleaning manufactured products, can pollute the air and waters, unless they are recycled or disposed of properly.

Likewise, there are numerous *byproducts* from manufacturing plants: (a) sand with additives from foundries; (b) water, oil, and various other fluids from heat-treating and facilities; (c) slag from foundries and welding operations; and (d) a wide variety of metallic and nonmetallic scrap produced in such operations as sheet forming, casting, and molding. Consider also the various effects of water and air pollution, acid rain, ozone depletion, hazardous wastes, landfill seepage, and global warming. Recycling efforts have gained increasing momentum over the years; aluminum cans, for example, are now recycled at a rate of 67% and plastics at around 9%.

Note that, as indicated below, the term *discarding* suggests that the product has reached the end of its useful life; however, it does not necessarily indicate that it has to be dumped into landfills.The particular manufacturing process and the operation of machinery can each have a significant environmental impact. Manufacturing operations generally produce some waste, such as:

1. Chips from machining and trimmed materials from sheet forming, casting, and molding operations
2. Slag from foundries and welding operations
3. Additives in sand used in sand-casting operations
4. Hazardous waste and toxic materials used in various products
5. Lubricants and coolants in metalworking and machining operations
6. Liquids from such processes as heat treating and plating
7. Solvents from cleaning operations
8. Smoke and pollutants from furnaces and gases from burning fossil fuels.

The adverse effects of these activities, their damage to the environment and to the Earth's ecosystem, and, ultimately, their effect on the quality of human life are now widely recognized. Major concerns involve global warming, greenhouse gases (carbon dioxide, methane, and nitrous oxide), acid rain, ozone depletion, hazardous wastes, water and air pollution, and contaminant seepage into water sources. One measure of the adverse impact of human activities is called the **carbon footprint**, which quantifies the amount of greenhouse gases produced in our daily activities.

The term **sustainable design and manufacturing** has become in common usage in all industrial activities, with major emphasis on **design for the environment** (DFE). Also called **environmentally conscious design and manufacturing** and **green design**, this approach considers *all* possible adverse environmental impacts of materials, processes, operations, and products, so that they can all be taken into account at the earliest stages of their design and production.

These goals also have led to the concept of **design for recycling** (DFR). Recycling may involve one of two basic activities:

- **Biological cycle**: Organic materials degrade naturally, and in the simplest version of a biological cycle, they lead to new soil that can sustain life. Thus, product design involves the use of organic materials, as well as ensuring that products function well for their intended life and can then be safely disposed of.

- **Industrial cycle**: The materials in the product are recycled and reused continuously. To demonstrate the economic benefits of this approach, it has been estimated that producing aluminum from scrap, instead of from bauxite ore, reduces production costs by as much as 66% and reduces energy consumption and pollution by more than 90%.

A basic principle of *design for recycling* is the use of materials and product design features that facilitate biological or industrial recycling. In the U.S. automotive industry, for example, about 75% of automotive parts (mostly metal) are now recycled, and there are continuing plans to recycle the rest as well, including plastics, glass, rubber, and foam. About 80% of the 300 million discarded automobile tires are reused in various ways.

Cradle-to-cradle Production. Also called *cradle-to-cradle* (C2C), manufacturing considers the impact of each stage of a product's life cycle, from the time natural resources are mined and processed into raw materials, through each stage of manufacturing products, their use and, finally, recycling. *Cradle-to-grave* production, also called *womb-to-tomb* production, has a similar approach, but it does not necessarily consider or take on the responsibility of recycling.

Cradle-to-cradle production emphasizes

1. Sustainable and efficient manufacturing activities, using clean technologies
2. Waste-free production
3. Using recyclable and nonhazardous materials
4. Reducing energy consumption
5. Using renewable energy, such as wind, solar, and ocean waves
6. Maintaining ecosystems by minimizing the environmental impact of all manufacturing activities
7. Using materials and energy sources that are available locally, so as to reduce energy use associated with their transport which, by and large, has an inherently high carbon footprint
8. Continuously exploring the reuse and recycling of materials, and perpetually trying to recirculate materials; also included is investigating the composting of materials whenever appropriate or necessary, instead of dumping them into landfills.

Guidelines for Sustainable Design and Manufacturing. In reviewing the activities described thus far, it can be noted that there are overarching relationships among the basic concepts of DFMA, DFD, DFE, and DFR. These relationships can be summarized as guidelines, rapidly accepted worldwide:

1. Reduce waste of materials, by refining product design, reducing the amount of materials in products, and select manufacturing processes that minimize scrap (such as forming instead of machining).
2. Reduce the use of hazardous materials in products and processes.
3. Investigate manufacturing technologies that make environmentally friendly and safe products and by-products.
4. Make improvements in methods of recycling, waste treatment, and reuse of materials.
5. Minimize energy use; whenever possible, encourage the use of renewable sources of energy. Select materials can have a major impact on the latent energy in products, as described in Section 40.5.
6. Encourage recycling by using materials that are a part of either industrial or biological cycling, but not both in the same product. Ensure proper handling and disposal of all waste of materials that are used in products, but are not appropriate for industrial or biological cycling.

I.5 Selection of Materials

An increasingly wide variety of materials are now available, each type having its own properties and manufacturing characteristics, advantages, limitations, and costs (Part I). The selection of materials for products (consumer or industrial) and their components is typically made in consultation with materials engineers; design engineers may also be sufficiently experienced and qualified to assist.

The general types of materials used, either individually or in combination with other materials, are the following:

- **Ferrous metals:** Carbon, alloy, stainless, and tool and die steels (Chapter 5)
- **Nonferrous metals:** Aluminum, magnesium, copper, nickel, titanium, superalloys, refractory metals, beryllium, zirconium, low-melting-point alloys, and precious metals (Chapter 6)
- **Plastics (polymers):** Thermoplastics, thermosets, and elastomers (Chapter 7)
- **Ceramics, glasses, glass ceramics, graphite, diamond,** and **diamond-like materials** (Chapter 8)
- **Composite materials:** Reinforced plastics and metal-matrix and ceramic-matrix composites (Chapter 9)
- **Nanomaterials** (Section 8.8)
- **Shape-memory alloys** (*smart materials*), **amorphous alloys, semiconductors**, and **superconductors** (Chapters 6, 18 and 28)

As new developments continue, selection of an appropriate material for a particular application from a very large variety of materials has become even more challenging. Furthermore, there are continuously shifting trends in the substitution of materials, driven not only by technological considerations, but also by economics.

Properties of Materials. *Mechanical properties* of interest in manufacturing generally include strength, ductility, hardness, toughness, elasticity, fatigue, and creep resistance (Chapter 2). *Physical properties* are density, specific heat, thermal expansion and conductivity, melting point, and electrical and magnetic properties (Chapter 3). Optimum designs often require a consideration of a combination of mechanical and physical properties. A typical example is the *strength-to-weight* and *stiffness-to-weight* ratios of materials for minimizing the weight of structural members. Weight minimization is particularly important for aerospace and automotive applications, in order to improve performance and fuel economy.

Table I.4: General Manufacturing Characteristics of Various Materials.

Alloy	Castability	Weldability	Machinability
Aluminum	E	F	E–G
Copper	G–F	F	G–F
Gray cast iron	E	D	G
White cast iron	G	VP	VP
Nickel	F	F	F
Steels	F	E	F
Zinc	E	D	E

Note: E, excellent; G, good; F, fair; D, difficult; VP, very poor. The ratings shown depend greatly on the particular material, its alloys, and its processing history.

Chemical properties include oxidation, corrosion, degradation, toxicity, and flammability. These properties play a significant role under both hostile (such as corrosive) and normal environments. *Manufacturing properties* indicate whether a particular material can be cast, formed, shaped, machined, joined, and heat treated with relative ease. As Table I.4 illustrates, no one material has the same manufacturing characteristics. Another important consideration is *appearance*, which includes such characteristics as surface texture, color, and feel, all of which can play a significant role in a product's acceptance by the public.

Availability. As emphasized throughout this book, the economic aspect of material selection is as important as technological considerations (Chapter 40). Availability of materials is a major concern in manufacturing. Furthermore, if materials are not available in the shapes needed, dimensions, surface texture, and quantities, materials substitution or additional processing of a particular material may well be required, all of which can contribute significantly to product cost.

Reliability of supply is important in order to meet production schedules. In automotive industries, for example, materials *must* arrive at a plant at appropriate time intervals (see also *just in time*, Section I.7). Reliability of supply is also important, considering the fact that most countries import numerous raw materials. The United States, for example, imports most of the cobalt, titanium, chromium, aluminum, nickel, natural rubber, and diamond that it needs. A country's self-reliance on resources, especially energy, is an often-expressed political goal, but challenging to achieve. Geopolitics (defined briefly as the study of the influence of a nation's physical geography on its foreign policy) also must thus be a consideration, particularly during periods of global instability or hostility.

Service Life. Everyone has directly experienced a shortened service life of a product, which often can be traced to one or more of the following: (a) improper selection of materials, (b) improper selection of production methods, (c) insufficient control of processing variables, (d) defective raw materials or parts, or manufacturing-induced defects, (e) poor maintenance of machinery and equipment, and (f) improper use of the product.

Generally, a product is considered to have failed when it

- stops functioning, due to the failure of one or more of its components, such as a broken shaft, gear, turbine blade, or a burned-out electric motor
- does not function properly or perform within its required specifications, due, for example, to worn gears or bearings
- becomes unreliable or unsafe for further use, as in the erratic behavior of a switch, poor connections in a printed-circuit board, or delamination of a composite material.

Material Substitution in Products. For a variety of reasons, numerous substitutions are often made in materials, as evidenced by a routine inspection and comparison of common products, such as home appliances, sports equipment, and automobiles. As a measure of the challenges faced in material substitution, consider the following examples: (a) metal vs. wooden handle for a hammer, (b) aluminum vs. cast-iron lawn chair, (c) copper vs. aluminum electrical wire, and (d) alloy steel vs. titanium submarine hull.

The following two case studies describe some details of the major factors involved in material substitution in common products.

Case Study I.2 U.S. Pennies

Billions of pennies are produced and put into circulation each year by the U.S. Mint. The materials used have undergone significant changes throughout their history, largely because of periodic material shortages and the resulting fluctuating cost of appropriate raw materials. The following table shows the chronological development of material substitutions in pennies:

1793–1837	100% copper
1837–1857	95% copper, 5% tin and zinc
1857–1863	88% copper, 12% nickel
1864–1962	95% copper, 5% tin and zinc
1943 (WW II years)	Steel, plated with zinc
1962–1982	95% copper, 5% zinc
1982–present	97.5% zinc, plated with copper

I.6 Selection of Manufacturing Processes

There is often more than one method that can be employed to produce a part from a given material. The following broad categories of manufacturing methods are all applicable for metallic as well as nonmetallic materials:

1. Casting (Fig. I.7a): Expendable mold and permanent mold (Part II).
2. Forming and shaping (Figs. I.7b through I.7d): Rolling, forging, extrusion, drawing, sheet forming, powder metallurgy, and molding (Part III).
3. Machining (Fig. I.7e): Turning, boring, drilling, milling, planing, shaping, broaching; grinding; ultrasonic machining; chemical, electrical, and electrochemical machining; and high-energy-beam machining (Part IV). This broad category also includes micromachining for producing ultraprecision parts (Part V).
4. Joining (Fig. I.7f): Welding, brazing, soldering, diffusion bonding, adhesive bonding, and mechanical joining (Part VI).
5. Finishing: Honing, lapping, polishing, burnishing, deburring, surface treating, coating, and plating (Chapters 26 and 34).
6. Microfabrication and nanofabrication: Technologies that are capable of producing parts with dimensions at the micro (one-millionth of a meter) and nano (one-billionth of a meter) levels; fabrication of microelectromechanical systems (MEMS) and nanoelectromechanical systems (NEMS), typically involving processes such as lithography, micromachining, etching, LIGA, and various specialized processes (Chapters 28 and 29).

Process Selection. The selection of a particular manufacturing process or, more often, sequence of processes, depends on the geometric features of the parts to be made, including the dimensional tolerances

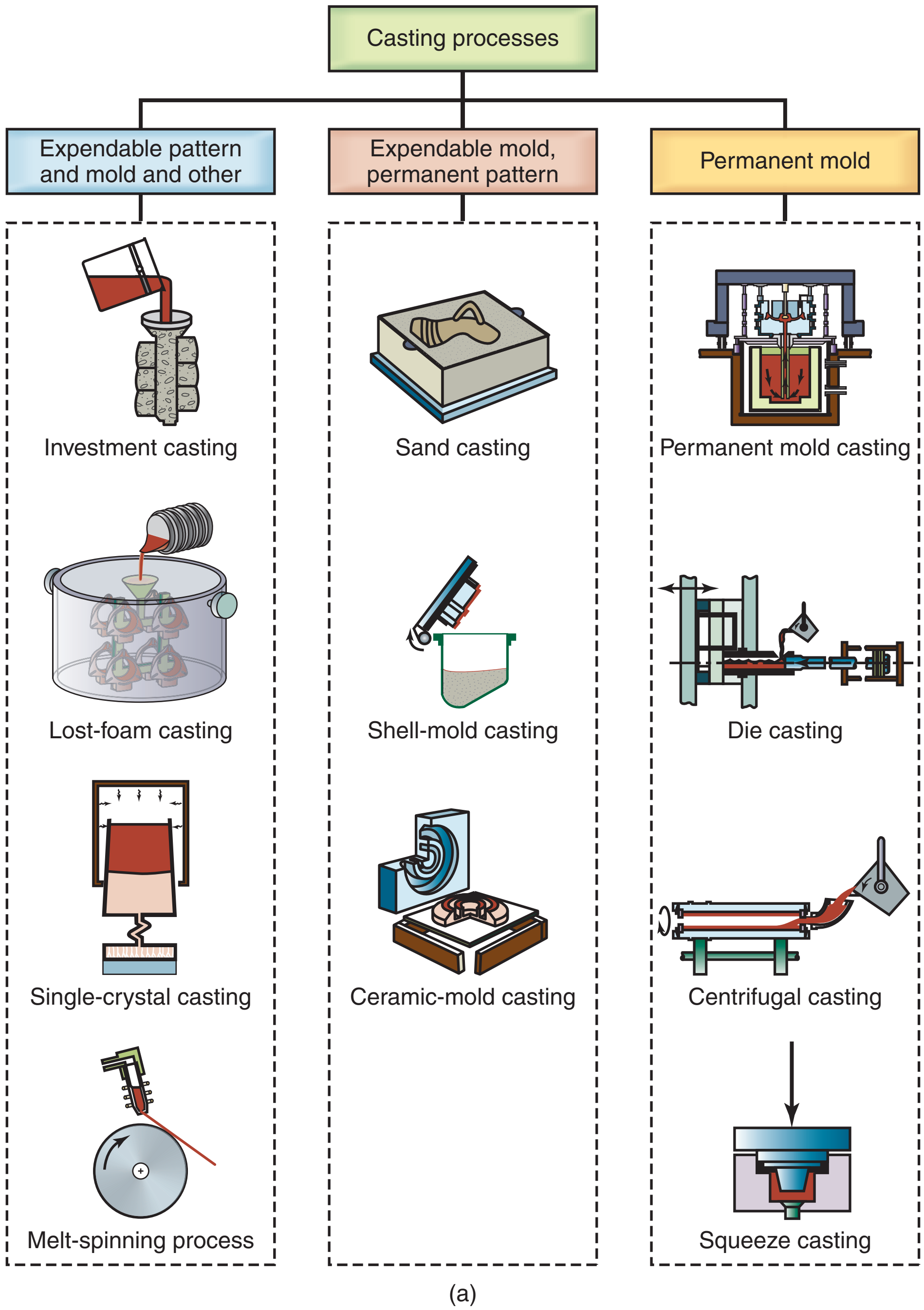


Figure I.7: (a) Schematic illustrations of various casting processes.

and surface texture required, and on numerous factors pertaining to the particular workpiece material and its manufacturing properties. To emphasize the challenges involved, consider the following two cases:

1. Brittle and hard materials cannot be shaped or formed without the risk of fracture, unless they performed at elevated temperatures, whereas these materials can be cast, machined, or ground with relative ease.

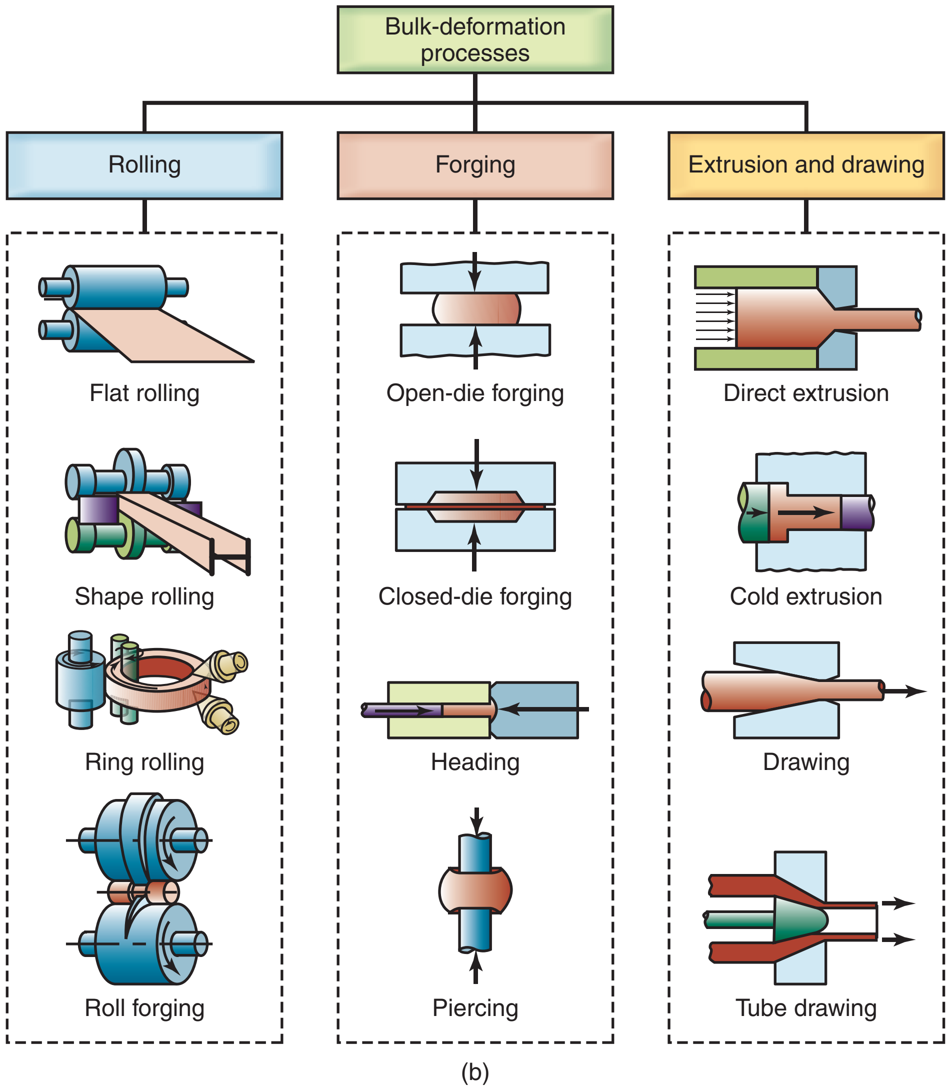


Figure I.7: (b) Schematic illustrations of various bulk-deformation processes.

2. Metals that have been preshaped at room temperature become less formable during subsequent processing, which, in practice, is often necessary to complete the part. This is because the metals have become stronger, harder, and less ductile than they were prior to processing them further.

There is a constant demand for new approaches to production challenges and, especially, for manufacturing cost reduction. For example, sheet-metal parts traditionally have been cut and fabricated using common mechanical tools, such as punches and dies. Although still widely used, some of these operations have been replaced by laser cutting (Fig. I.8). This method eliminates the need for hard tools, which typically have only fixed shapes, and can be expensive and time consuming to make.

The laser path in this operation is computer controlled, thereby increasing the operation's flexibility and its capability for accurately producing an infinite variety of shapes, repeatedly, and economically. Because of the high heat involved in using lasers, however, the surfaces produced have very different characteristics (such as texture and discoloration) than those produced by traditional methods. This difference can have

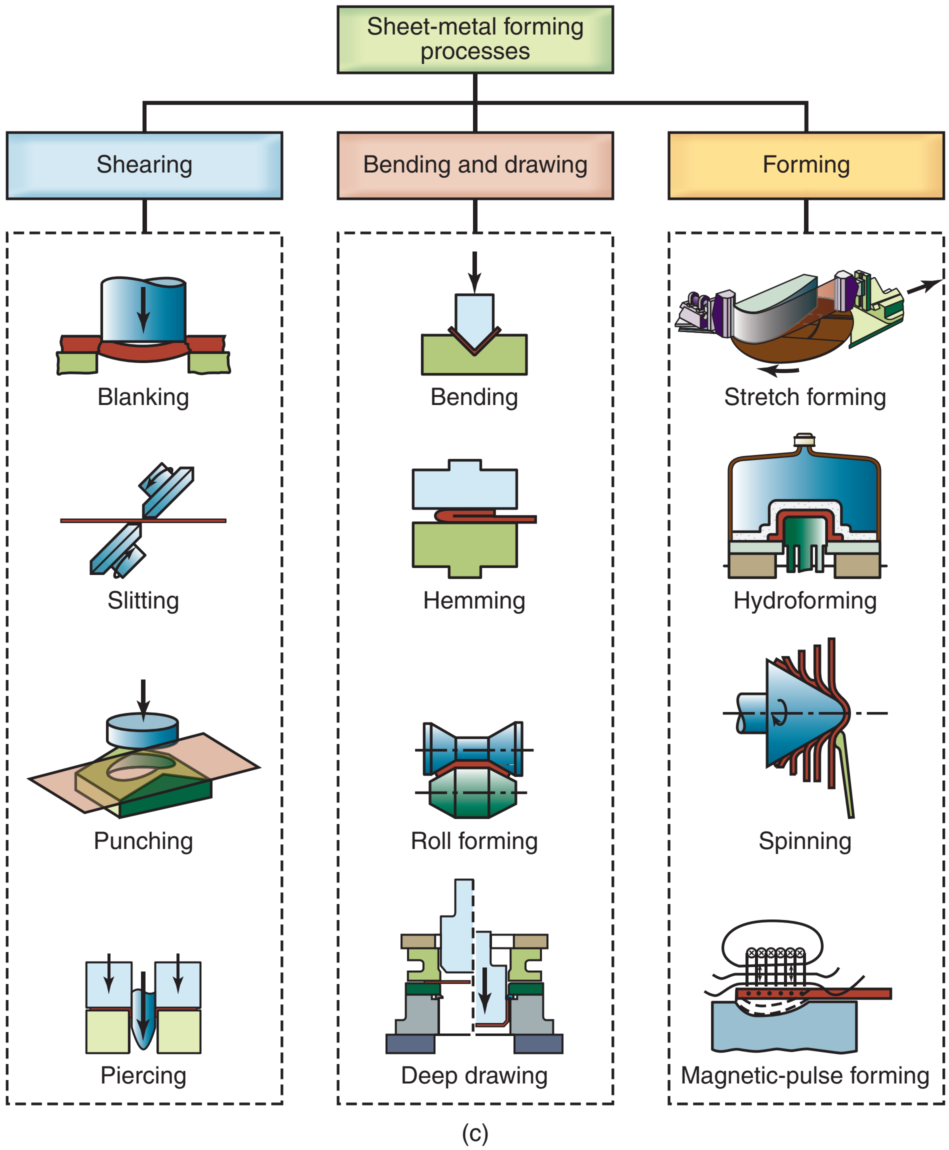


Figure I.7: (c) Schematic illustrations of various sheet-metal forming processes.

significant adverse effects, not only on appearance, but especially on its subsequent processing and in the service life of the product. Moreover, the inherent flexibility of the laser cutting process is countered by the fact that it is slower than traditional punching operations.

Several factors can have a major role in process selection, such as part size, shape complexity, and dimensional accuracy and surface finish required. For example:

- Flat parts and thin cross sections can be difficult to cast.
- Complex parts generally cannot be shaped easily and economically by such metalworking techniques as forging, whereas, depending on part size and level of complexity, the parts may be precision cast, fabricated and assembled from individual pieces, or produced by powder-metallurgy techniques.

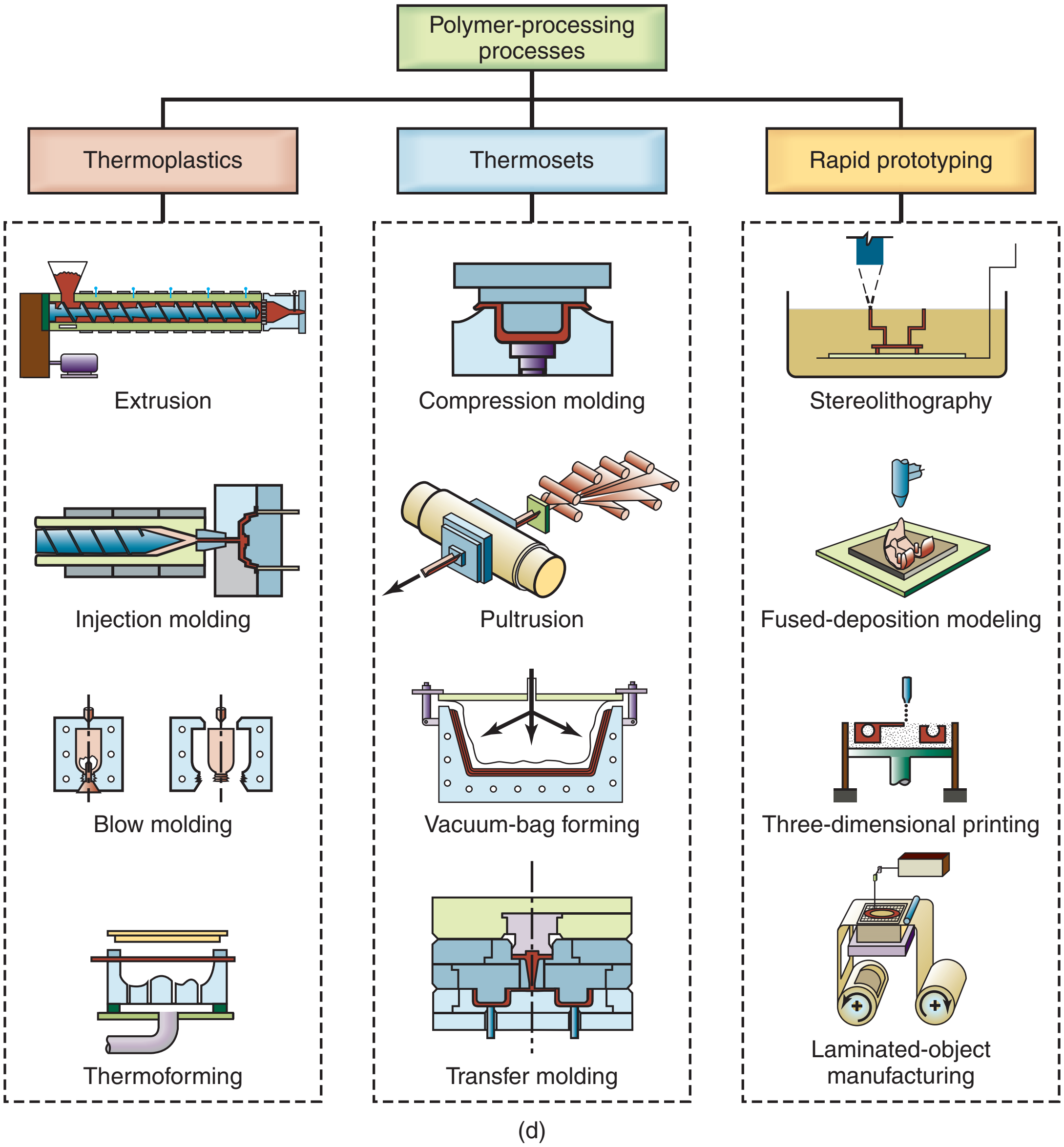


Figure I.7: (d) Schematic illustrations of various polymer-processing methods.

- Dimensional tolerances and surface finish in hot-working operations are not as fine as those obtained in operations performed at room temperature (cold working), because of the dimensional changes, distortion, warping, and surface oxidation due to elevated temperatures.

Part size and dimensional accuracy. The size, thickness, and shape complexity of a part have a major bearing on the process selected. Complex parts, for example, may not be formed easily and economically, whereas they may be produced by casting, injection molding, and powder metallurgy, or they may be fabricated and assembled from individual pieces. Likewise, flat parts with thin cross sections may not be cast easily. Dimensional tolerances and surface finish in hot-working operations cannot be as fine as those in cold-working operations, because dimensional changes, warping, and surface oxidation occur during processing at elevated temperatures. Also, some casting processes produce a better surface finish than others,

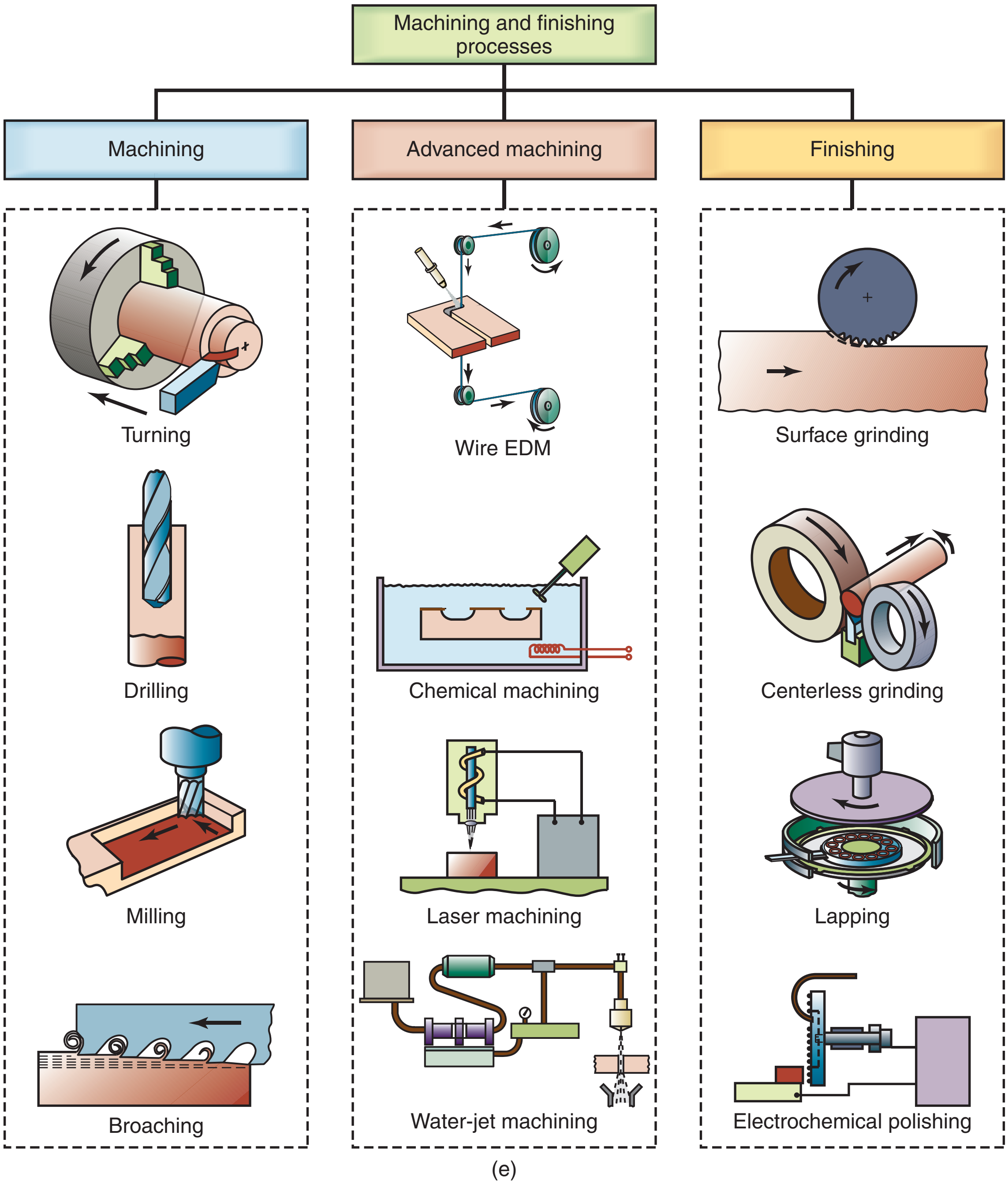


Figure I.7: (e) Schematic illustrations of various machining and finishing processes.

because of the different types of mold materials used. Moreover, the appearance of materials after they have been manufactured into products greatly influences their appeal to the consumer; color, surface texture, and feel are characteristics typically are considered when making a purchasing decision.

The size and shape of manufactured products vary widely. The main landing gear for the twin-engine, 400-passenger Boeing 777 jetliner, for example, is 4.3 m tall, and has three axles and six wheels. The main structure of the landing gear is made by forging, followed by several machining operations

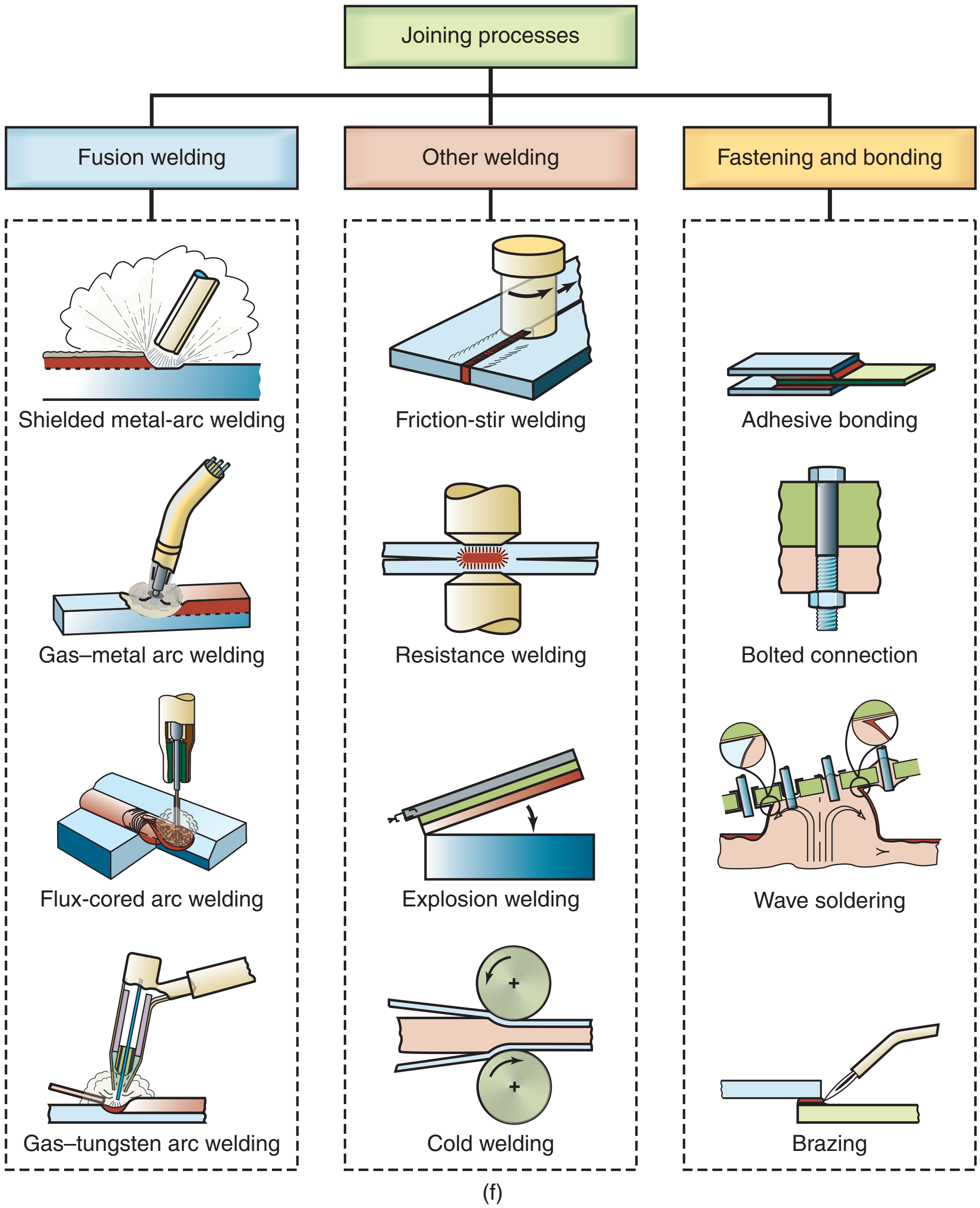


Figure I.7: (f) Schematic illustrations of various joining processes.

(Chapters 6, 8, and 9). At the other extreme is manufacturing microscopic parts and mechanisms. These components are produced through surface micromachining operations, typically using electron beam, laser beam, and wet and dry etching techniques, on materials such as silicon.

Process Substitution. It is common practice in industry that, for a variety of reasons and after a review of all appropriate and applicable processes, a particular production method, which may have been employed in the past, may well have to be substituted with another method. Consider, for example, the following

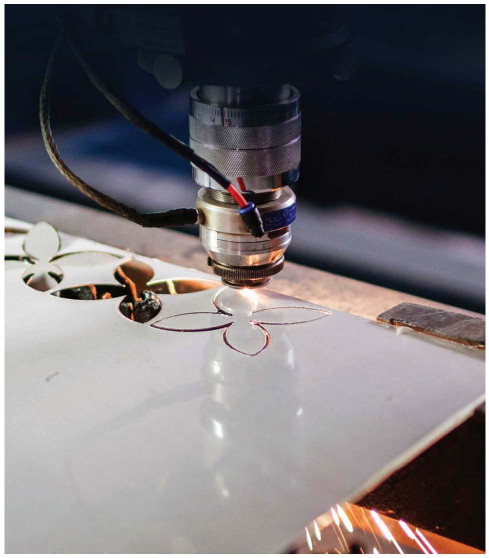


Figure I.8: Cutting sheet metal with a laser beam. *Source:* Courtesy of SPI Lasers UK Ltd.

products which can be produced by any of the sets of the following processes: (a) Cast vs. forged crankshaft, (b) stamped sheet-metal vs. forged or cast automobile wheels, (c) cast vs. stamped sheet-metal frying pan, (d) selective laser sintered vs. extruded or cast polymer bracket, and (e) welded vs. riveted sheet-metal safety hood for a machine.

Criteria for process selection include such factors as cost, maintenance required, whether the product is for industrial or consumer use, parameters to which the product will be subjected (such as external forces, impact, temperatures, and chemicals), environmental concerns that have to be addressed, and the product's appeal to the customer.

Net-shape and Near-net-shape Manufacturing. *Net-shape* and *near-net-shape manufacturing* together constitute an important methodology, by which a part is made in only one operation at or close to the final specific dimensions, tolerances, and surface finish. The difference between net shape and near net shape is a matter of degree, of how close the product is to its final dimensional surface-finish characteristics.

The necessity for and benefits of net-shape manufacturing can be appreciated from the fact that, in the majority of cases, more than one additional manufacturing operation or step is often necessary to produce the part. For example, a cast or forged crankshaft generally will not have the necessary dimensional surface finish characteristics, and will typically require additional processing, such as machining or grinding. These additional operations can contribute significantly to the cost of a product.

Typical examples of net-shape manufacturing include precision casting (Chapter 11), forging (Chapter 14), forming sheet metal (Chapter 16), powder metallurgy and injection molding of metal powders (Chapter 17), and injection molding of plastics (Chapter 19).

Ultraprecision Manufacturing. Dimensional accuracies for some modern equipment and instrumentation have now reached the magnitude of the atomic lattice (below 1 nm). Several techniques, including the use of highly sophisticated technologies (see *micromechanical and microelectromechanical device fabrication*, Chapter 29), are rapidly being developed to attain extreme accuracy. Also, mirror-like surfaces on metals can now be produced by machining, using a very sharp diamond with a nose radius of 250 μm as the cutting tool. The equipment is highly specialized, with very high stiffness, to minimize deflections, as well as vibration and chatter, during machining. It is operated in a room where the ambient temperature is controlled to within 1°C, in order to avoid thermal distortions of the machine.

Types of Production. The number of parts to be produced (such as the annual quantity) and the rate (the number of pieces made per unit time) are important economic considerations in determining optimum processes and types of machinery required. Note, for example, that beverage cans, door locks, and spark plugs are produced in numbers and at rates that are much higher than those for jet engines ship propellers.

A brief outline of the general types of production, in increasing order of annual quantities produced, are:

1. **Job shops**: Small lot sizes, typically less than 100, using general-purpose machines, such as lathes, milling machines, drill presses, and grinders, many now typically equipped with computer controls.

2. **Small-batch production**: Quantities from about 10 to 100, using machines similar to those in job shops.

3. **Batch production**: Lot sizes typically between 100 and 5000, using more advanced machinery with computer control.

4. **Mass production**: Lot sizes generally over 100,000, using special-purpose machinery, known as dedicated machines, and various automated equipment in a plant for transferring materials and parts in progress.

Case Study I.3 Saltshaker and Pepper Mill

The saltshaker and pepper mill set shown in Fig. I.9 consists of metallic and nonmetallic components. The main parts (the body) of the set are made by injection molding of a thermoplastic (Chapter 19), such as an acrylic, which has both transparency and other characteristics for this particular application and also is easy to mold. The round metal top of the saltshaker is made of sheet metal, has punched holes (Chapter 16), and is electroplated for improved appearance corrosion resistance (Section 34.9).

The knob on the top of the pepper mill is made by machining (Chapter 23) and is threaded on the inside to allow it to be screwed and unscrewed. The square rod connecting the top portion of the pepper mill to the two pieces shown at the bottom of the figure is made by rolling (Chapter 13). The two grinder components, shown at the bottom of the figure, are made of stainless steel. A design for manufacturing analysis indicated that casting or machining the two components would be too costly; consequently, it was determined that an appropriate and economical method would be the powder-metallurgy technique (Chapter 17).

I.7 Computer-integrated Manufacturing

Computer-integrated manufacturing (CIM), as the name suggests, integrates the software and the hardware needed for computer graphics, computer-aided modeling, and computer-aided design and manufacturing activities, from initial product concept through its production and distribution in the marketplace. This comprehensive and integrated approach began in the 1970s, and has been particularly effective because of its capability of making possible the following tasks:

- Responsiveness to rapid changes in product design modifications and to varying market demands
- Better use of materials, machinery, and personnel
- Reduction in inventory
- Better control of production and management of the total manufacturing operation.

Figure I.9: A saltshaker and pepper mill set. The two metal pieces (at the bottom) for the pepper mill are made by powder-metallurgy techniques. *Source:* Courtesy of the Metal Powder Industries Federation.

The following is a brief outline of the various elements in CIM, all described in detail in Chapters 38 and 39:

1. **Computer numerical control** (CNC). First implemented in the early 1950s, this is a method of controlling the movements of machine components by direct insertion of coded instructions in the form of numerical data.

2. **Adaptive control** (AC). The processing parameters in an operation are automatically adjusted to optimize the production rate and product quality and to minimize manufacturing costs. For example, in machining, forces, temperature, surface finish, and dimensions of the part are constantly monitored. If they move outside the specified range, the system automatically adjusts the relevant variables until all the parameters are within the specified range.

3. **Industrial robots.** Introduced in the early 1960s, industrial robots have rapidly been replacing humans, especially in operations that are repetitive, dangerous, and boring. As a result, variability in product quality decreases and productivity is improved. Robots are particularly effective in assembly operations; *intelligent robots* have been developed with *sensory perception* capabilities and movements that simulate those of humans. Recent innovations involve *cobots*, which are designed to work and interact with humans.

4. **Automated materials handling.** Computers have made possible highly efficient handling of materials and parts in various stages of completion (*work in progress*), as in moving a part from one machine to another, and then to points of inspection, to inventory, and, finally, to shipment.

5. **Automated assembly systems.** These systems have been developed to replace assembly by human operators, although humans still have to perform some of the operations. Depending on the type of product, assembly costs can be high; thus, products must be designed such that they can be assembled more easily and faster by automated machinery.

6. **Computer-aided process planning** (CAPP). By optimizing process planning, this system is capable of improving productivity, product quality, and consistency, thus reducing costs. Functions such as cost estimating and monitoring work standards (time required to perform a certain operation) are also incorporated into the system.

7. **Group technology** (GT). The concept behind group technology is that numerous parts can be grouped and produced by classifying them into *families* according to similarities in (a) design and (b) the manufacturing processes employed to produce them. In this way, part designs and processing plans can be standardized, and new parts, based on similar parts made previously, can be produced efficiently and economically.

8. **Just-in-time production** (JIT). The principle behind JIT is that (a) supplies of raw materials and parts are delivered to the manufacturer just in time to be used, (b) parts and components are produced just in time to be made into subassemblies, and (c) products are assembled and finished just in time to be delivered to the customer. As a result, inventory carrying costs are minimal, defects in components are detected right away, productivity is increased, and high-quality products are made and at low cost.

9. **Cellular manufacturing** (CM). This system utilizes workstations that consist of a number of *manufacturing cells*, each containing various production machines, all controlled by a central robot, with each machine performing a specific operation on the part, including inspection (Fig. I.10).

10. **Flexible manufacturing systems** (FMS). These systems integrate manufacturing cells into a large production facility, in which all cells are interfaced with a central computer. Although very costly, flexible manufacturing systems are capable of producing parts efficiently (although in relatively small quantities, because hard automation is still most efficient for mass production) and of quickly changing manufacturing sequences required for making different types of parts. Flexibility enables these systems to meet rapid changes in market demand for all types of products.

11. **Expert systems** (ES). Consisting basically of complex computer programs, these systems have the capability of performing a variety of tasks and solving difficult real-life problems (much as human experts would), including expediting the traditional iterative process.

12. **Artificial intelligence** (AI). Computer-controlled systems are capable of learning from experience and of making decisions that optimize operations and minimize costs, ultimately replacing human intelligence.

Figure I.10: Robotic arm production line. *Source:* Shutterstock/Andrey Armyagov.

13. **Artificial neural networks** (ANN). These networks are designed to simulate the thought processes of the human brain, with such capabilities as modeling and simulating production facilities, monitoring and controlling manufacturing operations, diagnosing problems in machine performance, and conducting financial planning and managing a company's manufacturing strategy.

I.8 Quality Assurance and Total Quality Management

Product quality is one of the most critical considerations in manufacturing, because it directly influences customer satisfaction, thus playing a crucial role in determining a product's success in the marketplace (Chapter 36). The traditional approach of inspecting products after they were made has largely been replaced by the recognition that *quality must be built into the product*, from its initial design through all subsequent steps of manufacturing and assembly operations.

Even small products typically undergo several manufacturing steps, and each step involves its own variations in performance, which can occur within a relatively short time. A production machine, for example, may perform differently when it is first turned on than after it begins to warm up or when the ambient temperature in the plant fluctuates. Consequently, *continuous control of processes online monitoring*) is a critical factor in maintaining product quality. The objective is to *control processes, not products*.

Quality assurance and *total quality management* (TQM) are widely recognized as being the responsibility of everyone involved in the design and manufacturing of products and their components. *Product integrity* is a term generally used to define the degree to which a product

- Functions reliably during its life expectancy (Table I.5)
- Is suitable for its intended purposes
- Can be maintained with relative ease.

Producing and marketing defective products can be very costly to the manufacturer, with costs varying by orders of magnitude, as shown in Table I.6.

Table I.5: Average Life Expectancy of Various Products.

Type of product	Life expectancy (years)
U.S. dollar bill	1.5
Personal computer	2
Car battery	4
Hair dryer	5
Automobile	8
Dishwasher	10
Kitchen disposal unit	10
Vacuum cleaner	10
Water heater (gas)	12
Clothes dryer (gas)	13
Clothes washer	13
Air-conditioning unit (central)	15
Manufacturing cell	15
Refrigerator	17
Furnace (gas)	18
Machinery	30
Nuclear reactor	40

Note: Significant variations can be expected, depending on the quality of the product and how well it has been maintained.

Table I.6: Relative Cost of Repair at Various Stages of Product Development and Sale.

Stage	Relative cost of repair
When part is being made	1
Subassembly of the product	10
Assembly of the product	100
Product at the dealership	1000
Product at the customer	10,000

Pioneers in quality control, particularly W.E. Deming (1900–1993), J.M. Juran (1904–2008), and G. Taguchi (1924–2012), all emphasized the importance of management's commitment to (a) product quality, (b) pride of workmanship at all levels of production, and (c) the necessity of using **statistical process control** (SPC) and **control charts** (Chapter 36). They also pointed out the importance of *online monitoring* and rapidly identifying the *sources of quality problems* in production, before even another defective part is produced. The major goal of control is to *prevent* defective parts from ever being made, rather than to inspect, detect, and reject defective parts *after* they have been made.

As an example of strict quality control, computer chips are now produced with such high quality that only a few out of a million chips may be defective. The level of defects is identified in terms of **standard deviation**, denoted by the symbol σ (the Greek letter *sigma*). Three sigma would result in 2700 defective parts per million, which is unacceptable in modern manufacturing. In fact, it has been estimated that at this level, no modern computer would function reliably. At **six sigma**, defective parts are reduced to only 3.4 per million parts made. This level has been reached through major improvements in manufacturing *process capabilities* in order to *reduce variability* in product quality.

Important developments in quality assurance include the implementation of **experimental design**, a technique by which the factors involved in a manufacturing operation and their interactions are studied simultaneously. For example, the variables affecting dimensional accuracy or surface finish in a machining operation can readily be identified, thus making it possible for appropriate *on-time preventive adjustments* to be taken.

Quality Standards. Global manufacturing and competitiveness have led to an obvious need for international conformity and consensus in establishing quality control methods. This need resulted in the establishment of the ISO 9000 standards series on quality management and quality assurance standards, as well as of the QS 9000 standards (Section 36.6), introduced in 1994. A company's registration for these standards, which is a *quality process certification* and not a product certification, means that the company conforms to consistent practices as specified by *its own* quality system. ISO 9000 and QS 9000 have permanently influenced the manner in which companies conduct business in world trade.

Human-factors Engineering. This topic deals with human–machine interactions, and thus it is an important aspect of manufacturing operations in a plant, as well as of products in their expected use. The human-factors approach is essential in the design and manufacture of safe products. It emphasizes **ergonomics**, defined as the study of how a workplace and the machinery and equipment in it, can best be designed and arranged for comfort, safety, efficiency, and productivity.

Examples of the need for proper ergonomic considerations are the following: (a) a mechanism that is difficult to operate manually, causing injury to the employee; (b) a poorly designed keyboard that causes pain to the user's hands and arms during its normal use (*repetitive stress syndrome*); and (c) a control panel on a machine which is difficult to reach or use safely and comfortably.

Product Liability. Designing and manufacturing safe products is an essential responsibility of the manufacturer. All those involved with product design, manufacture, and marketing must fully recognize the consequences of a product's failure, including failure due to foreseeable misuse of the product.

A product's malfunction or failure can cause bodily injury or even death, as well as financial loss to an individual, a bystander, or an organization. Known as *product liability*, the laws governing it vary from state to state and from country to country. Among numerous examples of products that could involve liability are the following:

- A grinding wheel that shatters and causes injury to a worker.
- A cable supporting a platform snaps, allowing the platform to drop, causing bodily harm or death.
- Automotive brakes that suddenly become inoperative, because of the failure of a particular component of the braking system.
- Production machinery that lacks appropriate safety guards.
- Electric and pneumatic tools that lack appropriate warnings and instructions for their safe use.

I.9 Lean Production and Agile Manufacturing

Lean production (Section 39.7) is a methodology that involves thorough assessment of each activity of a company. Its basic purpose is to minimize waste at *all* levels, and calling for the elimination of unnecessary operations that do not provide any added value to the product being made. This approach, also called *lean manufacturing*, identifies all of a manufacturer's activities and optimizes the processes used in order to *maximize added value*.

Lean production focuses on (a) the efficiency and effectiveness of each and every manufacturing step, (b) the efficiency of the machinery and equipment used, and (c) the activities of the personnel involved in each operation. This methodology also includes a comprehensive analysis of the *costs* incurred in each activity and the costs of productive and for nonproductive labor.

The lean production strategy requires a fundamental change in corporate culture, as well as having an understanding of the importance of *cooperation and teamwork* among a company's workforce and management. Lean production does not necessarily require cutting back on a company's physical or human resources. It aims at *continually* improving efficiency and profitability by removing all waste in the company's operations and dealing with any problems as soon as they arise.

I.10 Manufacturing Costs and Global Competition

Always critically important, economics of manufacturing has become even more so with (a) ever-increasing global competition and (b) the demand for high-quality products, generally referred to as *world-class manufacturing*, and at low prices. Typically, the *manufacturing cost* of a product represents about 40% of its *selling price*, which often is the overriding consideration in a product's marketability and general customer satisfaction. An approximate, but typical, breakdown of costs in modern manufacturing is given in Table I.7. As to be expected, the percentages indicated can vary significantly depending on product type.

Table I.7: Typical Cost Breakdown in Manufacturing.

Design	5%
Materials	50%
Manufacturing	
Direct labor	15%
Indirect labor	30%

The *total cost* of manufacturing a product generally consists of the following components:

1. **Materials.** Raw-material costs depend on the material itself and on supply and demand for that material. Low cost may not be the deciding factor if the cost of processing a particular material is higher than that for a more expensive material. For example, a low-cost piece of metal may require more time to machine or to shape than one of higher cost.

2. **Tooling.** Tooling costs include those for cutting tools, dies, molds, workholding devices, and fixtures. Some cutting tools cost as little as $2, others as much as $100 for cubic boron nitride and diamond. Depending on their size and the materials involved in making them, molds and dies can cost from only a few hundred dollars to over $2 million for a set of dies for stamping large sheet metal parts.

3. **Fixed.** Fixed costs include costs for energy, rent for facilities, insurance, and real-estate taxes.

4. **Capital.** Production machinery, equipment, buildings, and land are typical capital costs. Machinery costs can range from a few hundred to millions of dollars. Although the cost of computer-controlled machinery can be very high, such an expenditure may well be warranted if it reduces labor costs.

5. **Labor.** Labor costs consist of direct and indirect costs. *Direct labor*, also called *productive labor*, concerns the labor that is directly involved in manufacturing products. *Indirect labor*, also called *nonproductive labor* or *overhead*, pertains to servicing of the total manufacturing operation.

Direct-labor costs may be only 10% to 15% of the total cost (Table I.7), but it can be as much as 60% for labor-intensive products, such as clothing and other textiles, and products assembled from components such as toys and musical instruments. Reductions in the direct-labor share of manufacturing costs can be achieved by such means as extensive use of automation, computer control of all aspects of manufacturing, implementation of modern technologies, and increased efficiency of operations.

As expected and as shown in Table I.8, there continues to be a worldwide disparity in labor costs, by an order of magnitude. Today, numerous consumer products are manufactured or assembled in the Pacific Rim countries, especially China. Likewise, software and information technologies are often much less costly to develop in such countries as India and China than in the United States or Europe. As living standards continue to rise, however, labor costs, too, are beginning to rise significantly in these countries.

Table I.8: Approximate Relative Hourly Compensation for Workers in Manufacturing in 2010 (United States = 100).

Norway	166	Italy	96
Switzerland	153	Japan	92
Belgium	146	Spain	76
Denmark	131	New Zealand	59
Germany	126	Israel	58
Sweden	126	Singapore	55
Finland	122	Korea (South)	48
Austria	118	Argentina, Slovakia	36
Netherlands, Australia	118	Portugal	34
France,	117	Czech Republic	33
Ireland	104	Poland	23
United States	100	Mexico	18
Canada	97	China, India, Philippines	6

Note: Compensation can vary significantly with benefits. Data for China and India are estimates, use different statistical measures of compensation, and are provided for comparison purposes only. *Source:* U.S. Department of Labor.

Outsourcing. *Outsourcing*, defined as the purchase by a company of parts and/or labor from an outside source, either from another company or another country, in order to reduce design and manufacturing costs. In theory, this approach allows companies to concentrate more on their core competencies, and be able to optimize their critical technologies. Outsourcing, however, has several drawbacks, including its social impact and political implications of any ensuing lowered employment, especially in the European Union countries and the United States. In recent years, the costs of shipping and transport have increased and have become more uncertain; also, manufacturers often prefer to be located near their customers and/or suppliers. As a result, a **reshoring** trend has been observed, which involves relocating manufacturing activities to a few critical locations, usually near the customers.

I.11 Trends in Manufacturing

Several trends regarding various aspects of modern manufacturing are the following:

- Product *variety* and *complexity* continue to increase.
- Product *life cycles* are becoming shorter.
- Markets continue to become *multinational* and *global competition* has increased rapidly.
- Customers are consistently demanding *high-quality*, *reliability*, and *low-cost* products.
- Developments continue in the quality of materials and their selection, especially for improved *recyclability*.
- Machining is faster and more able to achieve better tolerances, because of innovative *control strategies* and suppression of chatter.
- The most *economical* and *environmentally friendly* manufacturing methods are being increasingly pursued, and *energy management* has become increasingly important.
- *Weight savings* continue with the use of materials with higher strength-to-weight and stiffness-to-weight ratios, particularly in the automotive, aerospace, and sporting industries. These materials include fiber reinforced composites as well as advanced metals such as aluminum-lithium alloys and hot stamped steels.
- Titanium, magnesium, aluminum and fiber-reinforced polymers are increasingly seen as essential technologies for meeting *fuel energy efficiency* goals in transportation applications.
- Improvements are being made in *predictive models* of the effects of material-processing parameters on product integrity, applied during a product's design stage.
- Developments in *ultraprecision manufacturing*, *micromanufacturing*, and *nanomanufacturing* continue, so that manufacturing ability is approaching the level of atomic dimensions.
- *Computer simulation*, modeling, and control strategies are being applied to all areas of manufacturing, design, performance and maintenance prediction.
- *Additive manufacturing* has become pervasive, with a wide range of equipment availability and at lower cost. Additive manufacturing technologies are increasingly being applied to the production of tooling and direct digital manufacturing. New additive manufacturing approaches and extension to new materials are continually under development; additive manufacturing machines are becoming faster, more reliable, and inexpensive.
- Advances in *optimization* of manufacturing processes and production systems are making them more reliable.

- *Lean production* and *information technology* are being implemented as powerful tools to help meet global challenges.
- Manufacturing activities are viewed not as individual, separate tasks, but as making up a *large system*, with all its parts being interrelated.
- It has now become common practice to *build quality into the product* at each stage of its production.
- Continued efforts are aimed at achieving higher levels of productivity and *eliminating* or *minimizing waste*, with optimum use of an organization's resources.
- Software continues to expand into *all* aspects of manufacturing, from control of machinery to controlling flow of materials and products to factories (*supply chain management*).
- *Sensors* of all types are being incorporated into machines, providing data for process validation and for historical information that can be stored for future reference. A term used to incorporate computer data into all parts of a product's lifecycle is the *digital thread*.
- *Lean production* and information technology are being implemented as powerful tools to help meet global challenges. Machine tools are increasingly capable of communicating using MTConnect, giving plant managers real-time information about factory floor operations.
- Advances in *communication* and *sensors* will lead to unprecedented access to data (as with cloud-based storage), improving control of the manufacturing enterprise, quality and efficiency, and management of complex global supply chains.

PART I
Fundamentals of Materials: Behavior and Manufacturing Properties

Part I of this text begins by describing the behavior and properties of materials, their manufacturing characteristics, and their applications, as well as their advantages and limitations that influence their selection in the design and manufacture of products.

To emphasize the importance of the topics to be described, review a typical automobile as an example of a common product that utilizes a wide variety of materials (Fig. I.1). These materials were selected not only because they possessed the desired properties and characteristics for the intended function of a specific part but also they were the ones that could be manufactured at the lowest cost.

Steel, for example, was chosen for much of the body because it is strong, relatively easy to shape, and relatively inexpensive. Plastics were chosen for many components because of such characteristics as light weight, resistance to corrosion, availability in a wide variety of colors, and ease of manufacturing into complex shapes and at low cost. Glass was chosen for all the windows because it is transparent, hard (hence scratch resistant), easy to shape, and easy to clean. Numerous similar observations can be made about each component of an automobile, ranging from tiny screws to engine blocks. Fuel efficiency and the need for improved performance have driven the substitution of materials in cars (such as aluminum, magnesium, and plastics) for steel and the use of composite materials for structural (load-bearing) components.

As stated in the General Introduction of this text, material selection for individual parts in a product requires a thorough knowledge and assessment of material properties, specific functions of the part, and manufacturing costs involved. A typical automobile is an assemblage of some 15,000 individual parts; consequently, by saving just one cent on the cost per part, such as by selecting a different material or manufacturing process, the cost of an automobile would be reduced by $150. This task thus becomes very challenging, especially with the ever-increasing variety of materials and manufacturing processes that are now available, as outlined in Fig. 1.2.

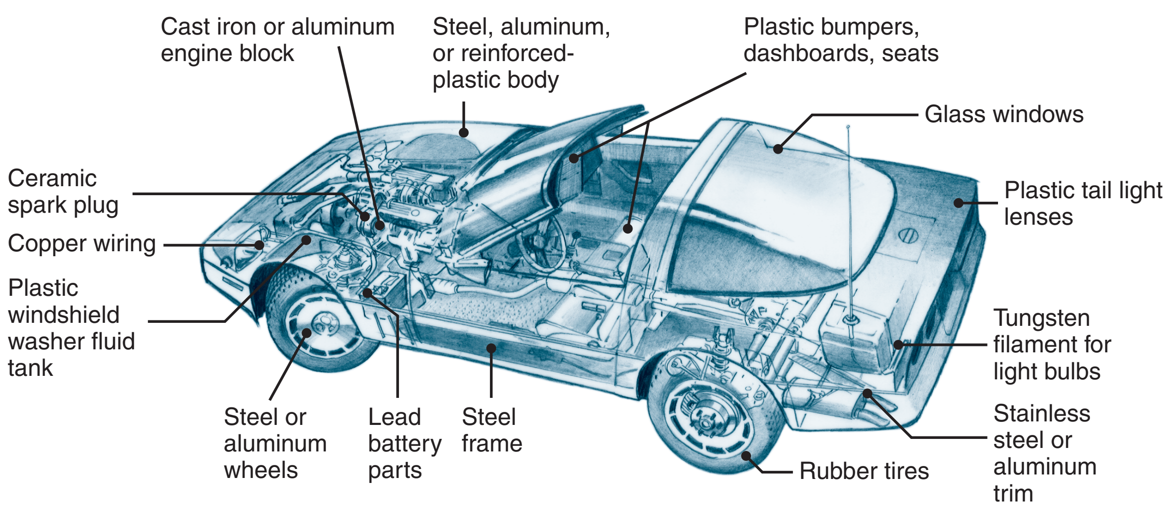


Figure I.1: An outline of the topics described in Part I.

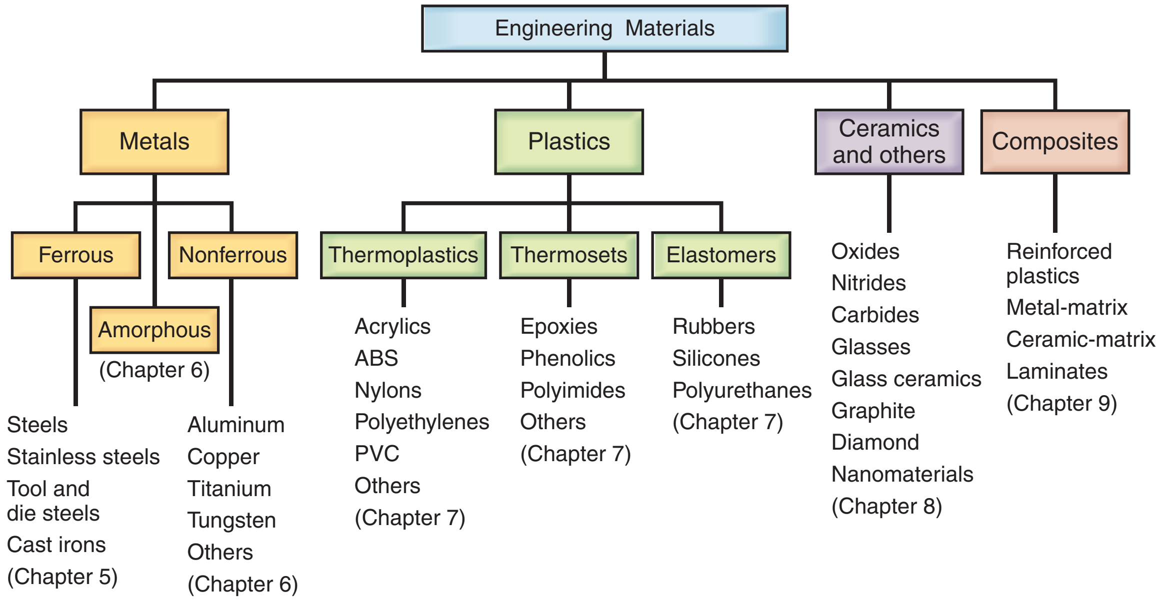


Figure 1.2: An outline of the engineering materials described in Part I.

A general outline of the topics described in Part I of this text is given in Fig. 1.3. The fundamental knowledge presented on the behavior, properties, and characteristics of materials will help the reader understand their relevance to all the manufacturing processes described in Parts II through IV.

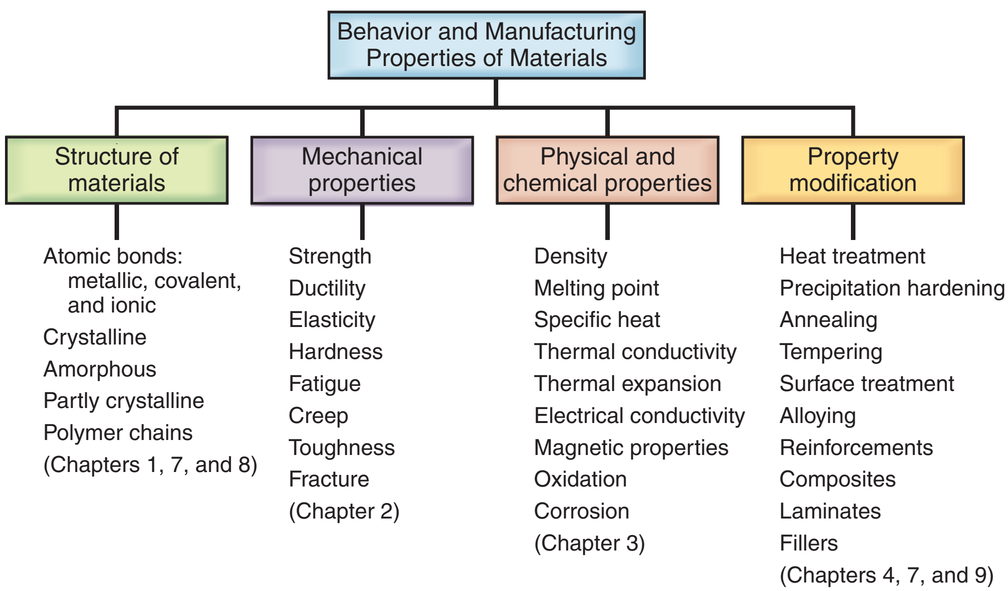


Figure 1.3: An outline of the behavior and the manufacturing properties of materials described in Part I.

Chapter 1

The Structure of Metals

- This chapter describes the crystalline structure of metals and explains how crystal structure determines properties and behavior.
- It begins with a review of the types of atomic bonds and their characteristics: ionic, covalent, and metallic.
- Metal structures and the arrangement of atoms within the structure are then examined; the types of imperfections in the crystal structure and their effects on material behavior are presented.
- The effects of grains and grain boundaries are examined, followed by a description of strain hardening and anisotropy of metals.

1.1 Introduction

Why are some metals hard and others soft? Why are some metals brittle, while others are ductile and can be shaped easily without fracture? Why can some metals withstand high temperatures while others cannot? Why does a piece of sheet metal behave differently when stretched in one direction versus another? These questions can be answered by studying the **atomic structure** of metals—that is, the arrangement of the atoms within the metals. This knowledge then serves as a guide to controlling and predicting the behavior and performance of metals in various manufacturing processes.

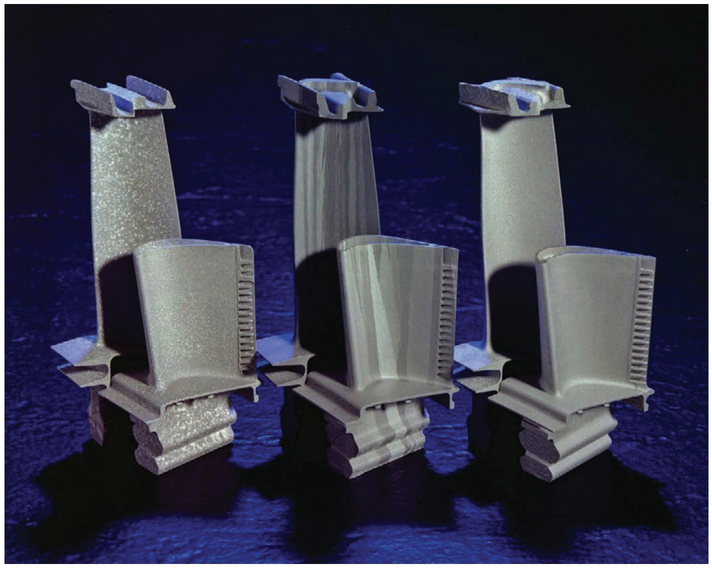


Figure 1.1: Turbine blades for jet engines, manufactured by three different methods: left: conventionally cast; center: directionally solidified, with columnar grains as can be seen from the vertical streaks, and right: single crystal. Although more expensive, single-crystal blades have properties at high temperatures that are superior to those of other blades. *Source:* Courtesy of NASA.

Understanding the structure of metals allows prediction and evaluation of their **properties** such as strength and stiffness, whereby appropriate selection for specific applications can then be made. For example, single-crystal turbine blades (Fig. 1.1) for use in jet engines have properties that are better than those for conventional blades. In addition to their atomic structure, several other factors also influence the properties and behavior of metals. These include the composition of the particular metal, impurities and vacancies in their atomic structure, grain size, grain boundaries, the presence of impurities and inclusions, the environment, surface condition of the metal, and the methods by which they are made into specific products.

The topics described in this chapter are outlined in Fig. 1.2. The structure and general properties of materials other than metals are described in Chapter 7 (on polymers), Chapter 8 (ceramics and glasses), and Chapter 9 (composite materials). The structure of metal alloys, the control of their structure, and heat-treatment processes are described in Chapter 4.

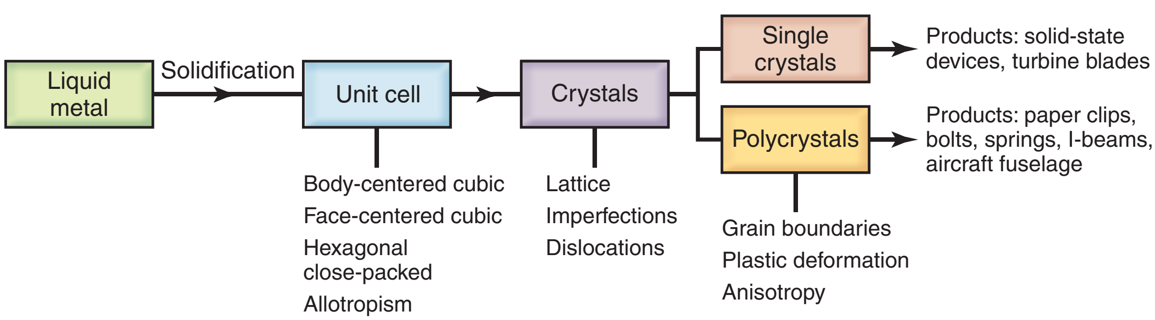


Figure 1.2: An outline of the topics described in Chapter 1.

1.2 Types of Atomic Bonds

All matter is made up of **atoms** consisting of a **nucleus** of *protons* and *neutrons* and surrounding clouds, or orbits, of *electrons*. The number of protons in the nucleus determines whether a particular atom will be metallic, nonmetallic, or semi-metallic. An atom with a balanced charge has the same number of electrons as protons; when there are too many or too few electrons, the atom is called an **ion**. An excess of electrons results in a negatively charged atom, referred to as an **anion**, while too few electrons results in a positively charged atom, called a **cation**. The number of electrons in the outermost orbit of an atom determines the *chemical affinity* of that atom for other atoms.

Atoms can transfer or share electrons; in doing so, multiple atoms can combine to form **molecules**. Molecules are held together by attractive forces called **bonds**, which act through electron interaction. The basic types of atomic attraction associated with electron transfer, called **primary bonds** or **strong bonds**, are the following:

- **Ionic bonds**. When one or more electrons from an outer orbit are transferred from one material to another, a strong attractive force develops between the two ions. An example is that of sodium (Na) and chlorine (Cl) in common table salt; it consists of Na^+ and Cl^- ions (hence the term *ionic bond*), which are strongly attracted to each other. The attraction is between all adjacent ions, allowing crystalline structures to be formed, as described in Section 1.3. Molecules with ionic bonds generally have low ductility and low thermal and electrical conductivity. Ionic bonding is the predominant bond in ceramic materials (Chapter 18).

- **Covalent bonds**. In a covalent bond, the electrons in outer orbits are shared by atoms to form molecules. The number of electrons shared is reflected by terms such as "single bond," "double bond." Polymers (Chapter 7) consist of large molecules covalently bonded together. Water (H_2O) and nitrogen gas (N_2) are common examples of molecules formed by covalent bonds. Solids formed by covalent bonding typically have low electrical conductivity and can have high hardness; diamond, a form of covalently bonded carbon, is an example.

- **Metallic bonds**. Metals have relatively few electrons in their outer orbits, and thus they cannot complete the outer shell when self-mated. Instead, metals and alloys form *metallic bonds*, whereby the available electrons are shared by all atoms in contact. The resultant electron cloud provides the attractive forces to hold the atoms together, resulting in generally high thermal and electrical conductivity.

In addition to the strong attractive forces associated with electrons, weak or secondary attractions exist between molecules; also referred to as **van der Waals** forces, these forces arise from the attraction of opposite charges without electron transfer. Water molecules, for example, consist of one oxygen atom and two smaller hydrogen atoms, located around 104° from each other. Although each molecule has a balanced, or neutral, charge, there are more hydrogen atoms on one side of the molecule (*i.e.*, it is a *dipole*), so that the molecule develops a weak attraction to nearby oxygen atoms on that side.

1.3 The Crystal Structure of Metals

When metals solidify from a molten state (Section 10.2), the atoms arrange themselves into various orderly configurations, called **crystals**; this arrangement is called **crystal structure** or **crystalline structure**. The *smallest* group of atoms exhibiting the characteristic **lattice structure** of a particular metal is known as a **unit cell**.

The three basic atomic arrangements in metals, and some examples of each, are

1. **Body-centered cubic (bcc)**; alpha iron, chromium, molybdenum, tantalum, tungsten, and vanadium.

2. **Face-centered cubic (fcc)**; gamma iron, aluminum, copper, nickel, lead, silver, gold, and platinum.

3. **Hexagonal close-packed (hcp)**; beryllium, cadmium, cobalt, magnesium, alpha titanium, zinc, and zirconium.

These structures are represented by the illustrations given in Figs. 1.3 through 1.5, in which each sphere represents an atom. The distance between the atoms in these crystal structures is on the order of 0.1 nm. The models shown are known as **hard-ball** or **hard-sphere** models; they can be likened to tennis balls arranged in various configurations in a box.

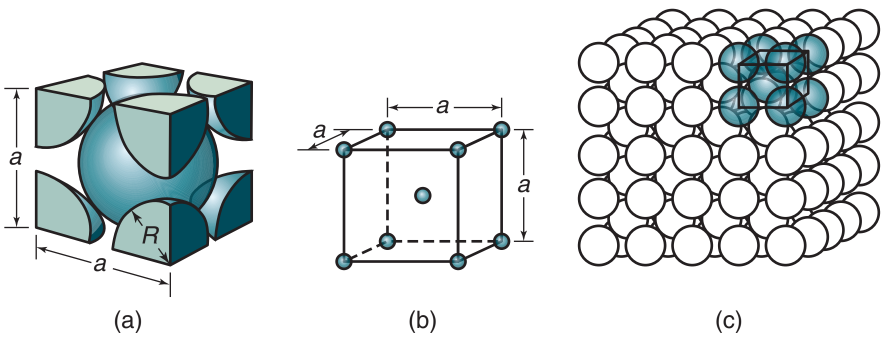


Figure 1.3: The body-centered cubic (bcc) crystal structure: (a) hard-ball model; (b) unit cell; and (c) single crystal with many unit cells.

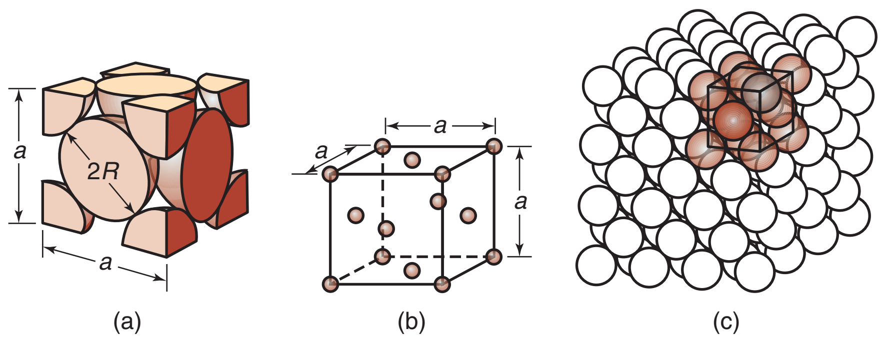


Figure 1.4: The face-centered cubic (fcc) crystal structure: (a) hard-ball model; (b) unit cell; and (c) single crystal with many unit cells.

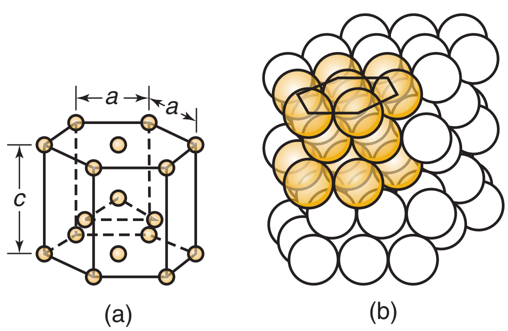


Figure 1.5: The hexagonal close-packed (hcp) crystal structure: (a) unit cell and (b) single crystal with many unit cells.

In the three structures described earlier, the hcp crystals have the most densely packed configurations, followed by fcc and then bcc. In the hcp structure, the top and bottom planes are called **basal planes**. All three arrangements can be modified by adding atoms of some other metal or metals, known as **alloying**, and often improving various properties of the metal.

The presence of more than one type of crystal structure in metals is known as **allotropism** or **polymorphism**, meaning "many shapes." Because the properties and behavior of a particular metal depend greatly on its crystal structure, allotropism is important in the heat treatment of metals and in metalworking and welding operations, described in Parts III and V, respectively. Single crystals of metals are now produced as ingots in sizes on the order of 1.5 m long and up to 300 mm in diameter, with such applications as gas turbine blades and semiconductors (Sections 11.5 and 28.4). Most metals used in manufacturing operations are polycrystalline, as described in Section 1.5.

1.4 Deformation and Strength of Single Crystals

When a single crystal is subjected to an external force, it first undergoes **elastic deformation** (Chapter 2); that is, it returns to its original shape when the force is removed. A simple analogy to this type of behavior is a helical spring that stretches when loaded and returns to its original shape when the load is removed. If the force is increased sufficiently, the crystal undergoes **plastic deformation** or **permanent deformation**; that is, it does not return to its original shape when the force is removed (see also *shape-memory alloys*, Section 6.14).

There are two basic mechanisms by which plastic deformation takes place in crystal structures. One mechanism involves a plane of atoms slipping over an adjacent plane (called the **slip plane**) under a **shear stress** (Fig. 1.6a). This behavior is much like sliding of playing cards against each other. *Shear stress* is defined as the ratio of the applied shearing force to the cross-sectional area being sheared.

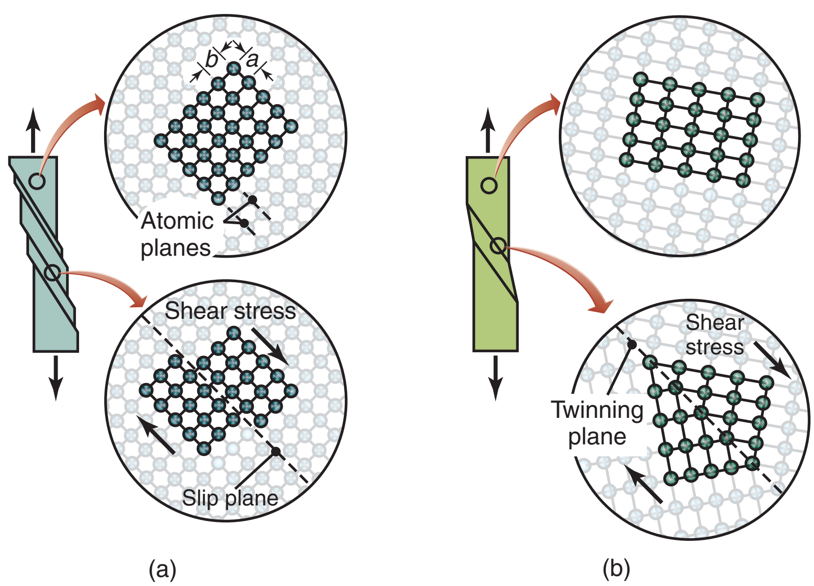


Figure 1.6: Permanent deformation of a single crystal under a tensile load. The highlighted grid of atoms emphasizes the motion that occurs within the lattice. (a) Deformation by slip. The b/a ratio influences the magnitude of the shear stress required to cause slip. (b) Deformation by twinning, involving the generation of a "twin" around a line of symmetry subjected to shear. Note that the tensile load results in a shear stress in the plane illustrated.

Just as it takes a certain force to slide playing cards against each other, a single-crystal metal requires a certain magnitude of shear stress (called **critical shear stress**) to undergo permanent deformation. Thus, there must be a shear stress of sufficient magnitude for plastic deformation to occur; otherwise, the deformation remains elastic.

The shear stress required to cause slip in single crystals is directly proportional to the ratio b/a in Fig. 1.6a, where a is the spacing of the atomic planes and b is inversely proportional to the atomic density in the atomic plane. As the ratio b/a decreases, the shear stress required to cause slip decreases; thus, slip in a single crystal takes place along planes of *maximum atomic density*. In other words, slip takes place in closely packed planes and in closely packed directions. Because the b/a ratio varies for different directions within the crystal, a single crystal exhibits different properties when tested in different directions, a property called **anisotropy**. An example is the behavior of plywood, which is much stronger in the planar direction than it is along its thickness direction.

The second and less common mechanism of plastic deformation in crystals is **twinning**, in which a portion of the crystal forms a mirror image of itself across the *plane of twinning* (Fig. 1.6b). Twins form abruptly and are the cause of the creaking sound (called "tin cry") that occurs when a tin or zinc rod is bent at room temperature; twinning usually occurs in hcp metals.

Slip Systems. The combination of a slip plane and slip direction is known as a *slip system*. Metals with five or more slip systems are generally ductile.

1. In **body-centered cubic** crystals, there are 48 possible slip systems; therefore, the probability is high that an externally applied shear stress will operate on one of these systems and cause slip. Because of the relatively high b/a ratio in this type of crystal, the required shear stress is high. Metals with bcc structures (such as titanium, molybdenum, and tungsten) generally have good strength and moderate ductility, but can have high ductility at elevated temperatures.

2. In **face-centered cubic** crystals, there are 12 slip systems. The probability of slip is moderate, and the shear stress required to cause slip is low because of the relatively low b/a ratio. These metals, such as aluminum, gold, copper, and silver, generally have moderate strength and good ductility.

3. The **hexagonal close-packed** crystal has three slip systems and, therefore, has a low probability of slip; however, additional slip systems become active at elevated temperatures. Metals with hcp structures, such as beryllium, magnesium, and zinc, are generally brittle at room temperature.

Note in Fig. 1.6a that the portions of the single crystal that have undergone slip have rotated from their original angular position toward the direction of the tensile force; note also that slip has taken place only along certain planes. It can be observed from electron microscopy that what appears to be a single slip plane is actually a **slip band**, consisting of several slip planes (Fig. 1.7).

1.4.1 Imperfections in the Crystal Structure of Metals

The actual strength of metals is approximately one to two orders of magnitude lower than the strength levels obtained from theoretical calculations. This discrepancy is explained in terms of **defects** and **imperfections** in the crystal structure. Unlike in idealized models described earlier, actual metal crystals contain a large number of defects and imperfections, generally categorized as:

1. *Point defects*, such as a **vacancy** (missing atom), an **interstitial atom** (extra atom in the lattice), or an **impurity** (foreign atom that has replaced the atom of the pure metal) (Fig. 1.8);

2. *Linear*, or *one-dimensional, defects*, called **dislocations** (Fig. 1.9);

3. *Planar*, or *two-dimensional, imperfections*, such as **grain boundaries** and **phase boundaries** (see Section 1.5);

4. *Volume*, or *bulk, imperfections*, such as **voids, inclusions** (nonmetallic elements, such as oxides, sulfides, and silicates), other **phases**, or **cracks**.

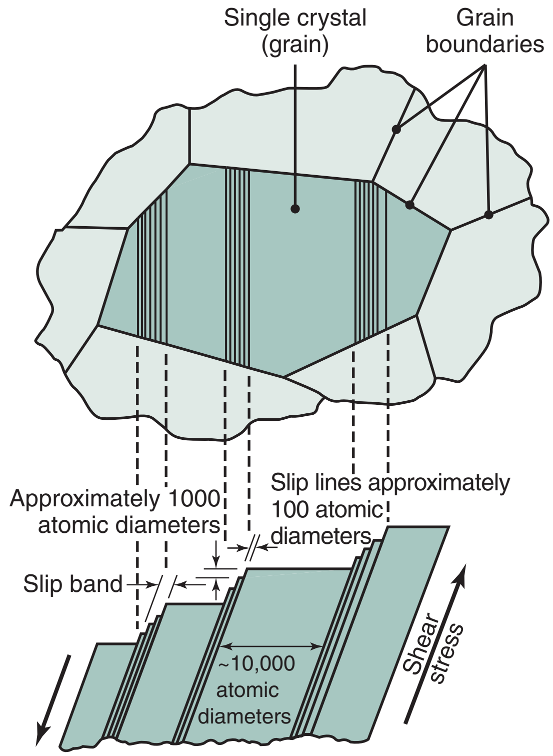


Figure 1.7: Schematic illustration of slip lines and slip bands in a single crystal (grain) subjected to a shear stress. A slip band consists of a number of slip planes. The crystal at the center of the upper illustration is an individual grain surrounded by several other grains.

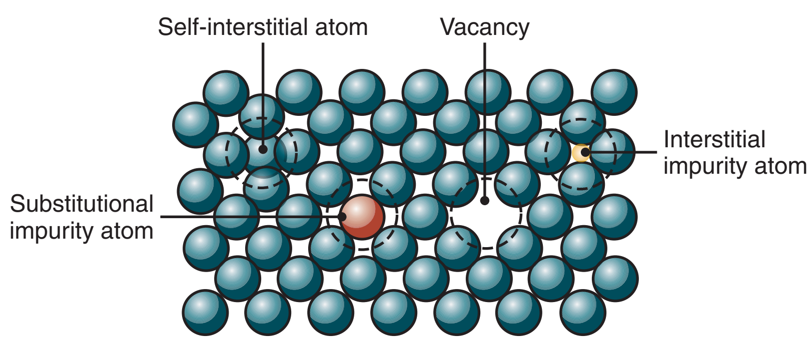


Figure 1.8: Schematic illustration of types of defects in a single-crystal lattice: self-interstitial, vacancy, interstitial, and substitutional.

Mechanical and electrical properties of metals, such as yield stress, fracture strength, and electrical conductivity, are adversely affected by the presence of defects; these properties are known as **structure sensitive**. By contrast, physical and chemical properties, such as melting point, specific heat, coefficient of thermal expansion, and elastic constants such as modulus of elasticity and modulus of rigidity (see Sections 2.2.1 and 2.4) are not sensitive to these defects; these properties are known as **structure insensitive.**

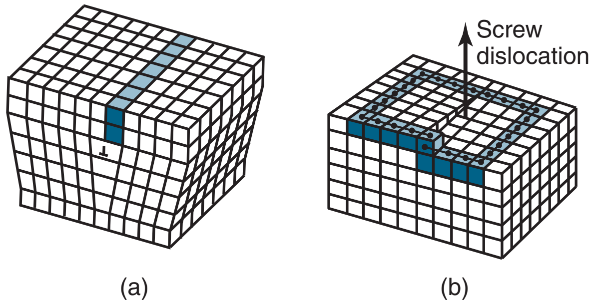


Figure 1.9: Types of dislocations in a single crystal: (a) edge dislocation and (b) screw dislocation.

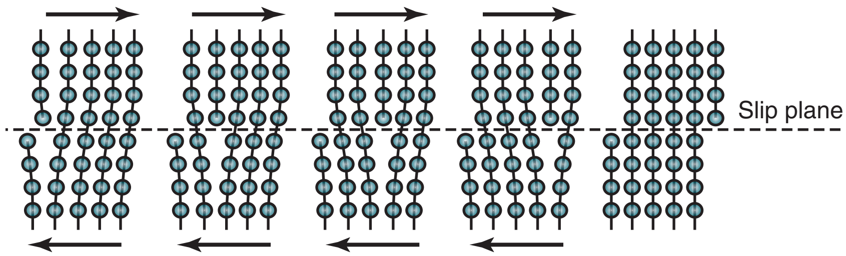


Figure 1.10: Movement of an edge dislocation across the crystal lattice under a shear stress. Dislocations help explain why the actual strength of metals is much lower than that predicted by theory.

Dislocations. First observed in the 1930s, *dislocations* are defects in the orderly arrangement of the atomic structure of a metal. Because a slip plane containing a dislocation (Fig. 1.10) requires much lower shear stress to allow slip than does a plane in a perfect lattice, dislocations are the most significant defects that explain the discrepancy between the actual and the theoretical strengths of metals.

There are two types of dislocations: **edge** and **screw** (Fig. 1.9). An analogy to the movement of an edge dislocation is the progress of an earthworm, which moves forward by means of a hump that starts at its tail and moves toward its head. Another analogy is moving a large carpet on a floor by first forming a hump at one end and gradually moving the hump to the other end. (Recall that the force required to move a carpet in this way is much lower than that required to slide the whole carpet along the floor.) Screw dislocations are so named because the atomic planes form a spiral ramp, like the threads on a screw or bolt.

1.4.2 Work Hardening (Strain Hardening)

Although the presence of a dislocation lowers the shear stress required to cause slip, dislocations can be:

1. Entangled and interfere with each other, and
2. Impeded by barriers, such as grain boundaries, impurities, and inclusions in the material.

The higher shear stress required to overcome entanglements and impediments thus results in an increase in the overall strength and the hardness of the metal, and is known as **work hardening** or **strain hardening**. The greater the deformation, the greater is the number of entanglements and, hence, the higher the increase in strength. Work hardening is a mechanism for strengthening of metals in metalworking processes at low to moderate temperatures. Typical examples are producing sheet metal for appliances and aircraft fuselages by *cold rolling* (Chapter 13), producing the head of a bolt by *heading* (Chapter 14), and strengthening wire by *drawing* it through a die at room temperature (Chapter 15).

1.5 Grains and Grain Boundaries

When a mass of molten metal begins to solidify, crystals form independently of each other at various locations within the liquid mass, and thus have *random* and unrelated orientations (Fig. 1.11). Each of these crystals eventually grows into a crystalline structure, or **grain**. Each grain consists of either a single crystal (for pure metals) or a polycrystalline aggregate (for alloys).

The number and size of the grains developed in a unit volume of the metal depends on the *rate* at which **nucleation** (the initial stage of crystal formation) takes place. The *median size* of the grains depends on (a) the number of different sites at which individual crystals begin to form (note that there are seven in Fig. 1.11a) and (b) the rate at which these crystals grow. If the nucleation rate is high, the number of grains in a unit volume of metal will be large, and thus grain size will be small. Conversely, if the crystal growth rate is high (as compared with their nucleation rate), there will be fewer grains per unit volume, and thus grain size will be larger. Generally, rapid cooling produces smaller grains, whereas slow cooling produces larger grains.

Note in Fig. 1.11d that the growing grains eventually interfere with and impinge upon one another; the interfaces that separate the individual grains are called **grain boundaries**. Note also that the crystallographic orientation changes abruptly from one grain to the next across the grain boundaries. Recall, from Section 1.4, that the behavior of a single crystal or a single grain is anisotropic. Because its many grains have random crystallographic orientations, the behavior of a polycrystalline metal is thus essentially **isotropic**; that is, its properties do not vary with direction.

1.5.1 Grain Size

Grain size has a major influence on the mechanical properties of metals. At room temperature, for example, a large grain size is generally associated with low strength, low hardness, and low ductility. Grains can be so large as to be visible with the naked eye; zinc grains on the surface of galvanized sheet steels are an example. Large grains also cause a rough surface appearance after the material has been plastically deformed, particularly in the stretching of sheet metals (see **orange peel**, Section 1.7). The yield strength, S_y, of the metal is the most sensitive property and is related to grain size by the empirical formula (known as the ***Hall-Petch** equation*)

$$S_y = S_{yi} + kd^{-1/2} \tag{1.1}$$

where S_{yi} is the yield stress for a large grained material, k is a constant, and d is the mean grain diameter. Equation (1.1) is valid for a temperature below the recrystallization temperature of the metal.

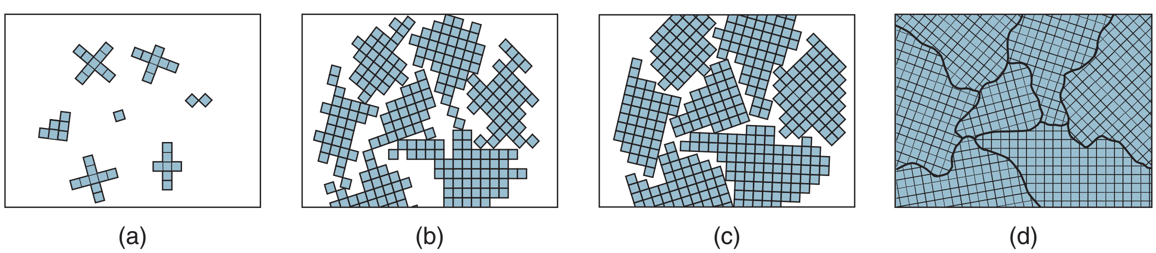


Figure 1.11: Schematic illustration of the stages during the solidification of molten metal; each small square represents a unit cell. (a) Nucleation of crystals at random sites in the molten metal; note that the crystallographic orientation of each site is different. (b) and (c) Growth of crystals as solidification continues. (d) Solidified metal, showing individual grains and grain boundaries; note the different angles at which neighboring grains meet each other.

Grain size is usually measured either by counting the number of grains in a given area or by counting the number of grains that intersect a prescribed length of a line (drawn randomly on an enlarged image of the grains), taken under a microscope on a polished and etched specimen. Software is available to automate these tasks. Grain size may also be determined by comparing such an image against a standard chart.

The American Society for Testing and Materials (ASTM) grain size number, n, is related to the number of grains, N, per square inch at a magnification of $100\times$ (equal to 0.0645 mm^2 of actual area) by the formula

$$N = 2^{n-1} \tag{1.2}$$

An International Standards Organization (ISO) equivalent to Eq. (1.2) is given by

$$N' = 8\left(2^{G_m}\right)$$

where G_m is the metric grain size number and N' is the number of grains per square millimeter at $1\times$ magnification.

Because grains are typically extremely small, numerous grains may occupy a small volume of metal (Table 1.1). Grain sizes between 5 and 8 are generally considered fine grained; a grain size of 7 is typically acceptable for sheet metals for making automotive bodies, appliances, and kitchen utensils (Chapter 16).

Example 1.1 Number of Grains in a Paper Clip

Given: A paper clip is made of wire that is 120 mm long and 0.75 mm in diameter, with an ASTM grain size of 9.

Find: Calculate the number of grains in the paper clip.

Solution: A metal with an ASTM grain size of 9 has 185,000 grains per mm^3 (see Table 1.1). The volume of the paper clip is

$$V = \frac{\pi}{4}d^2 l = \frac{\pi}{4}\left(0.75\right)^2\left(120\right) = 53.0 \text{ mm}^3$$

The total number of grains is calculated by multiplying the volume by the grains per mm^3, or No. of grains = (53.0 mm^3) (185,000 grains/mm^3) = 9.81 million.

1.5.2 Influence of Grain Boundaries

Grain boundaries have an important influence on the strength and ductility of metals, and because they interfere with dislocation movement also influence strain hardening. The magnitude of these effects depends on temperature, deformation rate, and the type and amount of impurities present *along* the grain boundaries.

Because the atoms along the grain boundaries are more disordered and hence packed less efficiently, grain boundaries are more reactive than the grains themselves. As a result, the boundaries have lower energy than the atoms in the orderly lattice within the grains; thus, they can be more easily removed or chemically bonded to another atom. For example, the surface of a piece of metal becomes rougher when etched or it is subjected to corrosive environments. (See also *end grains in forging*, Section 14.5.)

At elevated temperature, and in metals whose properties depend on the rate at which they are deformed, plastic deformation also takes place by means of *grain-boundary sliding*. The **creep** mechanism (elongation under stress over time, usually at elevated temperatures) involves grain-boundary sliding (see Section 2.8).

Grain-boundary embrittlement. When exposed to certain low-melting-point metals, a normally ductile and strong metal may crack when subjected to very low external stresses. Two examples of such behavior are (a) aluminum wetted with a mercury–zinc amalgam or with liquid gallium and (b) copper at elevated

Table 1.1: Grain Sizes.

ASTM No.	Grains/mm^2	Grains/mm^3
−3	1	0.7
−2	2	2
−1	4	5.6
0	8	16
1	16	45
2	32	128
3	64	360
4	128	1020
5	256	2900
6	512	8200
7	1024	23,000
8	2048	65,000
9	4096	185,000
10	8200	520,000
11	16,400	1,500,000

temperature wetted with lead or bismuth; these elements weaken the grain boundaries of the metal by *embrittlement*. The term **liquid-metal embrittlement** is used to describe such phenomena because the embrittling element is in a liquid state. However, embrittlement can also occur at temperatures well below the melting point of the embrittling element, known as **solid-metal embrittlement**.

Another embrittlement phenomenon is **hot shortness**, caused by local melting of a constituent or of an impurity along a grain boundary at a temperature below the melting point of the metal itself. When subjected to plastic deformation at elevated temperatures (*hot working*), the metal crumbles along its grain boundaries; examples are (a) antimony in copper, (b) leaded steels (Section 21.7.1), and (c) leaded brass. To avoid hot shortness, the metal is usually worked at a lower temperature to prevent softening and melting along the grain boundaries. **Temper embrittlement** in alloy steels is another form of embrittlement, caused by segregation (movement) of impurities to the grain boundaries (Section 4.11).

1.6 Plastic Deformation of Polycrystalline Metals

When a polycrystalline metal with uniform *equiaxed grains* (grains having equal dimensions in all directions) is subjected to plastic deformation at room temperature (called *cold working*), the grains become deformed and elongated, as shown schematically in Fig. 1.12. Deformation may be carried out by, for example, compressing the metal, as is done in a forging operation to make a turbine disk (Chapter 14) or by subjecting it to tension, as is done in stretch forming of sheet metal (Section 16.6). The deformation within each grain takes place by the mechanisms described in Section 1.4 for a single crystal.

During plastic deformation, the grain boundaries remain intact and mass continuity is maintained. The deformed metal exhibits higher strength than before, because of entanglement of dislocations with grain boundaries and with each other. The increase in strength depends on the degree of deformation (*strain*) to which the metal is subjected; the higher the deformation, the stronger the metal becomes. The strength increase is higher for metals with smaller grains, because they have a larger grain-boundary surface area per unit volume of metal and, hence, more entanglement of dislocations.

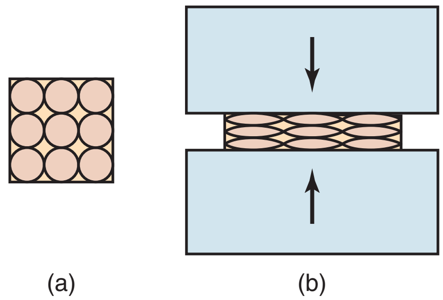


Figure 1.12: Plastic deformation of idealized (equiaxed) grains in a specimen subjected to compression (such as occurs in the forging or rolling of metals): (a) before deformation and (b) after deformation. Note the alignment of grain boundaries along a horizontal direction; this effect is known as *preferred orientation*.

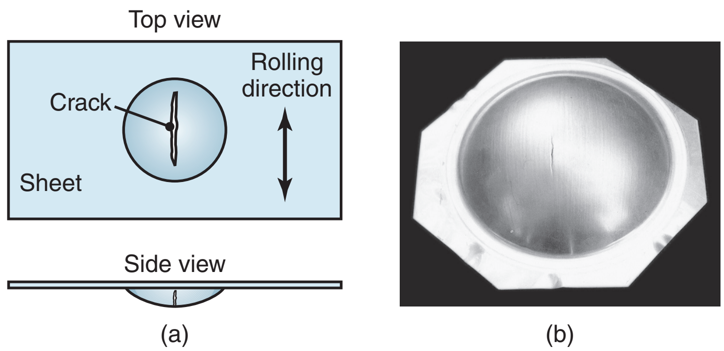


Figure 1.13: (a) Schematic illustration of a crack in sheet metal that has been subjected to bulging (caused, for example, by pushing a steel ball against the sheet). Note the orientation of the crack with respect to the rolling direction of the sheet; this sheet is anisotropic. (b) Aluminum sheet with a crack (vertical dark line at the center) developed in a bulge test; the rolling direction of the sheet was vertical. *Source:* After J.S. Kallend, Illinois Institute of Technology.

Anisotropy (Texture). Note in Fig. 1.12b that, as a result of plastic deformation, the grains have elongated in one direction and contracted in the other direction. Consequently, the metal has become *anisotropic*, and thus its properties in the vertical direction are different from those in the horizontal direction. The degree of anisotropy depends on the temperature at which deformation has taken place and on how uniformly the metal is deformed. Note from the crack direction shown in Fig. 1.13, for example, that the ductility of the cold-rolled sheet in the transverse direction is lower than in its rolling direction (see also Section 16.5).

Anisotropy influences both mechanical and physical properties of metals, described in Chapter 3. For example, sheet steel for electrical transformers is rolled in such a way that the resulting deformation imparts anisotropic magnetic properties to the sheet. This operation then reduces magnetic-hysteresis losses, thus improving the efficiency of transformers (see also *amorphous alloys*, Section 6.15). There are two general types of anisotropy in metals: *preferred orientation* and *mechanical fibering*.

Preferred Orientation. Also called **crystallographic anisotropy**, *preferred orientation* can be best described by referring to Fig. 1.6a. When a single-crystal metal is subjected to tension, the sliding blocks rotate toward the direction of the tensile force; as a result, slip planes and slip bands tend to align themselves with the general direction of deformation. Similarly, for a polycrystalline metal, with grains in random orientations, all slip directions tend to align themselves with the direction of the tensile force being applied. By contrast, slip planes under compression tend to align themselves in a direction perpendicular to the direction of the applied compressive force.

Mechanical Fibering. This is a type of anisotropy that results from the alignment of inclusions (*stringers*), impurities, and voids in the metal during deformation. Note that if the spherical grains shown in Fig. 1.12a were coated with impurities, these impurities would align themselves in a generally horizontal direction after deformation. Because the impurities weaken the grain boundaries, this piece of metal will now be less strong and less ductile when tested in the vertical direction. As an analogy, consider plywood, which is strong in tension along its planar direction but splits easily when subjected to tension in its thickness direction.

1.7 Recovery, Recrystallization, and Grain Growth

Recall that plastic deformation at room temperature causes (a) distortion of the grains and grain boundaries, leading to anisotropic behavior, (b) a general increase in strength, and (c) a decrease in ductility. These effects can be reversed and the properties of the metal brought back to their original levels by heating the metal to a specific temperature range for a given period of time—a process called **annealing** (described in detail in Section 4.11).

Three events take place, consecutively, during this process:

1. **Recovery**. During *recovery*, which occurs at a certain temperature range below the **recrystallization temperature** of the metal (described next), the stresses in the highly deformed regions of the metal are relieved. Subgrain boundaries begin to form (called **polygonization**), with no significant change in mechanical properties (such as hardness and strength, Fig. 1.14).

2. **Recrystallization**. This is the process in which, within a certain temperature range, new equiaxed and strain-free grains are formed, replacing the older grains. The temperature required for recrystallization ranges approximately between $0.3T_m$ and $0.5T_m$, where T_m is the melting point of the metal on the absolute scale.

   Generally, the recrystallization temperature is defined as the temperature at which complete recrystallization occurs within approximately one hour. Recrystallization decreases the density of dislocations, lowers the strength, and raises the ductility of the metal (Fig. 1.14). Lead, tin, cadmium, and zinc recrystallize at about room temperature; consequently, they do not usually work harden.

   The recrystallization temperature depends on the degree of prior cold work (work hardening): the more the cold work, the lower the temperature required for recrystallization. The reason is that, as the amount of cold work increases, the number of dislocations and the amount of energy stored in dislocations (called **stored energy**) also increase. This energy supplies some of the work required for recrystallization. Recrystallization is also a function of time, because it involves **diffusion**—the movement and exchange of atoms across grain boundaries. The effects of temperature, time, and the degree of plastic deformation by cold working on recrystallization are as follows:

   (a) For a constant degree of deformation by cold working, the time required for recrystallization decreases with increasing temperature.

   (b) The higher the prior cold work, the lower the temperature required for recrystallization.

   (c) The higher the amount of deformation, the smaller the grain size becomes during recrystallization; this effect is a commonly used method for converting a coarse-grained structure to one having a finer grain, and with improved properties.

   (d) Some anisotropy (due to preferred orientation) usually persists after recrystallization; to restore isotropy, a temperature higher than that required for recrystallization may be necessary.

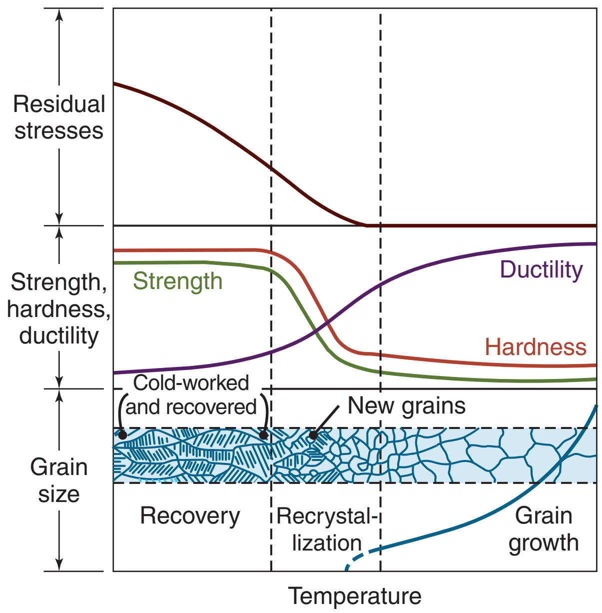


Figure 1.14: Schematic illustration of the effects of recovery, recrystallization, and grain growth on mechanical properties and on the shape and size of grains. Note the formation of small new grains during recrystallization. *Source:* After G. Sachs.

3. **Grain growth.** If the temperature is raised higher, the grains begin to grow (see lower part of Fig. 1.14), and their size may eventually exceed the original grain size. Called *grain growth*, this phenomenon adversely affects mechanical properties (Fig. 1.14). Large grains also produce a rough surface appearance on sheet metals, called **orange peel**, when they are stretched to form a part, or on the surfaces of a piece of metal when subjected to bulk deformation, such as compression in forging (Chapter 14).

1.8 Cold, Warm, and Hot Working

Cold working refers to plastic deformation carried out at room temperature. When deformation occurs above the recrystallization temperature, it is called **hot working**. *Cold* and *hot* are relative terms, as can be seen from the fact that deforming lead at room temperature is a hot-working process, because the recrystallization temperature of lead is about room temperature. As the name implies, **warm working** is carried out at intermediate temperatures; thus, warm working is a compromise between cold and hot working. The important technological differences in products that are processed by cold, warm, and hot working are described in Part III.

The temperature ranges for these three categories of plastic deformation are given in Table 1.2 in terms of a ratio, T/T_m, where T is the working temperature and T_m is the melting point of the metal, both on the absolute scale. Although it is dimensionless, this ratio is known as the **homologous temperature**.

Table 1.2: Homologous Temperature Ranges for Various Processes.

Process	T/T_m
Cold working	< 0.3
Warm working	0.3–0.5
Hot working	> 0.6

Summary

- There are three basic crystal structures in metals: body-centered cubic (bcc), face-centered cubic (fcc), and hexagonal close-packed (hcp). Grains made of these crystals typically contain various defects and imperfections, such as dislocations, vacancies, impurities, inclusions, and grain boundaries. Polycrystalline metals consist of many crystals, *or grains*, in random orientations.
- Plastic deformation in metals takes place by a slip mechanism. Although the theoretical shear stress required to cause slip is very high, actual required stresses are much lower because of the presence of dislocations (edge or screw type). Dislocations become entangled with one another or are impeded by barriers such as grain boundaries, impurities, and inclusions. As a result, the shear stress required to cause further slip is increased; consequently, the overall strength and hardness of the metal is also increased (through work hardening or strain hardening).
- Grain size has a significant effect on the strength of metals: the smaller the size, the stronger is the metal, and the larger the size, the more ductile is the metal. However, excessively large grains are generally associated with brittle behavior.
- Grain boundaries have a major influence on the behavior of metals, as boundaries can undergo embrittlement, severely reducing ductility at elevated temperatures (hot shortness). They are also responsible for the creep phenomenon, due to grain boundary sliding.
- Metals may be plastically deformed at room, warm, or high temperatures; their behavior and workability depend largely on whether deformation takes place below or above the recrystallization temperature of the metal. Deformation at room temperature (cold working) results in higher strength, but reduced ductility; generally, it also causes anisotropy (either preferred orientation or mechanical fibering), whereby the properties are different in different directions.
- The effects of cold working can be reversed by annealing; that is, heating the metal to a specific temperature range for a given period of time, thereby allowing the successive stages of recovery, recrystallization, and grain growth to take place.

Key Terms

Allotropism
Anisotropy
Basal plane
Body-centered cubic
Cold working
Covalent bond
Creep
Crystals
Dislocations
Elastic deformation
Embrittlement
Face-centered cubic
Grains
Grain boundaries
Grain growth
Grain size
Hall-Petch effect
Hexagonal close-packed
Homologous temperature
Hot shortness
Hot working
Imperfections
Ionic bond
Lattice structure
Mechanical fibering
Metallic bond

Nucleation
Orange peel
Plastic deformation
Polycrystals
Polygonization
Polymorphism
Preferred orientation
Primary bond
Recovery
Recrystallization
Secondary bond
Shear stress
Slip band
Slip plane
Slip system
Strain hardening
Structure insensitive
Structure sensitive
Texture
Twinning
Unit cell
Vacancy
van der Waals force
Warm working
Work hardening

Bibliography

Ashby, M.F., and Jones, D.R.H., **Engineering Materials**, Vol. 1, **An Introduction to Their Properties and Applications**, 3rd ed., Pergamon, 2005; Vol. 2, **An Introduction to Microstructures, Processing and Design**, Pergamon, 2005.

Ashby, M.F., **Materials Selection in Mechanical Design**, 4th ed., Pergamon, 2010.

Askeland, D.R., and Wright, W.J., **The Science and Engineering of Materials**, 7th ed., CENGAGE Learning, 2015.

Brandt, D.A., and Warner, J.C., **Metallurgy Fundamentals**, 5th ed., Goodheart-Wilcox, 2009.

Callister, W.D., Jr., and Rethwisch, D.G., **Materials Science and Engineering: An Introduction**, 9th ed., Wiley, 2013.

Hosford, W.F., **Physical Metallurgy**, 2nd ed., Taylor & Francis, 2010.

Shackelford, J.F., **Introduction to Materials Science for Engineers**, 8th ed., Pearson, 2014.

Review Questions

1.1. What is the difference between an atom and a molecule, and a molecule and a crystal?

1.2. Describe ionic, covalent, and metallic bonds.

1.3. Explain the difference between a unit cell and a single crystal.

1.4. In tables on crystal structures, iron is listed as having both a bcc and an fcc structure. Why?

1.5. Define anisotropy. What is its significance?

1.6. What effects does recrystallization have on the properties of metals?

1.7. What is strain hardening, and what effects does it have on the properties of metals?

1.8. Explain what is meant by structure-sensitive and structure-insensitive properties of metals.

1.9. Make a list of each of the major kinds of imperfection in the crystal structure of metals, and describe them.

1.10. What influence does grain size have on the mechanical properties of metals?

1.11. What is the relationship between the nucleation rate and the number of grains per unit volume of a metal?

1.12. What is a slip system, and what is its significance?

1.13. Explain the difference between recovery and recrystallization.

1.14. What is hot shortness, and what is its significance?

1.15. Explain the advantages and limitations of cold, warm, and hot working, respectively.

1.16. What is a slip band? Explain why they can be seen on the surface of a crystal.

1.17. Describe what the orange peel effect is. Explain why we may have to be concerned with the orange-peel effect on metal surfaces.

1.18. Some metals, such as lead, do not become stronger when worked at room temperature. Explain the reason.

1.19. Describe the difference between preferred orientation and mechanical fibering.

1.20. Differentiate between stress relaxation and stress relieving.

1.21. What is twinning? How does it differ from slip?

1.22. What is annealing?

1.23. Which one is larger—a unit cell or a grain?

1.24. Describe the different approaches used to measure grain size.

Qualitative Problems

1.25. Explain your understanding of why the study of the crystal structure of metals is important.

1.26. What is the significance of the fact that some metals undergo allotropism?

1.27. Is it possible for two pieces of the same metal to have different recrystallization temperatures? Is it possible for recrystallization to take place in some regions of a part before it does in other regions of the same part? Explain.

1.28. Describe your understanding of why different crystal structures exhibit different strengths and ductilities.

1.29. A cold-worked piece of metal has been recrystallized. When tested, it is found to be anisotropic. Explain the probable reason.

1.30. What materials and structures can you think of (other than metals) that exhibit anisotropic behavior?

1.31. Two parts have been made of the same material, but one was formed by cold working and the other by hot working. Explain the differences you might observe between the two.

1.32. Explain the importance of homologous temperature.

1.33. Do you think it might be important to know whether a raw material to be used in a manufacturing process has anisotropic properties? What about anisotropy in the finished product? Explain.

1.34. What is the difference between an interstitial atom and a substitutional atom?

1.35. Explain why the strength of a polycrystalline metal at room temperature decreases as its grain size increases.

1.36. Describe the technique you would use to reduce the orange-peel effect on the surface of workpieces.

1.37. What is the significance of the fact that such metals as lead and tin have a recrystallization temperature that is about room temperature?

1.38. It was stated in this chapter that twinning usually occurs in hcp materials, but Fig. 1.6b shows twinning in a rectangular array of atoms. Can you explain the discrepancy?

1.39. It has been noted that the more a metal has been cold worked, the less it strain hardens. Explain why.

1.40. Is it possible to cold work a metal at temperatures above the boiling point of water? Explain.

1.41. Comment on your observations regarding Fig. 1.14.

1.42. Is it possible for a metal to be completely isotropic? Explain.

1.43. Referring to Fig. 1.1, assume you can make a ball bearing from a single crystal. What advantages and disadvantages would such a bearing have?

1.44. Referring to Fig. 1.10, explain why edge dislocations cannot cross grain boundaries using appropriate sketches.

Quantitative Problems

1.45. How many atoms are in a single repeating cell of an bcc crystal structure? How many in a repeating cell of an hcp structure?

1.46. The atomic weight of gold is 196.97, meaning that 6.023×10^{23} atoms weigh 196.97 g. The density of gold is 19,320 kg/m^3, and pure golf forms fcc crystals. Estimate the diameter of a gold atom.

1.47. Plot the data given in Table 1.1 in terms of grains/mm^2 vs. grains/mm^3, and discuss your observations.

1.48. A strip of metal is reduced from 50 mm in thickness to 25 mm by cold working; a similar strip is reduced from 50 mm to 30 mm. Which of these cold-worked strips will recrystallize at a lower temperature? Why?

1.49. The ball of a ballpoint pen is 1.6 mm in diameter and has an ASTM grain size of 12. How many grains are there in the ball?

1.50. How many grains are on the surface of the head of a pin? Assume that the head of a pin is spherical with a 2-mm diameter and has an ASTM grain size of 12.

1.51. The unit cells shown in Figs. 1.3 through 1.5 can be represented by tennis balls arranged in various configurations in a box. In such an arrangement, the *atomic packing factor* (APF) is defined as the ratio of the sum of the volumes of the atoms to the volume of the unit cell. Show that the APF is 0.68 for the bcc structure and 0.74 for the fcc structure.

1.52. Show that the lattice constant a in Fig. 1.4a is related to the atomic radius by the formula $\alpha = 2\sqrt{2}R$, where R is the radius of the atom as depicted by the tennis-ball model.

1.53. Show that, for the fcc unit cell, the radius r of the largest hole is given by $r = 0.414R$. Determine the size of the largest hole for the iron atoms in the fcc structure.

1.54. A technician determines that the grain size of a certain etched specimen is 8. Upon further checking, it is found that the magnification used was $150\times$, instead of the $100\times$ that is required by the ASTM standards. Determine the correct grain size.

1.55. If the diameter of the aluminum atom is 0.28 nm, how many atoms are there in a grain of ASTM grain size 10?

1.56. The following data are obtained in tension tests of brass:

Grain size (μm)	Yield strength (MPa)
30	150
40	140
100	105
150	90
200	75

Does the material follow the Hall-Petch effect? If so, what is the value of k?

1.57. Does water have a homologous temperature? What is the highest temperature where water (ice) can get cold worked?

1.58. The atomic radius of iron is 0.125 nm, while that of a carbon atom is 0.070 nm. Can a carbon atom fit inside a steel bcc structure without distorting the neighboring atoms?

1.59. Estimate the atomic radius for the following materials and data: (a) Aluminum (atomic weight = 26.98 g/mol, density = 2700 kg/m^3); (b) silver (atomic weight = 107.87 g/mol, density = 10,500 kg/m^3); (c) titanium (atomic weight = 47.87 g/mol, density = 4506 kg/m^3).

1.60. A simple cubic structure involves atoms located at the cube corners that are in contact with each other along the cube edges. Make a sketch of a simple cubic structure, and calculate its atomic packing factor.

1.61. Estimate the ASTM grain size number for a 300-mm silicon wafer used to produce computer chips.

1.62. Pure copper and pure titanium follow the Hall–Petch equation. For copper, $S_{yi} = 24$ MPa and $k = 0.12$ MPa-m$^{1/2}$. For titanium, $S_{yi} = 80$ MPa and $k = 0.40$ MPa-m$^{1/2}$. (a) Plot the yield strength of these metals as a function of grain size for ASTM grain sizes of -3 to 11. (b) Explain which material would see greater strengthening from a reduction in grain size, as in cold working.

Synthesis, Design, and Projects

1.63. By stretching a thin strip of polished metal, as in a tension-testing machine, demonstrate and comment on what happens to its reflectivity as the strip is being stretched.

1.64. Draw some analogies to mechanical fibering—for example, layers of thin dough sprinkled with flour or melted butter between each layer.

1.65. Draw some analogies to the phenomenon of hot shortness.

1.66. Obtain a number of small balls made of plastic, wood, marble, or metal, and arrange them with your hands or glue them together to represent the crystal structures shown in Figs. 1.3–1.5. Comment on your observations.

1.67. Take a deck of playing cards, place a rubber band around it, and then slip the cards against each other to represent Figs. 1.6a and 1.7. If you repeat the same experiment with more and more rubber bands around the same deck, what are you accomplishing as far as the behavior of the deck is concerned?

1.68. Give examples in which anisotropy is scale dependent. For example, a wire rope can contain annealed wires that are isotropic on a microscopic scale, but the rope as a whole is anisotropic.

1.69. The movement of an edge dislocation was described in Section 1.4.1 by means of an analogy involving a hump in a carpet on the floor and how the whole carpet can eventually be moved by moving the hump forward. Recall that the entanglement of dislocations was described in terms of two humps at different angles. Use a piece of cloth placed on a flat table to demonstrate these phenomena.

1.70. If you want to strengthen a material, would you wish to have it consist of one grain, or would you want it to have grains that contain the minimum number of atoms? Explain.

Chapter 2

Mechanical Behavior, Testing, and Manufacturing Properties of Materials

Examples:

- This chapter examines the effects of external forces on the behavior of materials, and the test methods employed in determining mechanical properties.
- The tension test, described first, is commonly used for quantifying a number of material parameters, including elastic modulus, yield stress, ultimate strength, ductility, and toughness.
- Compression tests are important because they more closely simulate some metalworking processes; however, they have the unavoidable drawback of contributing the effects of friction to the test results.

- Bending tests are particularly useful for brittle materials; three- and four-point tests are in common use.
- Hardness and the variety of hardness tests and their range of applicability are then explored.
- Fatigue involves the failure of materials subjected to cyclic or repeating loads; creep is deformation due to the application of a constant load over an extended period of time. These phenomena are also discussed.
- This chapter ends with descriptions of the types of and the factors involved in failure and fracture of materials.

2.1 Introduction

In manufacturing operations, parts and components are formed into a wide variety of shapes by applying external forces to the workpiece, typically by means of a variety of tools and dies. Common examples of such operations are forging of turbine disks, extruding various components for aluminum ladders, drawing wire for making nails, and rolling metal to make sheets for appliances. Forming operations may be carried out at room temperature or at elevated temperatures, and at a low or a high rate of deformation. Many of these operations are also used in forming and shaping nonmetallic materials, such as plastics, ceramics, and composite materials.

As indicated in Tables 1.2 and 2.1, a wide variety of metallic and nonmetallic materials is now available, with an equally wide range of properties and characteristics. This chapter covers those aspects of mechanical properties and behavior of metals that are relevant to the design and manufacturing of products, and includes commonly used test methods employed in assessing various material properties.

2.2 Tension

The **tension test** is the most commonly used method for determining the *mechanical properties* of materials such as strength, ductility, toughness, elastic modulus, and strain-hardening exponent. The test first requires the preparation of a **test specimen**, as shown in Fig. 2.1a. Although most specimens are solid and

Table 2.1: Relative Mechanical Properties of Various Materials at Room Temperature, in Decreasing Order. Metals Are in Their Alloy Form.

Strength	Hardness	Toughness	Stiffness	Strength/Density
Glass fibers	Diamond	Ductile metals	Diamond	Reinforced plastics
Carbon fibers	Cubic boron nitride	Reinforced plastics	Carbides	Titanium
Kevlar fibers	Carbides	Thermoplastics	Tungsten	Steel
Carbides	Hardened steels	Wood	Steel	Aluminum
Molybdenum	Titanium	Thermosets	Copper	Magnesium
Steels	Cast irons	Ceramics	Titanium	Beryllium
Tantalum	Copper	Glass	Aluminum	Copper
Titanium	Thermosets		Ceramics	Tantalum
Copper	Magnesium		Reinforced plastics	
Reinforced thermosets	Thermoplastics		Wood	
Reinforced thermoplastics	Tin		Thermosets	
Thermoplastics	Lead		Thermoplastics	
Lead			Rubbers	

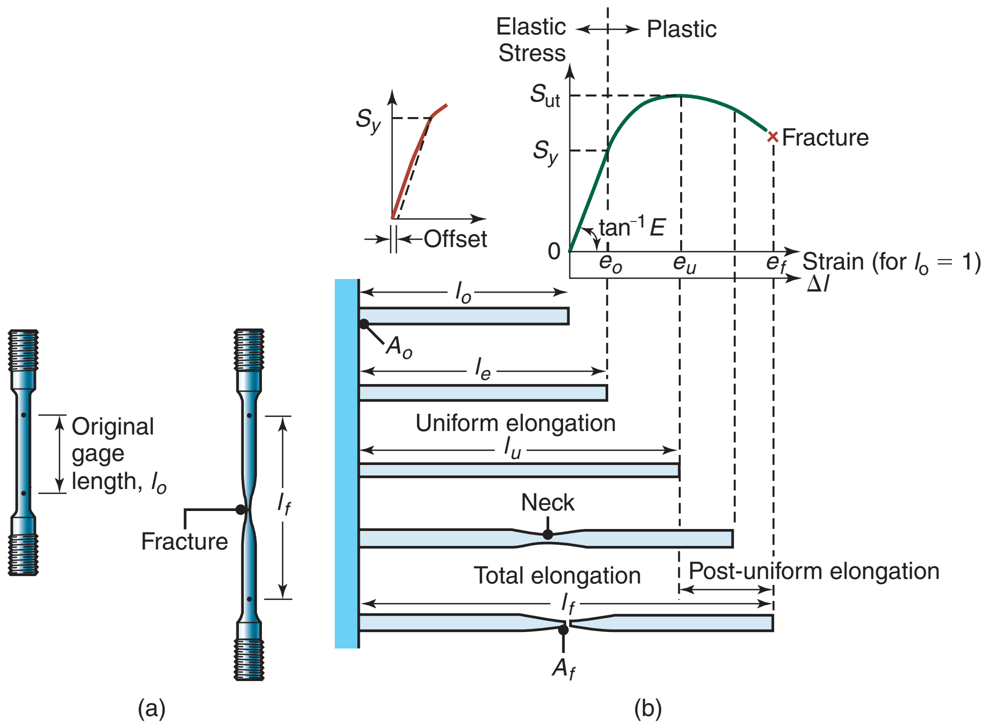


Figure 2.1: (a) A standard tensile-test specimen before and after pulling, showing original and final gage lengths. (b) Stages in specimen behavior in a tension test.

round in shape, they can also be flat or tubular. The specimen is prepared generally according to American Society for Testing and Materials (ASTM) specifications, although various other specifications are also available from corresponding organizations around the world.

Typically, the specimen has an **original gage length**, l_o, generally 50 mm, and a cross-sectional area, A_o, usually with a diameter of 12.5 mm. The specimen is mounted in the jaws of a tension-testing machine, equipped with various accessories and controls so that it can be tested at different temperatures and rates of deformation.

2.2.1 Stress–Strain Curves

A typical sequence of events in a tension test is shown in Fig. 2.1b. When the load is first applied, the specimen elongates in proportion to the load, a behavior called **linear elastic** (Fig. 2.2). If at this stage the load is removed, the specimen returns to its original length and shape, in a manner similar to stretching a rubber band and releasing it.

The **engineering stress** (also called **nominal stress**) is defined as the ratio of the applied load, P, to the original cross-sectional area, A_o, of the specimen:

$$\sigma = \frac{P}{A_o}. \tag{2.1}$$

The **engineering strain** is defined as

$$e = \frac{l - l_o}{l_o}, \tag{2.2}$$

where l is the instantaneous length of the specimen.

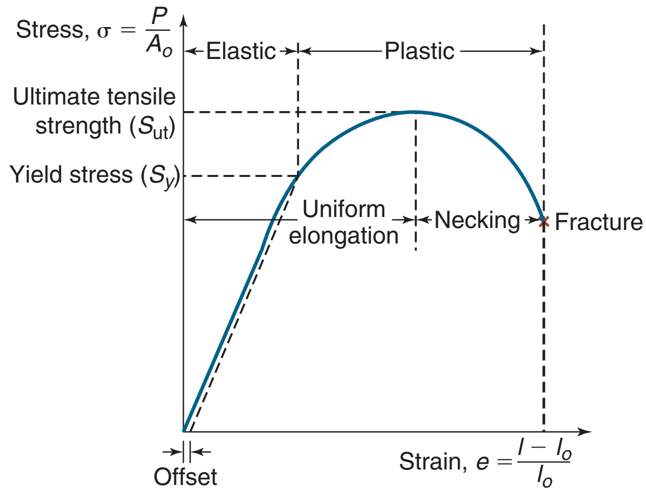


Figure 2.2: A typical stress–strain curve obtained from a tension test, showing various features.

As the load is increased further, the specimen begins to undergo *nonlinear* elastic deformation at a stress level called the *proportional limit*. At that point, the stress and strain are no longer proportional, as they were in the linear elastic region, but when unloaded, the specimen still returns to its original shape. **Permanent (plastic) deformation** occurs when the **yield stress**, S_y, of the material is reached. (The yield strength and other properties of various metallic and nonmetallic materials are given in Table 2.2.)

For soft and ductile materials, it may be difficult to determine the exact location on the stress–strain curve at which yielding occurs, because the slope of the curve begins to decrease slowly above the proportional limit. For such materials, S_y is usually defined by drawing a line with the same slope as the linear elastic curve, but that it is **offset** by a strain of 0.002, or 0.2% elongation. The yield strength is then defined as the stress where the offset line intersects the stress–strain curve. This simple procedure is shown on the left side in Fig. 2.2.

As the specimen begins to elongate under a continuously increasing load, its cross-sectional area decreases **permanently** and **uniformly** within its gage length. If the specimen is unloaded (from a stress level higher than the yield stress), the curve follows a straight line downward and parallel to the original slope of the curve, as shown in Fig. 2.3. As the load is increased further, the engineering stress eventually reaches a maximum and then begins to decrease (Fig. 2.2). The maximum engineering stress is called the **tensile strength**, or **ultimate tensile strength**, S_{ut}, of the material. (Values for S_{ut} for a variety of materials are given in Table 2.2.)

If the specimen is loaded beyond its ultimate tensile strength, it begins to **neck**, or *neck down*. The cross-sectional area of the specimen is no longer uniform along the gage length and is smaller in the necked region. As the test progresses, the engineering stress drops further and the specimen fractures at the necked region (Fig. 2.1a). The engineering stress at fracture is known as the **breaking** or **fracture strength**.

The ratio of stress to strain in the elastic region is called the **modulus of elasticity**, E, or **Young's modulus** (after T. Young, 1773–1829):

$$E = \frac{\sigma}{e}. \tag{2.3}$$

This linear relationship is known as **Hooke's law** (after R. Hooke, 1635–1703).

Table 2.2: Mechanical Properties of Various Materials at Room Temperature.

Materials	Elastic modulus (GPa)	Yield strength (MPa)	Ultimate tensile strength (MPa)	Elongation in 50 mm (%)	Poisson's ratio, ν
Metals (wrought)					
Aluminum and its alloys	69–79	35–550	90–600	45–4	0.31–0.34
Copper and its alloys	105–150	76–110	140–1310	65–3	0.33–0.35
Lead and its alloys	14	14	20–55	50–9	0.43
Magnesium and its alloys	41–45	130–305	240–380	21–5	0.29–0.35
Molybdenum and its alloys	330–360	80–2070	90–2340	40–30	0.32
Nickel and its alloys	180–214	105–1200	345–1450	60–5	0.31
Steels	190–210	205–1725	415–1750	65–2	0.28–0.33
Titanium and its alloys	80–130	344–1380	415–1450	25–7	0.31–0.34
Tungsten and its alloys	350–400	550–690	620–760	0	0.27
Zinc and its alloys	50	25–180	240–550	65–5	0.27
Nonmetallic materials					
Ceramics	70–1000	—	140–2600	0	0.2
Diamond	820–1050	—	60,000	—	0.2
Glass and porcelain	70–80	—	140	0	0.24
Silicon carbide (SiC)	200–500	—	310–400	—	0.19
Silicon nitride (Si_2N_4)	280–310	—	160–580	—	0.26
Rubbers	0.01–0.1	—	—	—	0.5
Thermoplastics	1.4–3.4	—	7–80	1000–5	0.32–0.40
Thermoplastics, reinforced	2–50	—	20–120	10–1	0–0.5
Thermosets	3.5–17	—	35–170	0	0.34–0.5
Boron fibers	380	—	3500	0	0.27
Carbon fibers	275–415	—	2000–3000	0	0.21–0.28
Glass fibers	73–85	—	3500–4600	0	0.22–0.26
Kevlar fibers	62–117	—	2800	0	0.36
Spectra Fibers	73–100	—	2400–2800	3	0.46

Note: In the upper part of the table, the lowest values for E, S_y, and S_{ut} and the highest values for elongation are for pure metals.

Note in Eq. (2.3) that, because engineering strain is dimensionless, E has the same units as stress. The modulus of elasticity is the slope of the elastic portion of the curve and indicates the **stiffness** of the material. The higher the elastic modulus, the higher is the load required to stretch the specimen to the same extent, and thus the stiffer is the material. Compare, for example, the stiffness of metal wire with that of a rubber band or plastic when they are both stretched.

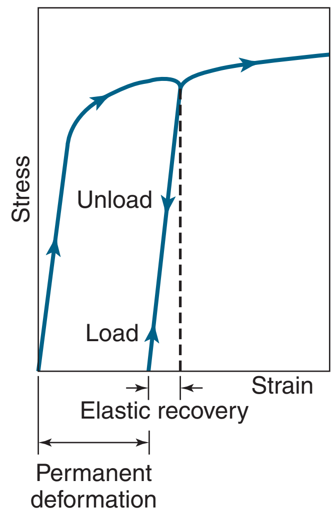


Figure 2.3: Schematic illustration of the loading and the unloading of a tensile-test specimen. Note that, during unloading, the curve follows a path parallel to the original elastic slope.

The elongation of the specimen under tension is accompanied by lateral contraction; this effect can easily be observed by stretching a rubber band. The absolute value of the ratio of the lateral strain to the longitudinal strain is known as **Poisson's ratio** (after S.D. Poisson, 1781–1840) and is denoted by the symbol ν.

2.2.2 Ductility

An important behavior observed during a tension test is **ductility**—the extent of plastic deformation that the material can undergo prior to fracture. There are two common measures of ductility. The first is the **total elongation** of the specimen, given by

$$\text{Elongation} = \frac{l_f - l_o}{l_o} \times 100, \tag{2.4}$$

where l_f and l_o are measured as shown in Fig. 2.1a. Note that the elongation is based on the original gage length of the specimen, and that it is calculated as a percentage.

The second measure of ductility is the **reduction of area**, given by

$$\text{Reduction of area} = \frac{A_o - A_f}{A_o} \times 100, \tag{2.5}$$

where A_o and A_f are, respectively, the original and final (fracture) cross-sectional areas of the test specimen. Thus, the ductility of a piece of chalk is zero, because it does not stretch at all or reduce in its cross section. By contrast, a ductile specimen, such as a pure metal (such as copper or gold) or thermoplastic, stretches and necks considerably before it fractures.

2.2.3 True Stress and True Strain

Recall that engineering stress is based on the *original* cross-sectional area, A_o, of the specimen. However, the *instantaneous* cross-sectional area of the specimen becomes smaller as it elongates, just as the area of a rubber band does when stretched. Thus, engineering stress does not represent the *actual* (or true) stress to which the specimen is subjected.

True stress is defined as the ratio of the applied load, P, to the actual (instantaneous, hence *true*) cross-sectional area, A, of the specimen:

$$\sigma = \frac{P}{A}. \tag{2.6}$$

For true strain, first consider the elongation of the specimen as consisting of increments of instantaneous change in length. Then, using calculus, it can be shown that the **true strain** (*natural* or *logarithmic strain*) is calculated as

$$\epsilon = \ln\left(\frac{l}{l_o}\right). \tag{2.7}$$

Note from Eqs. (2.2) and (2.7) that, for small values of strain, the engineering and true strains are essentially equal; however, they diverge rapidly as the strain increases. For example, when $e = 0.1$, $\epsilon = 0.095$, and when $e = 1$, $\epsilon = 0.69$.

Unlike engineering strains, true strains indicate actual physical phenomena in the deformation of materials. For example, consider a hypothetical situation where a compression specimen 50 mm in height is reduced, between flat platens, to a final height of zero. In other words, the specimen has been deformed to an infinite diameter. According to their definitions, the engineering strain that the specimen undergoes is $(0 - 50)/50 = -1$, but the true strain is $-\infty$. Note that the answer will be the same regardless of the original height of the specimen. Clearly, then, true strain describes the extent of deformation correctly, since the deformation is indeed infinite.

2.2.4 Construction of Stress–Strain Curves

The procedure for constructing an engineering stress–strain curve is to take the load–elongation curve (Fig. 2.4a; also, Fig. 2.2) and divide (1) the load (vertical axis) by the original cross-sectional area, A_o, and (2) the elongation (horizontal axis) by the original gage length, l_o. Because A_o and l_o are constants, the engineering stress–strain curve obtained, shown in Fig. 2.4b, has the same shape as the load–elongation curve shown in Fig. 2.4a. (In this example, $A_o = 36.1$ m^2 and $A_f = 10.3$ m^2.)

True stress–true strain curves are obtained similarly, by dividing the load by the instantaneous cross-sectional area, with the true strain calculated from Eq. (2.7); the result is shown in Fig. 2.4c. Note the *correction* to the curve, reflecting the fact that the specimen's necked region is subjected to three-dimensional tensile stresses, as described in more advanced texts. This state of stress gives higher stress values than the actual true stress; thus, to compensate for it, the curve must be corrected downward.

The true stress–true strain curve in Fig. 2.4c can be represented by the equation

$$\sigma = K\epsilon^n, \tag{2.8}$$

where K is the **strength coefficient** and n is the **strain-hardening** (or **work-hardening**) **exponent**. Typical values for K and n for several metals are given in Table 2.3.

When the curve shown in Fig. 2.4c is a log–log plot, it will be found that the curve is approximately a straight line (Fig. 2.4d). The slope of the curve is the exponent n. Thus, the higher the slope, the greater is the strain-hardening capacity of the material—that is, the stronger and the harder it becomes as it is strained.

True stress–true strain curves for a variety of metals are given in Fig. 2.5. When reviewed in detail, some differences between Table 2.3 and Fig. 2.5 will be noted. These discrepancies are due to the different sources of data and different specimens that have been involved in obtaining them. Note also that the elastic regions in the curves have been deleted, because the slope of the curve in this region is very high. Consequently, the point of intersection of each curve with the vertical axis in this figure can be considered to be the yield strength, S_y, of the material.

The area under the true stress–true strain curve at a particular strain is the energy per unit volume (**specific energy**) of the deformed material, and it indicates the work required to plastically deform a

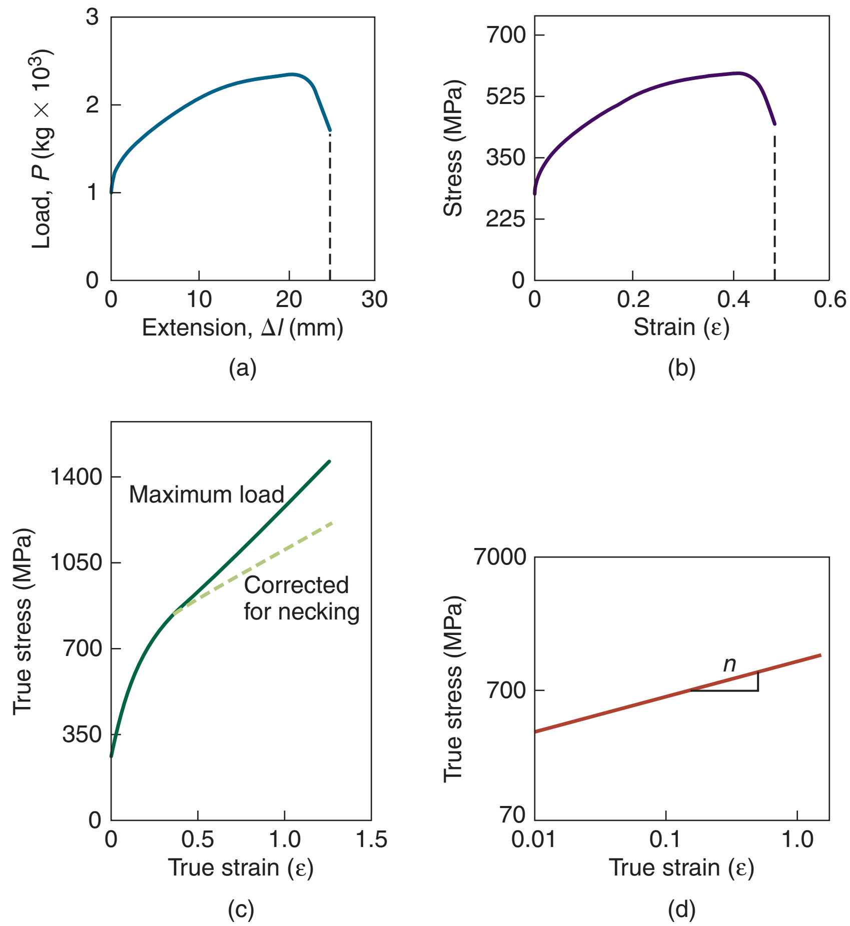


Figure 2.4: (a) Load–elongation curve in tension testing of a stainless steel specimen. (b) Engineering stress–engineering strain curve, drawn from the data in Fig. 2.4a. (c) True stress–true strain curve, drawn from the data in Fig. 2.4b. Note that this curve has a positive slope, indicating that the material is becoming stronger as it is strained. (d) True stress–true strain curve plotted on log–log paper and based on the corrected curve in Fig. 2.4c. The correction is due to the triaxial state of stress that exists in the necked region of the specimen.

unit volume of the material to that strain. The area under the true stress–true strain curve up to fracture is known as the material's **toughness**, that is, the amount of energy per unit volume that the material dissipates prior to fracture. Note that toughness involves *both* the height and width of the stress–strain curve of the material, whereas strength is related only to the *height* of the curve and ductility is related only to the *width* of the curve.

2.2.5 Strain at Necking in a Tension Test

As stated earlier, the onset of necking in a tension-test specimen corresponds to the ultimate tensile strength of the material. Note that the slope of the load–elongation curve at this point is zero, and it is there that the specimen begins to neck. The specimen cannot support the load being applied because the cross-sectional

Table 2.3: Typical Values for K and n for Selected Metals.

Material	K (MPa)	n
Aluminum		
1100–O	180	0.20
2024–T4	690	0.16
5052-O	202	0.13
6061–O	205	0.20
6061–T6	410	0.05
7075–O	400	0.17
Brass		
70–30, annealed	900	0.49
85–15, cold-rolled	580	0.34
Cobalt-base alloy, heat-treated	2070	0.50
Copper, annealed	315	0.54
Steel		
Low-C, annealed	530	0.26
1020, annealed	745	0.20
4135, annealed	1015	0.17
4135, cold-rolled	1100	0.14
4340, annealed	640	0.15
304 stainless, annealed	1275	0.45
410 stainless, annealed	960	0.10
Titanium		
Ti-6Al-4V, annealed, 20°C	1400	0.015
Ti-6Al-4V, annealed, 200°C	1040	0.026
Ti-6Al-4V, annealed, 600°C	650	0.064
Ti-6Al-4V, annealed, 800°C	350	0.146

area of the neck is becoming smaller at a rate that is higher than the rate at which the material becomes stronger (strain hardens).

The true strain at the onset of necking is numerically equal to the strain-hardening exponent, n, of the material. Thus, the higher the value of n, the higher the strain that a material can experience before it begins to neck. This observation is important, particularly in regard to sheet-metal-forming operations (Chapter 16) that involve the stretching of the workpiece material. It can be seen in Table 2.3 that annealed copper, brass, and stainless steel, for example, have high n values; this means that they can be stretched uniformly to a greater extent than the other metals listed in the table.

Example 2.1 Calculation of Ultimate Tensile Strength

Given: This example shows that the ultimate tensile strength, $S_{\rm ut}$, of a material can be calculated from its strength coefficient, K, and strain hardening exponent, n. Assume that a material has a true stress–true strain curve given by

$$\sigma = 700\epsilon^{0.5} \text{ MPa}$$

Find: Calculate the true ultimate tensile strength and the engineering S_{ut} of this material.

Solution: Recall that the necking strain corresponds to the maximum load; thus, the necking strain for this material is

$$\epsilon = n = 0.5.$$

Therefore, the *true* ultimate tensile strength is

$$\sigma = Kn^n = 700(0.5)^{0.5} = 495 \text{ MPa}.$$

The true area at the onset of necking is obtained from

$$\ln\left(\frac{A_o}{A_{\text{neck}}}\right) = n = 0.5.$$

Thus,

$$A_{\text{neck}} = A_o \epsilon^{-0.5},$$

and the maximum load, P, is

$$P = \sigma A_{\text{neck}} = \sigma A_o e^{-0.5},$$

where σ is the true ultimate tensile strength. Hence,

$$P = (495)(0.606)(A_o) = 300 A_o \text{ MN}$$

Since $S_{ut} = P/A_o$,

$$S_{ut} = 300 \text{ MPa}$$

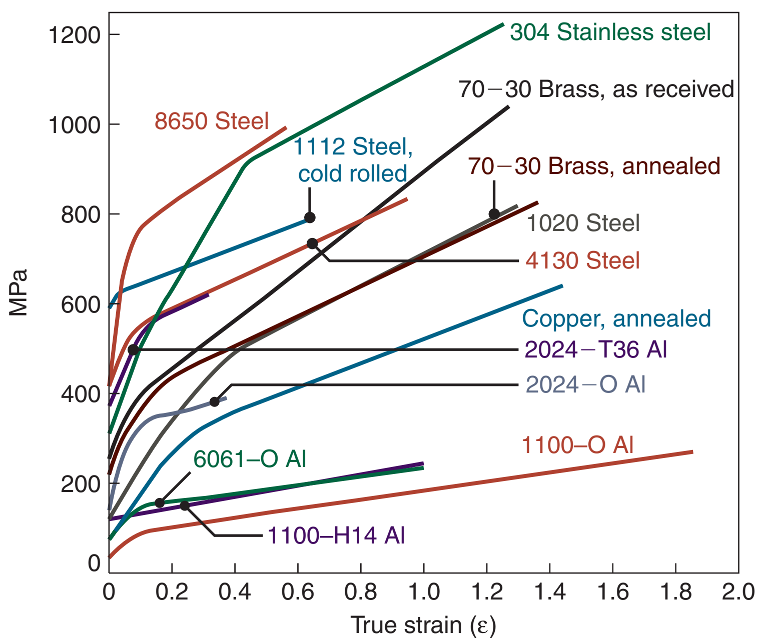


Figure 2.5: True stress–true strain curves in tension at room temperature for various metals. The curves start at a finite level of stress: The elastic regions have too steep a slope to be shown in this figure; thus, each curve starts at the yield strength, S_y, of the material.

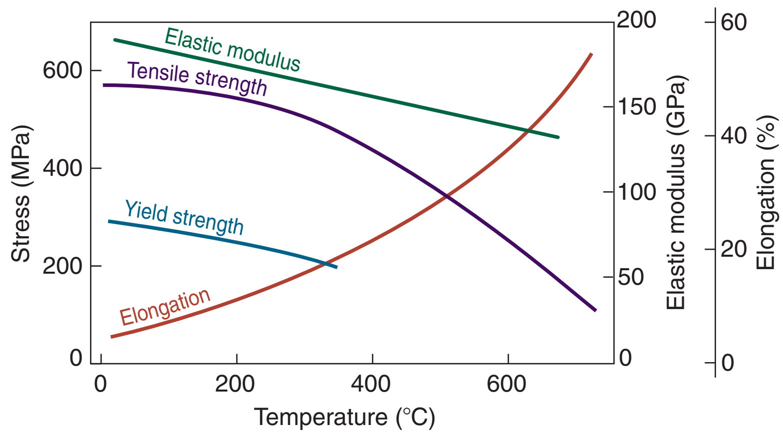


Figure 2.6: Effect of temperature on mechanical properties of carbon steel. Most materials display a similar temperature sensitivity for elastic modulus, yield strength, ultimate strength, and ductility.

2.2.6 Temperature Effects

Increasing the temperature generally has the following effects on stress–strain curves (Fig. 2.6):

a) The ductility and toughness increase

b) The yield strength and the modulus of elasticity decrease.

Temperature has a major influence on the magnitude of the strain-hardening exponent, n, for most metals, in that it increases with increasing temperature. However, this behavior is best described in conjunction with the rate of deformation, because increasing strain rate tends to decrease n.

2.2.7 Effects of Rate of Deformation and Strain Rate

Just as a balloon can be inflated or a rubber band stretched at different rates, materials in manufacturing processes can be shaped at different speeds. Some machines, such as hydraulic presses, form materials at low speeds, while others, such as mechanical presses, form them at high speeds.

The **deformation rate** in a tension test is the speed at which the specimen is being stretched, in units such as m/s. The **strain rate**, on the other hand, is a function of the specimen's length. For example, consider two rubber bands, one 20 mm and the other 100 mm long, respectively, that are stretched by 10 mm within a period of one second. The engineering strain in the shorter specimen is $\frac{10}{20} = 0.5$; the strain in the longer is $\frac{10}{100} = 0.1$. Thus, the strain rates are 0.5 s^{-1} and 0.1 s^{-1}, respectively. Although they are both being stretched at the same deformation rate, the short one is being stretched at a strain rate five times higher than that for the long one.

Deformation rates typically employed in various testing and metalworking processes, and the true strains involved, are given in Table 2.4. Because of the wide ranges encountered in practice, strain rates are generally stated in terms of orders of magnitude, such as 10^2 s^{-1}, 10^4 s^{-1}, and so on.

The typical effects that temperature and strain rate jointly have on the strength of metals are shown in Fig. 2.7. Note that increasing the strain rate increases the strength of the material, called **strain-rate hardening**. The slope of these curves is the **strain-rate sensitivity exponent**, m. The magnitude of m is determined from log–log plots, provided that the vertical and horizontal scales are the same (unlike those shown in Fig. 2.7). A slope of 45°, for example, would indicate a value of $m = 1$. The relationship is given by the equation

$$\sigma = C\dot{\epsilon}^m \tag{2.9}$$

Table 2.4: Typical Ranges of Strain and Deformation Rate in Manufacturing Processes.

Process	True Strain	Deformation rate (m/s)
Cold working		
Forging, rolling	0.1–0.5	0.1–100
Wire and tube drawing	0.05–0.5	0.1–100
Explosive forming	0.05–0.2	10–100
Hot working and warm working		
Forging, rolling	0.1–0.5	0.1–30
Extrusion	2–5	0.1–1
Machining	1–10	0.1–100
Sheet-metal forming	0.1–0.5	0.05–2
Superplastic forming	0.2–3	10^{-4}–10^{-2}

where C is the **strength coefficient** and $\dot{\epsilon}$ is the *true strain rate*, defined as the true strain that the material undergoes per unit time. Note that C has the units of stress and is similar to, but not to be confused with, the strength coefficient K in Eq. (2.8).

From Fig. 2.7, it can be seen that the sensitivity of strength of the material to strain rate increases with temperature; in other words, m increases with increasing temperature. Also note that the slope is relatively flat at room temperature; that is, m is very low. This condition holds for most metals, but not for those that recrystallize at room temperature, such as lead and tin. Typical ranges of m for metals are up to 0.05 for cold-working, 0.05 to 0.4 for hot-working, and 0.3 to 0.85 for super-plastic materials (see below).

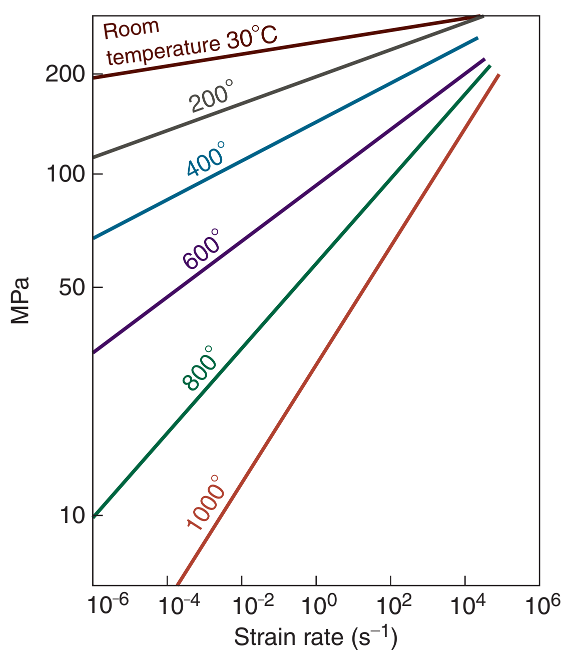


Figure 2.7: The effect of strain rate on the ultimate tensile strength for aluminum. Note that, as the temperature increases, the slopes of the curves increase; thus, strength becomes more and more sensitive to strain rate as temperature increases. *Source:* J.H. Hollomon.

The magnitude of the strain-rate sensitivity exponent significantly influences necking in a tension test. With increasing m, the material stretches further before it fails; thus, increasing m delays necking. Ductility enhancement caused by the high strain-rate sensitivity of some materials has been exploited in **superplastic forming** of sheet metal, as described in Section 16.10.

Superplasticity. The term *superplasticity* refers to the capability of some materials to undergo large uniform elongations prior to necking and to fracture in tension. The elongation may range from a few hundred percent to as much as 2000%. Common nonmetallic materials exhibiting superplastic behavior are bubble gum, glass at elevated temperatures (Section 8.4), and thermoplastics (Section 7.3). Thus, glass and thermoplastics, as examples, can successfully be formed into a wide variety of complex shapes. Among metals exhibiting superplastic behavior are very fine grained (10 to 15 μm) titanium alloys and alloys of zinc–aluminum; when heated, they can elongate to several times their original length.

2.2.8 Hydrostatic Pressure Effects

A variety of tests can be performed to determine the effect of hydrostatic pressure on the mechanical properties of materials. Tests at pressures up to 3.5 GPa indicate that increasing the hydrostatic pressure substantially increases the strain at fracture, both for ductile and for brittle materials. The beneficial effect of hydrostatic pressure has been exploited in metalworking processes, especially in hydrostatic extrusion (Section 15.4.2) and in compaction of metal powders (Section 17.3).

2.2.9 Radiation Effects

In view of the use of various metals and alloys in nuclear applications, extensive studies have been conducted on radiation's effects on mechanical properties. Typical changes in the properties of steels and other metals exposed to doses of high radiation are increased yield strength, tensile strength, hardness, and decreased ductility and toughness.

2.3 Compression

Numerous metalworking processes in manufacturing, such as forging, rolling, and extrusion (Part III), are performed whereby the workpiece is subjected to compressive forces. The **compression test**, in which the specimen is subjected to a compressive load, gives information that is essential in estimating forces and power requirements in these processes. The test is usually carried out by compressing a solid cylindrical specimen between two well-lubricated flat dies (*platens*). Because of friction between the test specimen and the platens, the specimen's cylindrical surface bulges, an effect called **barreling** (see Fig. 2.8). The height-to-diameter ratio of the specimen should be typically less than 3:1 in order to avoid buckling during the test (see also Section 14.4 on *heading*).

Because of barreling, the specimen's cross-sectional area varies along its height, and thus developing stress–strain curves in compression can be challenging. Furthermore, since friction dissipates energy (Section 33.4), the compressive force is higher than it otherwise would be in order to overcome friction. With effective lubrication, friction can be minimized and thus a reasonably constant cross-sectional area can be maintained during the test.

When the results of compression and tension tests on ductile metals are compared, the true stress–true strain curves coincide. This behavior, however, is not the case for *brittle* materials, which are generally much stronger and more ductile in compression than in tension (see Table 8.2).

If a specimen is subjected first to tension, deformed plastically, the load is released, and then a compressive load is applied, the yield strength in compression is found to be lower than that in tension, a behavior known as the **Bauschinger effect** (after J. Bauschinger, reported in 1881). This behavior is exhibited to varying degrees by all metals and alloys. The phenomenon is also called **strain softening** or **work softening**, because of the lowered yield strength in the direction opposite to that of the original load application.

Figure 2.8: Barreling in compressing a round solid cylindrical specimen (7075-O aluminum) between flat dies. Barreling is caused by friction at the die–specimen interfaces, which retards the free flow of the material (see also Fig. 14.3).

Disk Test. For *brittle* materials such as ceramics and glasses (Chapter 8), the **disk test** can be used, in which a disk is subjected to a diametral compression force between two hardened flat platens (Fig. 2.9). When the specimen is loaded as shown, tensile stresses are developed perpendicular to the vertical centerline along the disk. Fracture then initiates and the disk splits vertically in half. The *tensile stress*, σ, in the disk is uniform along the centerline; it can be calculated from the formula

$$\sigma = \frac{2P}{\pi dt}, \tag{2.10}$$

where P is the load at fracture, d is the diameter of the disk, and t is its thickness. In order to avoid premature failure at the contact points, thin strips of soft metal are placed between the disk and the two platens. The strips also protect the platens from being damaged during the test. The phenomenon of fracture at the centerline of the specimen has been utilized in making *seamless tubing* (Section 13.5).

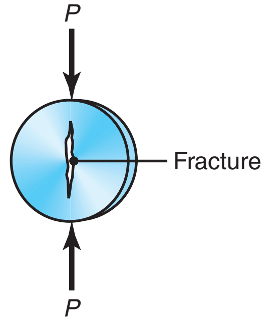


Figure 2.9: Disk test on a brittle material, showing the direction of loading and the fracture path.

2.4 Torsion

In addition to undergoing tension and compression, a workpiece may be subjected to shear strains (Fig. 2.10), such as in the punching of holes in sheet metals (Section 16.2), in swaging (Section 14.4), and in machining operations (Section 21.2). The method generally used to directly determine properties of materials in **shear** is the **torsion test**. This test is usually performed on a thin tubular specimen, in order to develop an approximately uniform stress and strain distribution along its cross section.

A torsion test specimen typically has a reduced cross section in order to confine the deformation to a narrow zone. The **shear stress** can then be calculated from the formula

$$\tau = \frac{T}{2\pi r^2 t}, \tag{2.11}$$

where T is the torque applied, r is the average radius of the tube, and t is the thickness of the tube at its narrow cross section.

The **shear strain** can be calculated from the formula

$$\gamma = \frac{r\phi}{l}, \tag{2.12}$$

where l is the length of the tube section and ϕ the **angle of twist** in radians.

The ratio of the shear stress to the shear strain in the elastic range is known as the **shear modulus** or **modulus of rigidity**, G. The modulus, G, is a quantity related to the modulus of elasticity, E, by the formula

$$G = \frac{E}{2(1+\nu)}. \tag{2.13}$$

The angle of twist, ϕ, to fracture in the torsion of solid round bars at elevated temperatures has been found to be useful in estimating the forgeability of metals (Section 14.5). The greater the number of twists prior to failure, the better is the forgeability.

2.5 Bending (Flexure)

Preparing specimens from brittle materials can be difficult because of the challenges involved in shaping, machining, and finishing them to appropriate final dimensions. Furthermore, the specimens are typically sensitive to surface defects (such as scratches and notches), and clamping brittle specimens for testing can

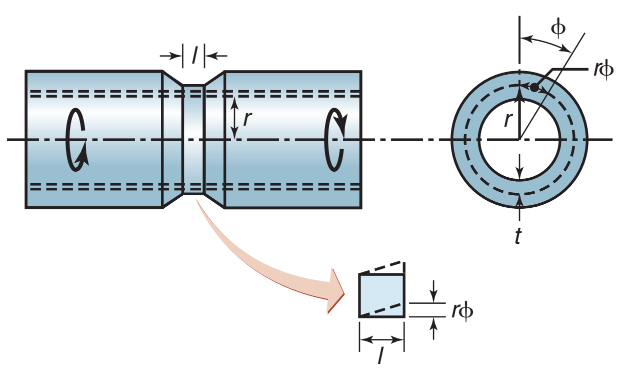


Figure 2.10: A typical torsion-test specimen, mounted between the two heads of a testing machine and twisted. Note the shear deformation of an element in the reduced section of the specimen.

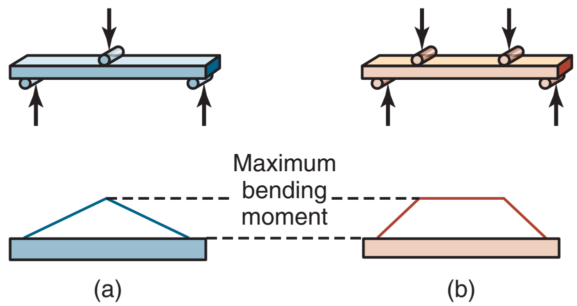


Figure 2.11: Two bend-test methods for brittle materials: (a) three-point bending, and (b) four-point bending. The two areas shown above the beams represent the bending-moment diagrams, described in texts on the mechanics of solids. Note the region of constant maximum bending moment in (b); by contrast, the maximum bending moment occurs only at the center of the specimen in (a).

be difficult. Improper alignment of the test specimen also can result in nonuniform stress distribution along its cross section.

A commonly used test method for brittle materials is the **bend** or **flexure test**; it usually involves a specimen that has a rectangular cross section and supported in the manner shown in Fig. 2.11. The load is applied vertically, at either one point or two points. Consequently, these tests are referred to as **three-point** and **four-point bending**, respectively. The longitudinal stresses in the specimens are tensile at their lower surfaces and compressive at their upper surfaces. The stresses developed can be calculated using simple beam equations, described in texts on the mechanics of solids.

The stress at fracture in bending is known as the **modulus of rupture** or **transverse rupture strength** (see Table 8.2). Note that, because of the larger volume of material subjected to the same bending moment in Fig. 2.11b, there is a higher probability that defects exist within this volume than exist in the point shown in Fig. 2.11a. Consequently, the four-point test predicts a lower modulus of rupture than the three-point test.

2.6 Hardness

Hardness is generally defined as *resistance to permanent indentation*; thus, steel is harder than aluminum, and aluminum is harder than lead. Hardness is not a fundamental property, because the resistance to indentation depends on the shape of the indenter and on the load applied. Hardness is a commonly used property; it gives a general indication of the strength of the material and of its resistance to scratching and to wear.

2.6.1 Hardness Tests

Several test methods, using different indenter materials and shapes (Fig. 2.12), have been developed to measure the hardness of materials. The most commonly used hardness tests are described next.

Brinell Test. Introduced by J.A. Brinell, in 1900, this test involves pressing a steel or tungsten-carbide ball, 10 mm in diameter, against a surface (Fig. 2.13). The *Brinell hardness number* (HB) is defined as the ratio of the applied load, P, to the curved surface area of the indentation. The harder the material tested, the smaller is the impression; a 1500 kg or 3000 kg load is usually recommended in order to obtain impressions sufficiently large for accurate measurement of hardness.

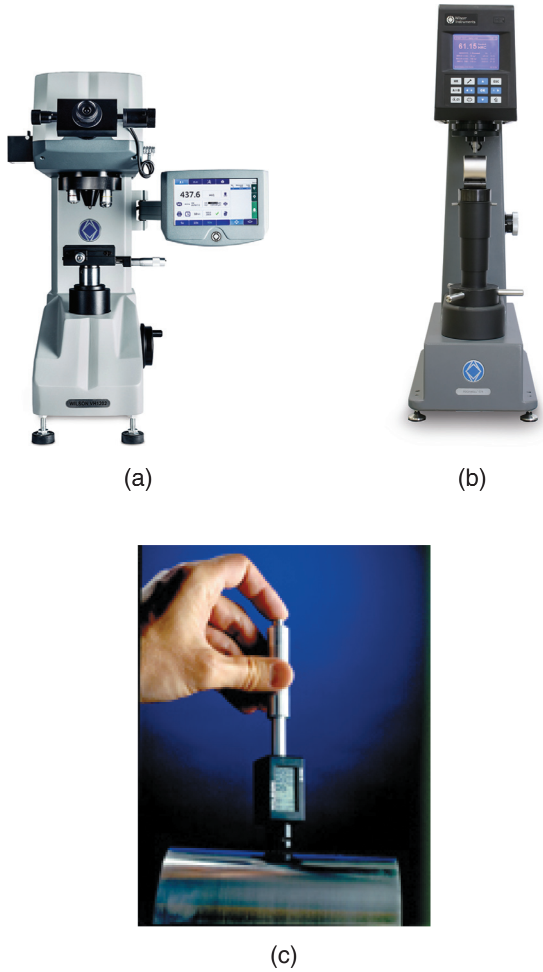


Figure 2.12: A selection of hardness testers. (a) A Micro Vickers hardness tester; (b) Rockwell hardness tester; (c) Leeb tester. *Source:* (a) and (b) Courtesy of Buehler (c) Courtesy of Wilson® Instruments.

Depending on the surface condition of the material tested, two types of impression develop (Fig. 2.14). In annealed metals, the impression generally has a rounded profile along its periphery (Fig. 2.14a). In cold-worked metals, they usually have a sharp profile (Fig. 2.14b). The correct method of measuring the indentation diameter, d, is shown in the figure.

The indenter has a finite elastic modulus, hence it undergoes *elastic* deformation under the applied load. As a result, hardness measurements may not be as accurate as expected, depending on the indenter material. One method for minimizing this effect is using tungsten-carbide balls (Section 22.4), which, because of their higher modulus of elasticity, distort less than steel balls do. These indenters are usually recommended for materials with a Brinell hardness number higher than 500.

Rockwell Test. Developed by S.P. Rockwell, in 1922, this test measures the *depth* of penetration instead of the diameter of the indentation. The indenter is pressed onto the surface, first with a *minor load* and

Test	Indenter	Shape of indentation: Side view	Shape of indentation: Top view	Load, P	Hardness number
Brinell	10-mm steel or tungsten-carbide ball	D, d	d	500 kg 1500 kg 3000 kg	$HB = \frac{2P}{(\pi D)\left(D - \sqrt{D^2 - d^2}\right)}$
Vickers	Diamond pyramid	136°	L	1–120 kg	$HV = \frac{1.854P}{L^2}$
Knoop	Diamond pyramid	t $L/b = 7.11$ $b/t = 4.00$	b, L	25 g–5 kg	$HK = \frac{14.2P}{L^2}$
Rockwell A, C, D	Diamond cone	120° $t = $ mm		60 kg 150 kg 100 kg	HRA, HRC, HRD $= 100 - 500t$
B, F, G	1.6-mm diameter steel ball	$t = $ mm		100 kg 60 kg 150 kg	HRB, HRF, HRG, HRE $= 130 - 500t$
E	3.2-mm diameter steel ball			100 kg	

Figure 2.13: General characteristics of hardness-testing methods and formulas for calculating hardness.

then with a *major load*. The difference in the depths of penetration is a measure of the hardness of the material. Figure 2.13 shows some of the common Rockwell hardness scales for engineering materials and the indenters used. Rockwell **superficial hardness** tests, using the same type of indenters but at lighter loads, also are available.

Vickers Test. This test, developed in 1922 and previously known as the *diamond pyramid hardness* test, uses a pyramid-shaped diamond indenter (Fig. 2.13) and at a load that ranges from 1 kg to 120 kg. The *Vickers hardness number* is indicated as HV. The impressions are typically less than 0.5 mm on the diagonal. This test gives essentially the same hardness number regardless of the load, and is suitable for testing materials with a wide range of hardness, including heat-treated steels. Test procedures have been developed to perform tests using atomic force microscopes and nanoindenters, where indentation depths are as low as 20 nm.

Knoop Test. This test, developed by F. Knoop, in 1939, uses a diamond indenter in the shape of an elongated pyramid (Fig. 2.13). The applied load generally ranges from 25 g to 5 kg. The *Knoop hardness number* is indicated as HK. Because of the light loads applied, this test is a **microhardness** test, and is suitable for very small or very thin specimens, and for such brittle materials as carbides, ceramics, and glass.

The Knoop test is also used for measuring the hardness of individual grains or components in a metal alloy. Because the size of the indentation is typically in the range of 0.01–0.10 mm, surface preparation is

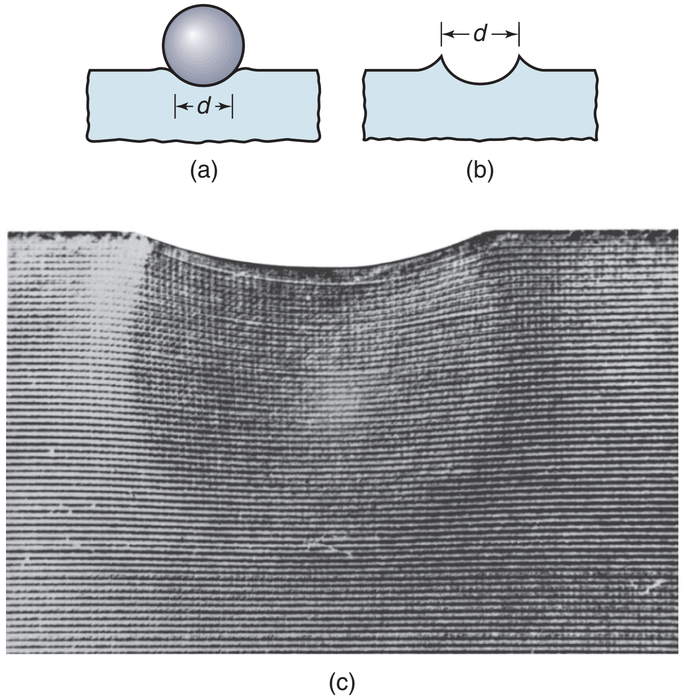


Figure 2.14: Indentation geometry in Brinell hardness testing: (a) annealed metal; (b) work-hardened metal; (c) deformation of mild steel under a spherical indenter. Note that the depth of the permanently deformed zone is about one order of magnitude larger than the depth of indentation. For a hardness test to be valid, this zone should be fully developed in the material. *Source:* After M.C. Shaw and C.T. Yang.

important. The hardness number obtained depends on the applied load, therefore test results should always cite the load employed.

Scleroscope and Leeb Tests. The *scleroscope* (from the Greek *skleros*, meaning hard) is an instrument in which a diamond-tipped indenter (*hammer*), enclosed in a glass tube, is dropped onto the specimen from a certain height. The hardness is related to the *rebound* of the indenter: the higher the rebound, the harder is the material tested. The impression made by a scleroscope is very small. Since reliable results with a scleroscope can be difficult to obtain, an electronic version, called a **Leeb**, or Equotip, test, has been developed (Fig. 2.12d). A carbide hammer impacts the surface, and the incident and the rebound velocities are electronically measured. The *Leeb number* is then calculated and is usually converted to Rockwell or Vickers hardness.

Mohs Hardness. Developed in 1822 by F. Mohs, this test is based on the capability of one material to scratch another. The Mohs hardness number is based on a scale from 1 to 10, with 1 being the measure for talc and 10 for diamond (the hardest substance known; see also Section 8.7). Thus, a material with a higher Mohs hardness number always scratches the one with a lower number. Soft materials typically have a number between 2 and 3, hardened steels about 6, and aluminum oxide (used in cutting tools and abrasives) of 9. Although the Mohs scale is qualitative and is used mainly by mineralogists, it correlates well with Knoop hardness.

Shore Test. The hardness of such materials as rubbers, plastics, and soft and elastic nonmetallic materials is generally measured by a Shore test, with an instrument called a **durometer** (from the Latin *durus*, meaning hard). An indenter is first pressed against the surface and a constant load is rapidly applied. The *depth* of penetration is then measured after one second. There are two different scales for this test. For Type A, a blunt indenter is used at an applied load of 1 kg. This method is typically used for softer materials. Type D has a sharper indenter and at a load of 5 kg; it is used for harder materials. The hardness numbers in these two tests range from 0 to 100.

Hot Hardness. The hardness of materials at elevated temperatures (see Fig. 22.1) is an important factor in such applications as cutting tools and for dies in hot-working and casting operations. Hardness tests can be performed using conventional testers, with some modifications such as enclosing the specimen and indenter in a small electric furnace.

2.6.2 Hardness and Strength

Because hardness is the resistance to *permanent* indentation, it can be likened to performing a compression test on a small area on a material's surface (Fig. 2.14c). It has been shown that the hardness of a cold-worked metal, for example, is about three times its yield strength S_y (using the same units). For annealed metals, the hardness is about five times S_y.

A relationship has been established between the ultimate tensile strength (S_{ut}) and the Brinell hardness number (HB) for steels, measured for a load of 3000 kg. In SI units, the relationship is given by

$$S_{ut} = 3.5\,(\text{HB}), \tag{2.14}$$

where S_{ut} is in MPa.

2.6.3 Hardness-testing Procedures

For a hardness test to be reliable, the **zone of deformation** under the indenter (Fig. 2.14c) must be allowed to develop freely. Consequently, the *location* of the indenter (with respect to the location of the *edges* of the specimen to be tested) and the *thickness* of the specimen are important considerations. Generally, the location should be at least two diameters of the indentation from the edge of the specimen, and the thickness of the specimen should be at least 10 times the depth of penetration of the indenter. Successive indentations on the same surface of the workpiece should be far enough apart so as not to interfere with each other.

The indentation should be sufficiently large to give a representative hardness value for the bulk material. If (a) hardness variations, if any, must be detected in a small surface area, or (b) the hardness of individual constituents in a matrix or in an alloy is to be determined, the indentations must be very small (such as those obtained in Knoop or Vickers tests, using light loads). While *surface preparation* is not critical for the Brinell test, it is important for the Rockwell test and even more important for the other hardness tests, because of the small sizes of the indentations. Surfaces may have to be polished to allow correct measurement of the impression's dimensions.

The hardness values obtained from different tests can be interrelated, and converted, using Fig. 2.15. Care should be exercised in using these charts because of the variables involved in material characteristics and in the shape of the indentation used.

Example 2.2 Calculation of Modulus of Resilience from Hardness

Given: A piece of steel is highly deformed at room temperature. Its hardness is found to be 300 HB.

Find: Estimate the area of the elastic portion of the stress–strain curve up to the yield point (that is, the *resilience*) for this material if the yield strength is one-third the Brinell hardness.

Solution: Since the steel has been subjected to large strains at room temperature, it may be assumed that its stress–strain curve has flattened considerably, thus approaching the shape of a perfectly plastic curve. Since the yield strength is one-third the Brinell hardness,

$$S_y = \frac{300}{3} = 100 \text{ kg/mm}^2 = 981 \text{ MPa}.$$

The area under the stress–strain curve is

$$\text{Modulus of resilience} = \frac{S_y^2}{2E}.$$

From Table 2.2, $E = 210$ GPa for steel. Hence,

$$\text{Modulus of resilience} = \frac{(981 \times 10^6)^2}{2\,(210 \times 10^9)} = 2.29 \text{ MNm/m}^3.$$

2.7 Fatigue

The components of manufacturing equipment, such as tools, dies, gears, cams, shafts, and springs, are often subjected to rapidly fluctuating (cyclic or periodic) loads, in addition to static loads. **Cyclic stresses** develop by fluctuating mechanical loads, such as (a) on gear teeth or in reciprocating sliders, (b) by rotating machine elements under constant bending stresses, as is commonly encountered in shafts, or (c) by thermal stresses, as when a die at room temperature comes into repeated contact with hot workpieces, and then begins to cool down between successive contacts. Under any of these conditions, the component may fail at a stress level *below* that at which failure would occur under static loading. Upon inspection, failure is found to be associated with cracks that develop and grow with every stress cycle. The cracks propagate through the part until a critical crack length is reached and the part fractures. Known as **fatigue failure**, this phenomenon is responsible for the majority of failures in mechanical components.

Fatigue test methods involve testing specimens under a variety of states of stress, usually in a combination of tension and bending. The test is carried out at various *stress amplitudes* (S); the number of cycles (N) to cause total failure of the specimen or part is then recorded. Stress amplitude is defined as the maximum stress, in tension and compression, to which the specimen is subjected.

Typical plots, called ***S-N* curves**, are shown in Fig. 2.16. These curves are based on complete reversal of the stress—that is, maximum tension, then maximum compression, then maximum tension, and so on—such as that imposed by bending a piece of wire alternately in one direction and then the other. Tests may also be performed on a rotating shaft in four-point bending (Fig. 2.11b.) With some materials, the S-N curve becomes horizontal at low stress levels; the maximum stress to which the material can be subjected without fatigue failure, regardless of the number of cycles, is known as the **endurance limit** or **fatigue limit**.

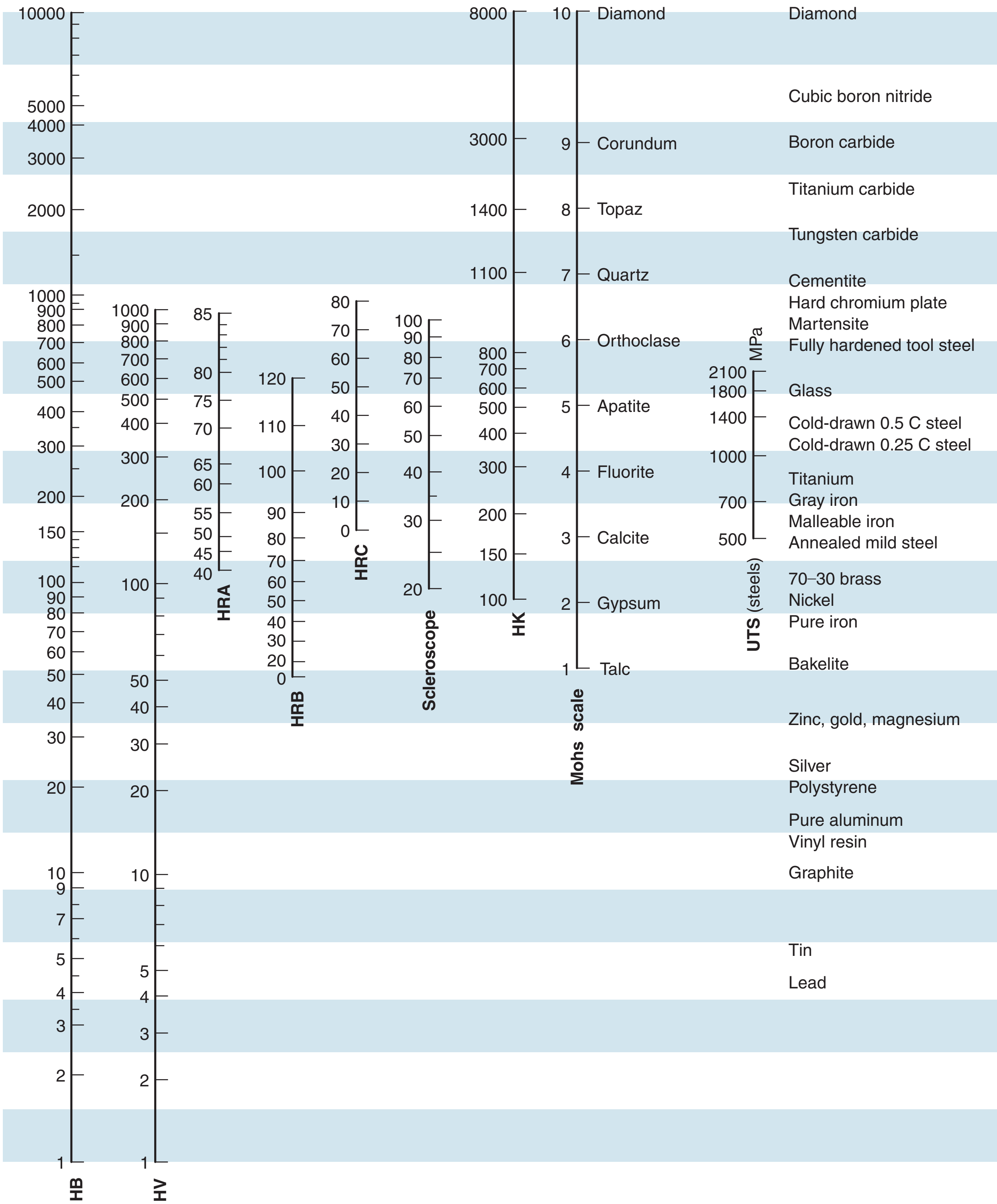


Figure 2.15: Chart for converting various hardness scales; note the limited range of most of the scales. Because of the many factors involved, these conversions are approximate.

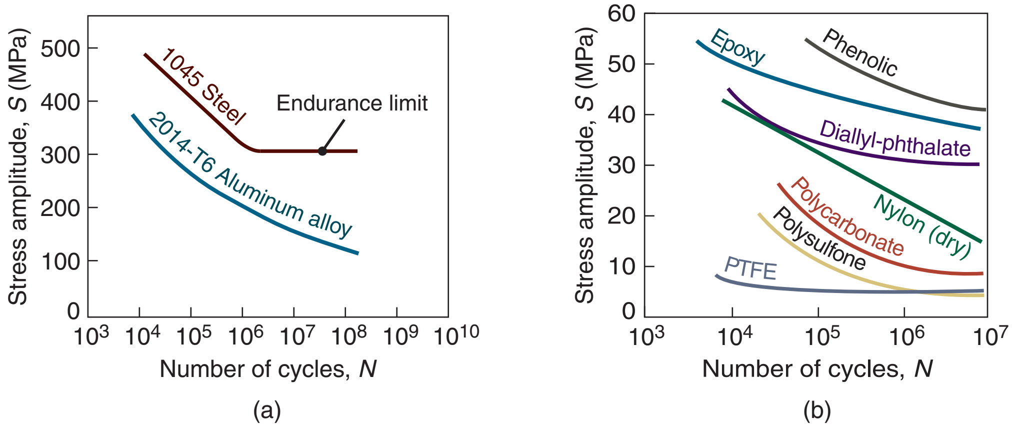


Figure 2.16: (a) Typical S-N curves for two metals. Note that, unlike steel, aluminum does not have an endurance limit. (b) S-N curves for common polymers.

Although several materials, especially steels, have a specific endurance limit, others, especially aluminum alloys, do not have such a limit and the S-N curve continues its downward trend. For metals exhibiting such behavior, the fatigue strength is specified at a certain number of cycles, such as 10^7. The useful service life of the component can then be specified. The endurance limit for metals can be approximately related to their ultimate tensile strength (Fig. 2.17). For carbon steels, for example, the endurance limit is usually 0.4–0.5 times the tensile strength.

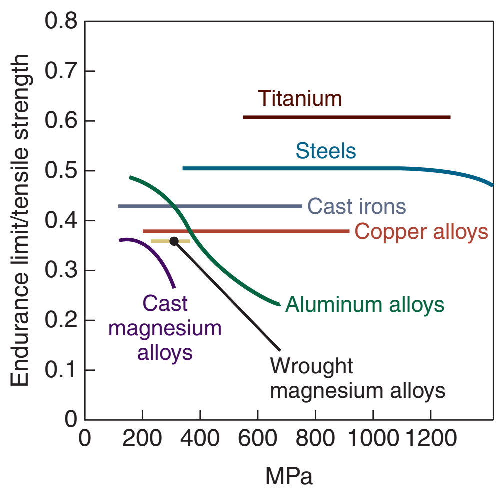


Figure 2.17: Ratio of endurance limit to tensile strength for various metals, as a function of tensile strength. Because aluminum does not have an endurance limit, the correlations for aluminum are based on a specific number of cycles, as is seen in Fig. 2.16.

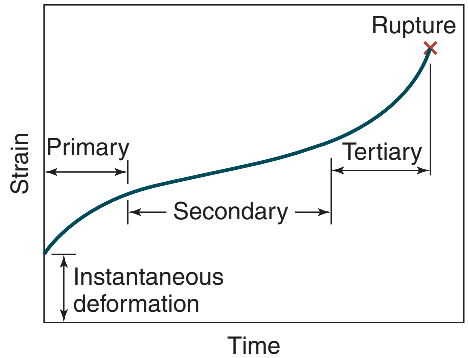


Figure 2.18: Schematic illustration of a typical creep curve. The linear segment of the curve (secondary) is used in designing components for a specific creep life.

2.8 Creep

Creep is the permanent elongation of a material under a static load maintained for a period of time. This phenomenon occurs in metals as well as nonmetallic materials, such as thermoplastics and rubbers, and it can occur at any temperature; lead, for example, creeps under a constant tensile load at room temperature. For metals and their alloys, creep of any significance occurs at elevated temperatures, beginning at about 200°C for aluminum alloys and at about 1500°C for refractory alloys.

The mechanism of creep at elevated temperature in metals is generally attributed to **grain-boundary sliding** (see Section 1.5). It is especially important in high-temperature applications, such as gas-turbine blades and various components in jet engines and rocket motors. High-pressure steam lines, nuclear-fuel elements, and furnace components are likewise subject to creep. Creep can also occur in tools and dies that are subjected to high stresses at elevated temperatures during hot-working operations, such as forging and extrusion.

The *creep test* typically consists of subjecting a specimen to a constant tensile load (hence constant engineering stress) at elevated temperature and measuring the changes in length at various time increments. A creep curve typically consists of *primary, secondary,* and *tertiary stages* (Fig. 2.18). During the test, the specimen eventually fails by necking and fracture, called **rupture** or *creep rupture*. As expected, the creep rate increases with specimen temperature and the applied load.

Design against creep usually requires knowledge of the secondary (linear) range and its slope, because the creep rate can be determined reliably only when the curve has a constant slope. Resistance to creep generally increases with the melting temperature of the material. Thus, stainless steels, superalloys, and refractory metals and their alloys are commonly used in applications where resistance to creep is required.

Stress Relaxation. Closely related to creep, the stresses resulting from external loading of a structural component decrease in magnitude over a period of time, even though the dimensions of the component remain constant. A typical example is the decrease in tensile stress of a wire in tension between two fixed points (as in the wires in a piano or a violin). Other examples include stress relaxation in rivets, bolts, guy wires, and various similar parts, either under tension, compression, or flexure. Stress relaxation is particularly common in thermoplastics (Section 7.3) and at a microscale in hot isostatic pressing (Section 17.3.2).

2.9 Impact

In numerous machinery components and manufacturing operations, materials are subjected to **impact** or **dynamic loading**, such as in the heading operation for nails and bolt heads (Section 14.4). A typical *impact*

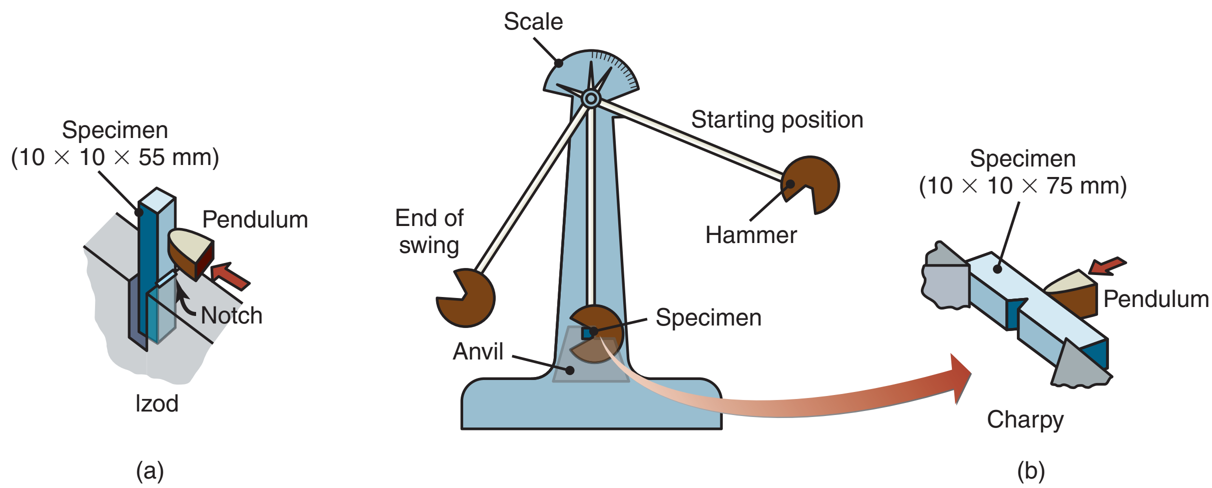


Figure 2.19: Impact test specimens. (a) Izod; (b) Charpy.

test for determining high-strain-rate properties of materials consists of placing a notched specimen in an impact tester and breaking it with a swinging pendulum (Fig. 2.19).

In the **Charpy test**, the specimen is supported at both ends, while in the **Izod test** it is supported at one end, as in a cantilever beam. From the swing of the pendulum, the energy dissipated in breaking the specimen can be obtained and is known as the **impact toughness** of the material. Unlike hardness-test conversions (Fig. 2.15), no quantitative relationships have yet been established between the Charpy and the Izod tests. Impact tests are particularly useful in determining the ductile–brittle transition temperature of materials (Section 2.10.1).

Materials that have high impact resistance generally also have high strength and high ductility, hence high toughness. Sensitivity of materials to surface defects (**notch sensitivity**) is an important factor, as it can significantly lower impact toughness, particularly in heat-treated metals and in ceramics and glasses.

2.10 Failure and Fracture of Materials

Failure is one of the most important aspects of material behavior, because it directly influences material selection and the method(s) of manufacture, and determines the service life of a component. Because of the many factors involved, failure and fracture of materials is a complex area of study. This section focuses only on those aspects of failure that are of particular significance to selecting and processing materials. There are two general types of failure:

1. **Fracture**, through either internal or external cracking. Fracture is further subclassified into two general categories: *ductile* and *brittle* (Figs. 2.21 and 2.22).

2. **Buckling**, as shown in Fig. 2.20b.

Although failure of materials is generally regarded as undesirable, some products are designed so that failure is essential for their proper function. Typical examples are (a) beverage or food containers, with pop tops which are opened by shearing the sheet metal along a scored profile; (b) shear pins on shafts, to prevent damage of machinery in the case of overloads; (c) perforated sheet, to ease tearing along a specific path, as in packaging; and (d) metal or plastic screw caps for beverage bottles, to ease their removal.

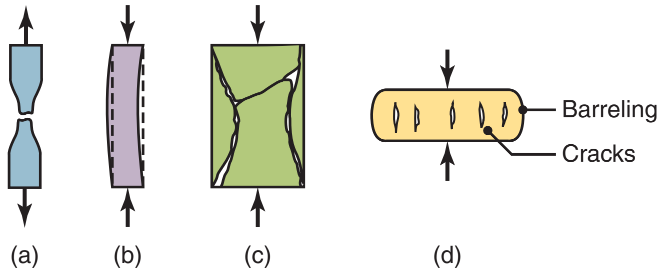


Figure 2.20: Schematic illustration of types of failures in materials: (a) necking and fracture of ductile materials; (b) buckling of ductile materials under a compressive load; (c) fracture of brittle materials in compression; (d) cracking on the barreled surface of ductile materials in compression.

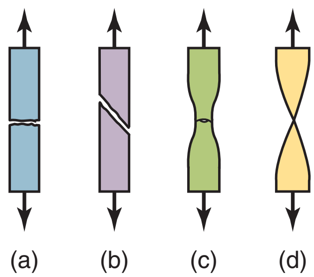


Figure 2.21: Schematic illustration of the types of fracture in tension: (a) brittle fracture in polycrystalline metals; (b) shear fracture in ductile single crystals—see also Fig. 1.6a; (c) ductile cup-and-cone fracture in polycrystalline metals; (d) complete ductile fracture in polycrystalline metals, with 100% reduction of area.

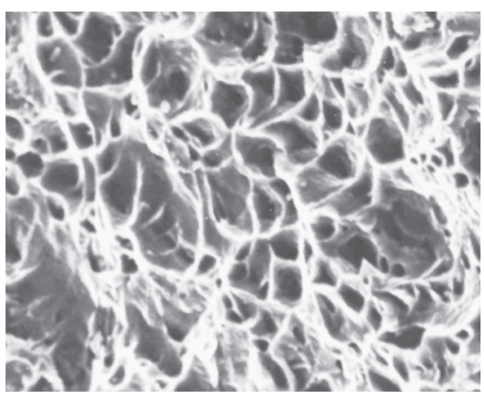


Figure 2.22: Surface of ductile fracture in low-carbon steel, showing dimples. Fracture is usually initiated at impurities, inclusions, or preexisting voids (microporosity) in the metal. *Source:* After K.-H. Habig and D. Klaffke.

2.10.1 Ductile Fracture

Ductile fracture is characterized by *plastic deformation*, which precedes failure (Fig. 2.20a). In a tension test, highly ductile materials such as gold and lead may neck down to a point before failing (Fig. 2.21d). Most metals and alloys, however, neck down to a finite cross-sectional area and then fail. Ductile fracture typically takes place along planes on which the shear stress is a maximum. In torsion, for example, a ductile metal fractures along a plane perpendicular to the axis of twist; that is, the plane on which the shear stress is a maximum. Fracture in shear, by contrast, is due to extensive slip along slip planes within the grains (see Fig. 1.7).

Close examination of the surface of ductile fracture (Fig. 2.22) shows a *fibrous* pattern with *dimples*, as if a number of very small tension tests have been carried out over the fracture surface. Failure is initiated with the formation of tiny *voids*, usually around small inclusions or preexisting voids, which then *grow* and *coalesce*, developing into microcracks which then continue to grow in size, eventually leading to fracture.

In a tension-test specimen, fracture begins at the center of the necked region, resulting in the growth and coalescence of cavities (Fig. 2.23). The central region thus becomes one large crack, as can be seen in the midsection of the tension-test specimen in Fig. 2.23d. The crack then propagates to the periphery of the necked region and results in total failure. Because of its appearance, the fracture surface of a tension-test ductile specimen is called a **cup-and-cone fracture**.

Effects of Inclusions. Because they are nucleation sites for voids, *inclusions* influence ductile fracture and, consequently, the workability of metals. Inclusions may consist of impurities of various kinds and of second-phase particles, such as oxides, carbides, and sulfides. The extent of their influence depends on such factors as their shape, hardness, distribution, and their fraction of the total volume. The greater the volume fraction of inclusions, the lower will be the ductility of the material.

Voids and porosity can also develop during processing of metals, such as porosity in castings (Section 10.6.1) and metalworking processes, such as drawing and extrusion (Chapter 15). Two factors affect void formation:

1. The strength of the bond at the interface between an inclusion and the matrix. If the bond is strong, there is lower tendency for void formation during plastic deformation.

2. The hardness of the inclusion. If the inclusion is soft, such as manganese sulfide, it will conform to the overall shape change of the workpiece during plastic deformation. If the inclusion is hard (as, for example, in carbides and oxides; see also Section 8.2), it could lead to void formation (Fig. 2.24). Hard and brittle inclusions may also break up into smaller particles during plastic deformation.

The alignment of inclusions during plastic deformation leads to **mechanical fibering** (Section 1.6). Subsequent processing of such a material must therefore involve considerations of the proper direction of working the material in order to develop maximum ductility and strength.

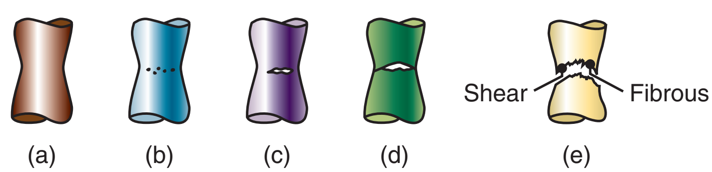


Figure 2.23: Sequence of events in the necking and fracture of a tensile-test specimen: (a) early stage of necking; (b) small voids begin to form within the necked region; (c) voids coalesce, producing an internal crack; (d) the rest of the cross section begins to fail at the periphery, by shearing; (e) the final fracture, known as a cup- (top fracture surface) and-cone- (bottom surface) fracture, surfaces.

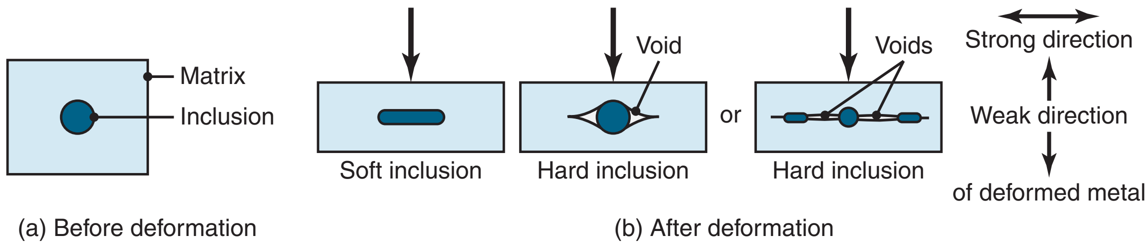


Figure 2.24: Schematic illustration of the deformation of soft and hard inclusions and of their effect on void formation in plastic deformation. Note that, because they do not conform to the overall deformation of the ductile matrix, hard inclusions can cause internal voids.

Transition Temperature. Metals may undergo a sharp change in ductility and toughness across a narrow temperature range, called the *transition temperature* (Fig. 2.25). This phenomenon occurs mostly in body-centered cubic and in some hexagonal close-packed metals; it is rarely exhibited by face-centered cubic metals. The transition temperature depends on such factors as (a) the composition, microstructure, and grain size of the material, (b) the surface finish and the shape of the specimen, and (c) the deformation rate. High rates, abrupt changes in workpiece shape, and the presence of surface notches raise the transition temperature of a material.

Strain Aging. *Strain aging* is a phenomenon in which carbon atoms in steels segregate to dislocations, thereby pinning the dislocations and, in this way, increasing the resistance to their movement. As a result, strength is increased and ductility is reduced. Instead of taking place over several days at room temperature, strain aging can occur in just a few hours at a higher temperature, called *accelerated strain aging*. An example of accelerated strain aging in steels is **blue brittleness**, so named because it occurs in the blue-heat range, where the steel develops a bluish oxide film. Blue brittleness causes a significant decrease in ductility and toughness, and an increase in the strength of plain-carbon and of some alloy steels.

2.10.2 Brittle Fracture

Brittle fracture occurs with little or no significant plastic deformation. In tension, brittle fracture takes place along the crystallographic plane (**cleavage plane**) on which the normal tensile stress is a maximum. Brittle fracture does not usually occur with face-centered cubic metals, but it is not uncommon with body-centered cubic and some hexagonal close-packed metals. In general, low temperature and a high rate of deformation promote brittle fracture.

In a polycrystalline metal under tension, the fracture surface has a bright granular appearance (unlike the fibrous appearance in ductile fracture), because of the changes in the direction of the cleavage planes as

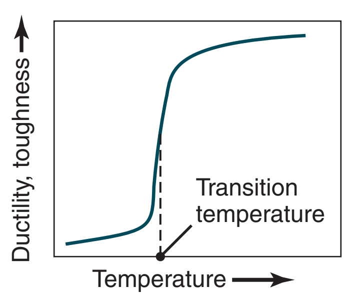


Figure 2.25: Schematic illustration of transition temperature in metals.

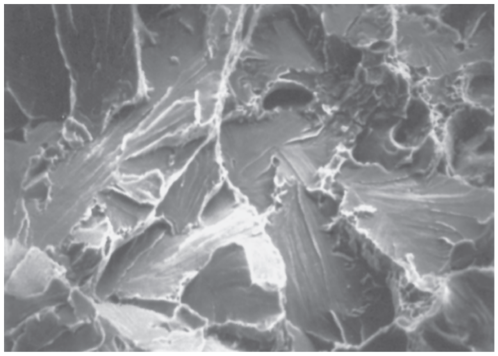


Figure 2.26: Fracture surface of steel that has failed in a brittle manner. The fracture path is transgranular (through the grains). Magnification: $200\times$. *Source:* After B.J. Schulze and S.L. Meiley and Packer Engineering Associates, Inc.

the crack propagates from one grain to another (Fig. 2.26). Brittle fracture, in compression, is more complex; fracture may even follow a path that is theoretically at an angle of 45° to the direction of the applied force.

Examples of fracture along a cleavage plane are the splitting of rock salt and the peeling of layers of mica. With mica, graphite, and other lamellar solids, tensile stresses normal to the cleavage plane, caused by pulling, initiate and control the propagation of fracture. Materials such as chalk, gray cast iron, and concrete fail in tension in the manner shown in Fig. 2.21a. In torsion, they fail along a plane at an angle of 45° to the axis of twist (Fig. 2.10)—that is, along a plane on which the tensile stress is a maximum.

Defects. An important factor in fracture is the presence of *defects*, such as scratches, flaws, and preexisting external or internal cracks. Under tension, the sharp tip of the crack is subjected to high tensile stresses, which then lead the crack to propagate rapidly.

The presence of defects explains why brittle materials exhibit weakness in tension as compared to their strength in compression (see Table 8.2). For example, the ratio of compressive to tensile strength is on the order of 10 for rocks and similar materials, 5 for glass, and 3 for gray cast iron. Under tensile stresses, cracks propagate rapidly, causing what is known as *catastrophic failure*.

In polycrystalline metals, the fracture paths most commonly observed are **transgranular** (**transcrystalline** or **intragranular**); that is, the crack propagates *through* the grain. In **intergranular** fracture, the crack propagates *along* the grain boundaries (Fig. 2.27). It generally occurs when the grain boundaries are soft, contain a brittle phase, or they have been weakened by liquid- or solid-metal embrittlement (Section 1.5.2).

Fatigue Fracture. *Fatigue fracture* generally occurs in a brittle manner and is associated with cyclic loads. Minute external or internal cracks develop at preexisting flaws or at various defects in the material. These cracks then propagate with each load cycle and eventually lead to total and sudden failure of the part. The surface in fatigue fracture is generally characterized by the term **beach marks**, because of its appearance (Fig. 2.28). Under high magnification (typically at more than $1000\times$), a series of **striations** can be seen on fracture surfaces, each beach mark consisting of several striations.

Improving Fatigue Strength. Fatigue life is greatly influenced by the method of surface preparation (Fig. 2.29). The fatigue strength of manufactured products can be improved by the following methods:

1. Inducing compressive residual stresses on surfaces—for example, by shot peening or roller burnishing (Section 34.2).

2. Case hardening (surface hardening) by various means (Section 4.10).

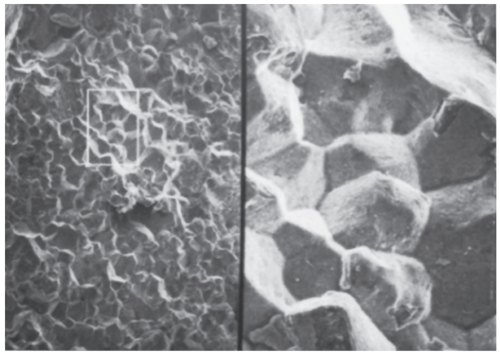


Figure 2.27: Intergranular fracture, at two different magnifications. Grains and grain boundaries are clearly visible in this micrograph. The fracture path is along the grain boundaries. Magnification: left, $100\times$; right, $500\times$. *Source:* After B.J. Schulze and S.L. Meiley and Packer Engineering Associates, Inc.

3. Producing a fine surface finish, thereby reducing the detrimental effects of scratches, notches, and other surface imperfections.
4. Selecting appropriate materials and ensuring that they are free from significant amounts of inclusions, voids, and impurities.

Conversely, the following factors and processes can *reduce* fatigue strength:

1. Tensile residual stresses on the surface (see Section 2.11).
2. Decarburization.
3. Surface pits (such as due to corrosion), that act as stress raiser.
4. Hydrogen embrittlement.
5. Galvanizing.
6. Electroplating.

Stress–Corrosion Cracking. An otherwise ductile metal can fail in a brittle manner by *stress–corrosion cracking* (also called **stress cracking** or **season cracking**). Parts that are free from defects may develop cracks,

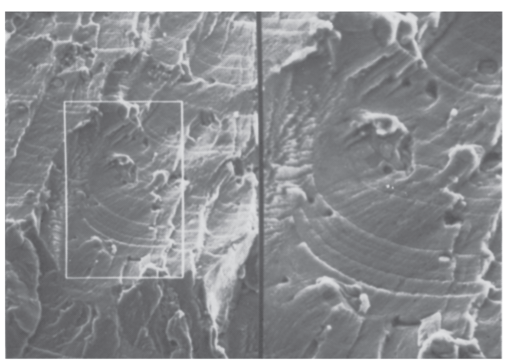


Figure 2.28: Typical fatigue-fracture surface on metals, showing beach marks. Magnification: left, $500\times$; right, $1000\times$. *Source:* After B.J. Schulze and S.L. Meiley and Packer Engineering Associates, Inc.

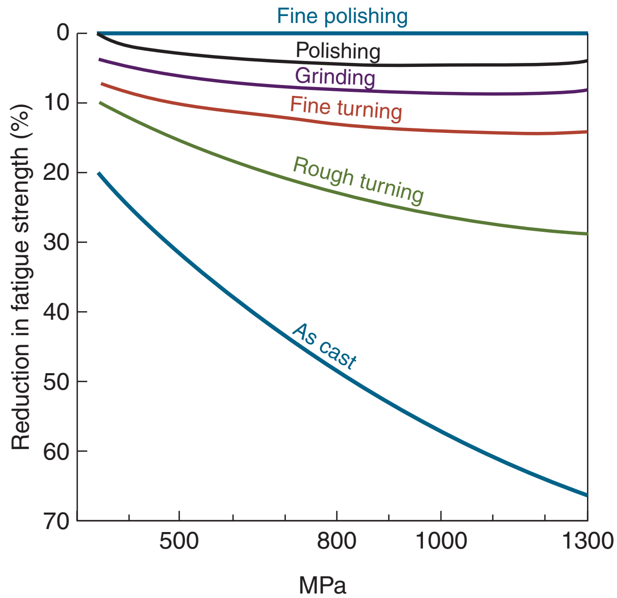


Figure 2.29: Reductions in the fatigue strength of cast steels subjected to various surface-finishing operations. Note that the reduction becomes greater as the surface roughness and the strength of the steel increase. *Source:* M.R. Mitchell.

either over time or even soon after they are made. Crack propagation may be either intergranular or transgranular. The susceptibility of metals to stress–corrosion cracking depends mainly on the material, the presence and magnitude of *tensile residual stresses*, and the environment (such corrosive media as salt water or chemicals).

Brass and austenitic stainless steels are among metals that are highly susceptible to stress cracking. The usual procedure to avoid stress–corrosion cracking is to *stress relieve* the part just after it is formed. Full annealing (Section 4.11) may also be done, but this treatment reduces the strength of cold-worked parts.

Hydrogen Embrittlement. The presence of hydrogen can reduce ductility and can cause severe embrittlement and premature failure in metals and their alloys, as well as in nonmetallic materials. Called *hydrogen embrittlement*, this phenomenon is especially severe in high-strength steels. Possible sources of hydrogen arise during melting of the metal in preparation for casting, pickling (Section 13.3), and electrolysis in electroplating (Section 34.9). Other sources of hydrogen are water vapor in the atmosphere and moisture on electrodes and in fluxes used during welding. Oxygen also can cause embrittlement, particularly in copper alloys.

2.11 Residual Stresses

Residual stresses are stresses that remain within a part after it has been shaped and all the external forces (applied through tools and dies) are removed. Residual stresses also may develop when workpieces are subjected to plastic deformation that is not uniform throughout the part. A typical example is the bending of a metal bar (Fig. 2.30). Note that the external bending moment first produces a stress distribution that varies linearly through the thickness (Fig. 2.30a). As the moment is increased, the outer fibers in the bar

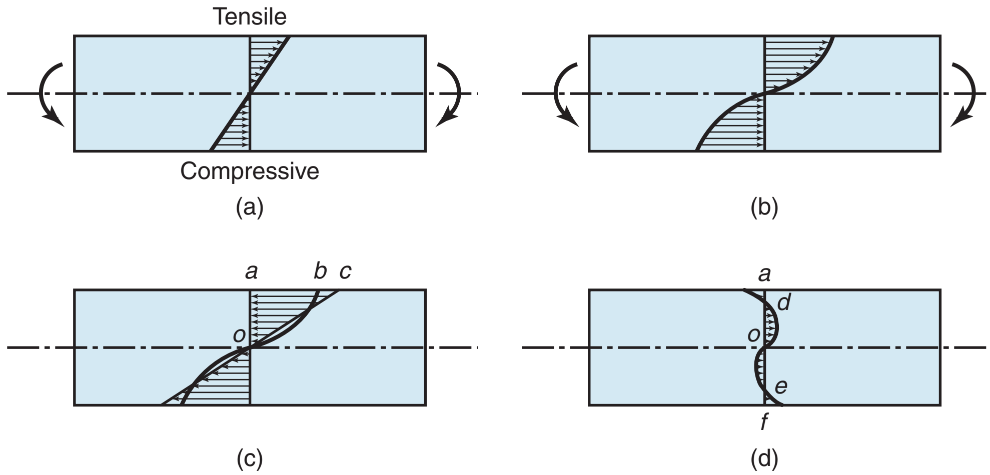


Figure 2.30: Residual stresses developed in bending a beam having a rectangular cross section. Note that the horizontal forces and moments caused by residual stresses in the beam must be balanced internally. Because of nonuniform deformation, especially during cold-metalworking operations, most parts develop residual stresses.

reach a stress level sufficiently high to cause yielding. For a typical strain-hardening material, the stress distribution shown in Fig. 2.30b is eventually reached, and the bar has now undergone permanent bending.

Consider the effects of removing the external bending moment on the bar. This operation is equivalent to applying an equal but opposite moment to the bar; thus, the moments of the areas oab and oac in Fig. 2.30c must be equal. Line oc, which represents the opposite bending moment, is linear, because all unloading and recovery is *elastic* (see Fig. 2.3). The difference between the two stress distributions gives the residual stress pattern within the bar, as is shown in Fig. 2.30d. Note the presence of compressive residual stresses in layers ad and oe, and tensile residual stresses in layers do and ef. Because there are now no external forces applied on the bar, the internal forces resulting from these residual stresses must be in static equilibrium. It should be noted that although this example involves residual stresses only in the longitudinal direction of the bar, in most cases residual stresses are three dimensional, and more difficult to analyze.

The removal of a layer of material from the surfaces of the bar, such as by machining or grinding, will disturb the equilibrium of the residual stresses shown in Fig. 2.30d. The bar will then acquire a new radius of curvature in order to balance the internal forces. Such disturbances of residual stresses cause **warping** of parts (Fig. 2.31). (Residual stresses may also be disturbed by *relaxation* of these stresses over a period of time; see below.)

Residual stresses also can be developed by *temperature gradients* within the part, such as occur during cooling of castings or a hot forgings. The local expansions and contractions caused by temperature gradients within the part will produce a nonuniform deformation, such as described above in the permanent bending of a beam.

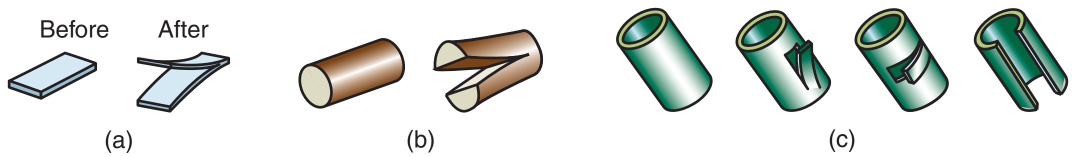


Figure 2.31: Distortion of parts with residual stresses after cutting or slitting: (a) flat sheet or plate; (b) solid round rod; (c) thin-walled tubing or pipe.

Tensile residual stresses on the surface of a part are generally undesirable, because they lower the fatigue life and fracture strength of the parts made. This is because a surface with tensile residual stresses cannot sustain as large of tensile stresses from external forces before fracture as if it were stress free. The reduction in strength is particularly characteristic of brittle or less ductile materials, in which fracture takes place with little or no plastic deformation preceding fracture. Tensile residual stresses can also lead, over a period of time, to *stress cracking* or *stress–corrosion cracking* of parts made (Section 2.10.2).

Compressive residual stresses on a surface, on the other hand, are generally desirable. In fact, in order to increase the fatigue life of components, compressive residual stresses can be purposefully imparted to surfaces by techniques such as *shot peening* or *surface rolling* (Section 34.2).

Reduction and Elimination of Residual Stresses. Residual stresses can be reduced or eliminated either by *stress-relief annealing* (Section 4.11) or by further *plastic deformation* of the part, such as by stretching. Given sufficient time, residual stresses may also diminish at room temperature by *relaxation*. The time required for relaxation can be greatly reduced by raising the workpiece temperature.

2.12 Work, Heat, and Temperature

Almost all the mechanical work in plastic deformation is converted into **heat**. However, the conversion is not complete, because a portion of the work is stored within the deformed material as elastic energy, known as **stored energy** (Section 1.7). This energy is generally 5–10% of the total energy input. In some metal alloys, however, it may be as high as 30%.

In a simple frictionless deformation process, and assuming that work is completely converted into heat, the theoretical (adiabatic) *temperature rise*, ΔT, in the workpiece is given by

$$\Delta T = \frac{u}{\rho c}, \tag{2.15}$$

where u is the **specific energy** (work of deformation per unit volume), ρ is the density, and c is the specific heat of the material. It can be noted that higher temperatures are associated with large areas under the stress–strain curve and with smaller values of specific heat. However, such physical properties as specific heat and thermal conductivity (Chapter 3) may also depend on temperature; thus, they must be taken into account in the calculations.

The temperature rise for a true strain of 1 (such as occurs in a 27 mm-high specimen when it is compressed down to 10 mm) can be calculated as: for aluminum aluminum, 75°C; copper, 140°C; low-carbon steel, 280°C; and titanium 570°C. In actual metalworking operations, however, heat is lost to the environment, to the tools and dies, and to the lubricants or coolants used, if any, in the process. If deformation is performed at high speed, the heat losses will be relatively low over that brief period. If, on the other hand, the process is carried out slowly, the actual temperature rise will be only a fraction of the calculated value.

Summary

- Manufacturing processes include shaping materials by plastic deformation; consequently, such mechanical properties as strength (yield strength, S_y, and ultimate tensile strength, S_{ut}); modulus of elasticity, E; ductility (total elongation and reduction of area); hardness; and the energy required for plastic deformation are important factors. These properties depend on the particular material and on its condition, temperature, deformation rate, surface condition, and the environment.

- The tensile test is the most commonly used to determine mechanical properties. From these tests, true stress–true strain curves are constructed to then determine the strength coefficient (K), the strain-hardening exponent (n), the strain-rate sensitivity exponent (m), and the toughness of materials.

- Compression tests are subject to inaccuracy due to the presence of friction and barreling of the specimen. Torsion tests typically are conducted on tubular specimens and subjected to twisting. Bending or flexure tests are commonly used for brittle materials to determine their modulus of rupture or the transverse rupture strength.
- Several hardness tests are available to determine the resistance of a material to permanent indentation or to scratching. Hardness is related to strength and wear resistance of a material, but hardness is not a fundamental property.
- Fatigue tests indicate the endurance limit or fatigue limit of materials, *i.e.*, the maximum stress to which a material can be subjected without fatigue failure, regardless of the number of cycles. Some materials have no endurance limit; instead, their allowable stress is reported with respect to the number of loading cycles.
- Creep is the permanent elongation of a component under a static load maintained for a period of time; failure is by rupture (necking and fracturing).
- Impact tests determine the energy required to completely break a specimen, called the impact toughness of the material. These tests are also useful for determining the transition temperature of materials.
- Failure and fracture constitute an important aspect of a material's behavior when subjected to deformation during manufacturing operations. Ductile fracture is characterized by plastic deformation preceding fracture; it requires a considerable amount of energy. Brittle fracture can be catastrophic, because it is not preceded by plastic deformation; however, it requires much less energy than does ductile fracture. Impurities, inclusions, and voids play a major role in the fracture of metals and alloys.
- Residual stresses are those that remain in a workpiece after it has been plastically deformed and then has all external forces removed. Surface tensile residual stresses are generally undesirable; they may be reduced or eliminated by stress-relief annealing, further plastic deformation, or by relaxation over a period of time.

Key Terms

Bauschinger effect
Blue brittleness
Brittle fracture
Buckling
Charpy test
Compression
Creep
Defects
Deformation rate
Disk test
Ductile fracture
Ductility
Durometer
Elongation
Engineering strain
Engineering stress
Fatigue
Fatigue failure
Flexural test
Fracture
Hardness
Impact loading
Inclusions
Izod test
Leeb test
Microhardness
Modulus of elasticity
Modulus of rigidity

Modulus of rupture
Poisson's ratio
Reduction of area
Residual stresses
Rupture
Shear
Shear modulus
Shore test
Strain aging
Strain-hardening exponent
Strain rate
Strain-rate sensitivity exponent
Strain softening
Strength coefficient
Stress–corrosion cracking
Stress relaxation
Superplasticity
Tension
Torsion test
Toughness
Transition temperature
True strain
True stress
Ultimate tensile strength
Yield stress

Bibliography

Ashby, M.F., **Materials Selection in Mechanical Design**, 4th ed., Pergamon, 2010.
ASM Handbook, Vol. 8: **Mechanical Testing and Evaluation**, ASM International, 2000.
Atlas of Stress–Strain Curves, 2nd ed., ASM International, 2002.
Beer, F.P., Johnston, E.R., DeWolf, J.T., and Mazurek, D., **Mechanics of Materials,** 7th ed., McGraw-Hill, 2014.
Courtney, T.H., **Mechanical Behavior of Materials**, 2nd ed., Waveland Press, 2005.
Dowling, N.E., **Mechanical Behavior of Materials: Engineering Methods for Deformation, Fracture, and Fatigue**, 4th ed., Prentice Hall, 2012.
Herrmann, K. (ed.), **Hardness, Testing: Principles and Applications,** ASM International, 2011.
Hosford, W.F., **Mechanical Behavior of Materials**, 2nd ed., Cambridge, 2009.
Wulpi, D.J., **Understanding How Components Fail**, 2nd ed., ASM International, 1999.

Review Questions

2.1. What is tension? What is shear?

2.2. Distinguish between engineering stress and true stress.

2.3. In a stress–strain curve, what is the proportional limit? Is it different than the yield point?

2.4. Describe the events that occur when a specimen undergoes a tension test. Sketch a plausible stress–strain curve, and identify all significant regions and points between them. Assume that loading continues up to fracture.

2.5. What is ductility, and how is it measured?

2.6. In the equation $\sigma = K\epsilon^n$, which represents the stress–strain curve for a material, and what is the significance of the exponent n?

2.7. What is strain-rate sensitivity, and how is it measured?

2.8. What test can measure the properties of a material undergoing shear strain?

2.9. What kind of loading is applied by scissors while cutting paper?

2.10. What testing procedures can be used to measure the properties of brittle materials, such as ceramics and carbides?

2.11. Describe the differences between brittle and ductile fracture.

2.12. What is hardness? Is it different from hardness number? Explain.

2.13. Describe the features of a Rockwell hardness test.

2.14. What is a Leeb test? How is it different from a Rockwell A test?

2.15. Differentiate between stress relaxation and creep.

2.16. Describe the difference between elastic and plastic behavior.

2.17. Explain what uniform elongation means in tension testing.

2.18. Describe the difference between deformation rate and strain rate. What unit does each one have?

2.19. Describe the difficulties involved in conducting a compression test.

2.20. What is Hooke's law, Young's modulus, and Poisson's ratio?

2.21. Describe the difference between transgranular and intergranular fracture.

2.22. What is the reason that yield strength is generally defined as a 0.2% offset strength?

2.23. Why does the fatigue strength of a specimen or part depend on its surface finish?

2.24. Explain how you would determine whether or not a material has an endurance limit.

2.25. If striations are observed under microscopic examination of a fracture surface, what do they suggest regarding the mode of fracture?

2.26. What is an Izod test? Why are Izod tests useful?

2.27. Why does temperature increase during deformation?

2.28. What is a residual stress? How can residual stresses be removed?

Qualitative Problems

2.29. On the same scale for stress, the tensile true stress–true strain curve is higher than the engineering stress–engineering strain curve. Explain whether this condition also holds for a compression test.

2.30. Explain why it is difficult to break a sheet of paper in tension, but easy to cut it with scissors.

2.31. What are the similarities and differences between deformation and strain?

2.32. Can a material have a negative Poisson's ratio? Give a rationale for your answer.

2.33. Referring to Table 2.2, explain why there can be so much variation in the strength and elongation in a class of metal alloys.

2.34. Referring to Table 2.2, explain why the stiffness of diamond has so much variation.

2.35. It has been stated that the higher the value of m, the more diffuse the neck is, and likewise, the lower the value of m, the more localized the neck is. Explain the reason for this behavior.

2.36. Explain why materials with high m values, such as hot glass and taffy, when stretched slowly, undergo large elongations before failure. Consider events taking place in the necked region of the specimen.

2.37. Explain if it is possible for stress–strain curves in tension tests to reach 0% elongation as the gage length is increased further.

2.38. With a simple sketch, explain whether it is necessary to use the offset method to determine the yield stress, S_y, of a material that has been highly cold worked.

2.39. Explain why the difference between engineering strain and true strain becomes larger as strain increases. Does this difference occur for both tensile and compressive strains? Explain.

2.40. Consider an elastomer, such as a rubber band. This material can undergo a large elastic deformation before failure, but after fracture it recovers completely to its original shape. Is this material brittle or ductile? Explain.

2.41. If a material (such as aluminum) does not have an endurance limit, how then would you estimate its fatigue life?

2.42. What role, if any, does friction play in a hardness test? Explain.

2.43. Which hardness tests and scales would you use for very thin strips of metal, such as aluminum foil? Explain.

2.44. Consider the circumstance where a Vickers hardness test is conducted on a material. Sketch the resulting indentation shape if there is a residual stress on the surface.

2.45. Which of the two tests, tension or compression, would require a higher capacity of testing machine, and why?

2.46. In a Brinell hardness test, the resulting impression is found to be an ellipse. Give possible explanations for this result.

2.47. List and explain briefly the conditions that induce brittle fracture in an otherwise ductile metal.

2.48. List the factors that you would consider in selecting a hardness test. Explain why.

2.49. List two situations where a material's toughness is important from a design standpoint.

2.50. On the basis of Fig. 2.5, if a metal tension-test specimen is pulled and broken rapidly, where would the temperature be highest, and why?

2.51. Comment on the temperature distribution if the specimen in Question 2.50 is pulled very slowly.

2.52. Comment on your observations regarding the contents of Table 2.2.

2.53. Is the disk test applicable to a ductile material? Why or why not?

2.54. Refer to Table 2.4, and note the true strain encountered by a material in different manufacturing processes. Explain why some typical strains are large and others are small.

2.55. Refer to Table 2.4, and sketch the original and deformed shape of a 25 mm specimen subjected to the largest typical strain for each process. What are your observations regarding strains that can be achieved?

2.56. If a tension test on carbon steel is conducted at room temperature, and then with a bath of boiling water, would you expect the strength to be different? Explain.

2.57. What hardness test is suitable for determining the hardness of a thin ceramic coating on a piece of metal?

2.58. Wire rope consists of many wires that bend and unbend as the rope is run over a sheave. A wire-rope failure is investigated, and it is found that some of the wires, when examined under a scanning electron microscope, display cup-and-cone failure surfaces, while others display transgranular fracture surfaces. Explain these observations.

2.59. A statistical sampling of Rockwell C hardness tests are conducted on a material, and it is determined that the material is defective because of insufficient hardness. The supplier claims that the tests are flawed because the diamond-cone indenter was probably dull. Is this a valid claim? Explain.

2.60. In a Brinell hardness test, the resulting impression is found to be elliptical. Give possible explanations for this result.

2.61. In the machining of an extruded aluminum block to produce a smart phone case, it is seen that there is significant warpage after machining. Explain why. What would you do to reduce this warpage?

2.62. Some coatings are extremely thin—some as thin as a few nanometers. Explain why even the Knoop test is not able to obtain reliable results for such coatings. Recent investigations have attempted to use highly polished diamonds (with a tip radius around 5 nm) to indent such coatings in atomic force microscopes. What concerns would you have regarding the appropriateness of the results?

2.63. Select an appropriate hardness test for each of the following materials, and justify your answer:

1. Cubic boron nitride
2. Lead
3. Cold-drawn 0.5%C steel
4. Diamond
5. Caramel candy
6. Granite.

2.64. Referring to Fig. 2.13, the material for testers is either steel, tungsten carbide, or diamond. Why isn't diamond used for all of the tests?

Quantitative Problems

2.65. A paper clip is made of wire 1 mm in diameter. If the original material from which the wire is made is a rod 50 mm in diameter, calculate the longitudinal engineering and true strains that the wire has undergone during processing.

2.66. A 150-mm-long strip of metal is stretched in two steps, first to 250 mm and then to 500 mm. Show that the total true strain is the sum of the true strains in each step; in other words, the true strains are additive. Show that, in the case of engineering strains, the strains cannot be added to obtain the total strain.

2.67. Identify the two materials in Fig. 2.5 that have the lowest and the highest uniform elongations. Calculate these quantities as percentages of the original gage lengths.

2.68. Plot the ultimate strength versus stiffness for the materials listed in Table 2.2, and prepare a three-dimensional plot for these materials where the third axis is their maximum elongation in 50 mm.

2.69. If you remove the layer of material ad from the part shown in Fig. 2.30d—for instance, by machining or grinding—which way will the specimen curve? (*Hint:* Assume that the part shown in sketch d in the figure is composed of four horizontal springs held at the ends. Thus, from the top down, you have compression, tension, compression, and tension springs.)

2.70. Prove that the true strain at necking equals the strain hardening exponent.

2.71. Percent elongation is always defined in terms of the original gage length, such as 50 mm. Explain how percent elongation would vary as the gage length of the tensile-test specimen increases. (*Hint:* Recall that necking is a *local* phenomenon.)

2.72. Make a sketch showing the nature and distribution of residual stresses in Fig. 2.31a and b, prior to the materials being cut. (*Hint:* Assume that the split parts are free from any stresses; then force these parts back to the shape they originally had.)

2.73. You are given the K and n values of two different metals. Is this information sufficient to determine which metal is tougher? If not, what additional information do you need?

2.74. A cable is made of two strands of different materials, A and B, and cross sections as follows:

For material A, $K = 500$ MPa, $n = 0.6$, $A_o = 0.00060\ \text{m}^2$.
For material B, $K = 300$ MPa, $n = 0.6$, $A_o = 0.00030\ \text{m}^2$.

Calculate the maximum tensile force that this cable can withstand prior to necking.

2.75. On the basis of the information given in Fig. 2.5, calculate the ultimate tensile strength (engineering) of 304 stainless steel.

2.76. In a disk test performed on a specimen 30 mm in diameter and 8 mm thick, the specimen fractures at a stress of 180 MPa. What was the load on at fracture?

2.77. A piece of steel has a hardness of 275 HB. Calculate its tensile strength, in MPa.

2.78. A metal has the following properties: $S_{\text{ut}} = 500$ MPa and $n = 0.25$. Calculate its strength coefficient, K.

2.79. Using only Fig. 2.5, calculate the maximum load in tension testing of an annealed copper specimen with an original diameter of 10 mm.

2.80. Estimate the modulus of resilience for a highly cold worked piece of steel having a hardness of 300 HB; for a piece of highly cold worked copper with a hardness of 100 HRB.

2.81. A metal has a strength coefficient $K = 600$ MPa and $n = 0.25$. Assuming that a tensile-test specimen made from this metal begins to neck at a true strain of 0.25, show that the ultimate tensile strength is 362 MPa.

2.82. Plot the true stress–true strain curves for the materials listed in Table 2.3.

2.83. The design specification for a metal requires a minimum hardness of 80 HRA. If a Rockwell test is performed and the depth of penetration is 80 μm, is the material acceptable?

2.84. Calculate the major and minor pyramid angles for a Knoop indenter, and compare your results with those obtained from Vickers and Rockwell A indenters.

2.85. If a material has a target hardness of 300 HB, what is the expected indentation diameter?

2.86. A Rockwell A test was conducted on a material and a penetration depth of 0.15 mm was recorded. What is the hardness of the material? What material would typically have such a hardness value? If a Brinell hardness test were to be conducted on this material, give an estimate of the indentation diameter if the load used was 1500 kg.

2.87. For a cold-drawn 0.5% carbon steel, will a Rockwell C test or a Brinell test at 500 kg result in a deeper penetration?

2.88. A material is tested in tension. Over a 25-mm gage length, the engineering strain measurements are 0.01, 0.02, 0.03, 0.04, 0.05, 0.1, 0.15, 0.2, 0.5, and 1.0. Plot the true strain versus engineering strain for these readings.

2.89. Calculate the work done in frictionless compression of a solid cylinder 40 mm high and 15 mm in diameter to a reduction in height of 50% for the following materials: (a) 1100-O aluminum; (b) annealed copper; (c) annealed 304 stainless steel; and (d) annealed 70-30 brass.

2.90. A bar 2 m long is bent and then stress relieved. The radius of curvature to the neutral axis is 1 m. The bar is 30 mm thick and is made of an elastic, perfectly plastic material with $S_y = 500$ MPa and $E = 207$ GPa. Calculate the length to which this bar should be stretched so that, after unloading, it will become and remain straight.

2.91. Take a cubic piece of metal with a side length l_o and deform it plastically to the shape of a rectangular parallelepiped of dimensions l_1, l_2, and l_3. Assuming that the material is rigid and perfectly plastic, show that volume constancy requires that the following expression be satisfied: $\epsilon_1 + \epsilon_2 + \epsilon_3 = 0$.

Synthesis, Design, and Projects

2.92. List and explain the desirable mechanical properties of (a) an automobile body panel, (b) a paper clip, (c) a leaf spring for a truck, (d) a bracket for a bookshelf, (e) a backpack shoulder strap, (f) a wire coat hanger, (g) the clip for a pen, and (h) a lens for an optical microscope.

2.93. When making a hamburger, you may have observed the type of cracks shown in Fig. 2.20d. What would you do to avoid such cracks? [*Note:* Test hamburger patties by compressing them at different temperatures, and observe the crack path (*i.e.*, the path through the fat particles, the meat particles, or their interface).]

2.94. An inexpensive claylike material called Silly Putty® is generally available in stores that sell toys and games. Obtain a sample and perform the following experiments: (a) Shape it into a ball, and drop it onto a flat surface. (b) Reround the ball and place a heavy book on it for one minute. (c) Shape the putty into a long rod, and pull on it—first slowly, then very quickly. Describe your observations, referring to the specific sections in this chapter where each particular observation is relevant.

2.95. Make individual sketches of the mechanisms of testing machines that, in your opinion, would be appropriate for tension, for torsion, and for compression testing of specimens at different rates of deformation. What modifications would you make on these machines to include the effects of temperature on material properties?

2.96. In tension testing of specimens, mechanical and electronic instruments are typically used to measure elongation. Make sketches of instruments that would be suitable for this purpose, commenting on their accuracy. What modifications would you make to these instruments to include the use of specimens at elevated temperatures?

2.97. Obtain small pieces of different metallic and nonmetallic materials, including stones. Rub them against each other, observe the scratches made, and order them in a manner similar to the Mohs hardness numbering system.

2.98. Demonstrate the stress-relaxation phenomenon by tightly stretching thin plastic strings between two nails placed at the ends of a long piece of wood. Pluck the strings frequently, to test the tension as a function of time. Repeat the test at a higher temperature by placing the fixture in an oven set on low.

2.99. Demonstrate the impact toughness of a piece of round chalk by first using a triangular file to produce a V-notch on the cylindrical surface (as shown in Fig. 2.19a) and then bending the chalk to break it.

2.100. Using a large rubber band and a set of weights, obtain the force–displacement curve for the rubber band. Is the result different from the stress–strain curves shown in Fig. 2.4? Explain.

2.101. Design a test protocol to obtain the work of plastic deformation by measuring the temperature rise in a workpiece, assuming that there is no heat loss and that the temperature distribution is uniform throughout. If the specific heat of the material decreases with increasing temperature, will the work of deformation calculated using the specific heat at room temperature be higher or lower than the actual work done? Explain.

2.102. Find or prepare some solid circular pieces of brittle materials, such as chalk, ceramics, etc. and subject them to the type of test shown in Fig. 2.9 by using the jaws of a simple vise. Describe your observations as to how the materials fracture. Repeat the tests, using ductile materials, such as clay, soft metals, and describe your observations.

2.103. Take several rubber bands and pull them at different temperatures, including from a frozen state. Comment on their behavior such as ductile or brittle.

2.104. Devise a simple fixture for conducting the bend tests shown in Fig. 2.11. Test sticks of various brittle materials by loading them with dead weights until they break. Verify the statement in the text that the specimens on the right in the figure will fracture sooner than the ones on the left.

2.105. By pushing a small ball bearing against the top surfaces of various materials, such as clay and dough, observe the shape of the indentation with a magnifier, referring to those shapes shown in Fig. 2.14a and b.

2.106. Describe your observations regarding Fig. 2.14c.

2.107. Embed a small steel ball in a soft block of material such as clay, and compress the clay as shown in Fig. 2.24a. Then cut the clay carefully along the center plane and observe the deformation of the material. Repeat the experiment by embedding a small round jelly bean in the clay and deforming the material. Comment on your observations.

2.108. A penny-shaped piece of soft metal is brazed to the ends of two flat, round steel rods of the same diameter as the piece. The assembly is then subjected to uniaxial tension. What is the state of stress to which the soft metal is subjected? Explain.

2.109. Devise a simple experiment, and perform tests on materials commonly found around the house by bending them at different temperatures for a qualitative assessment of their transition temperature, as shown in Fig. 2.25.

2.110. Obtain some solid and some tubular metal pieces, and slit them as shown in Fig. 2.31. Comment on whether there are any residual stresses in the parts prior to slitting them.

2.111. Explain how you would obtain an estimate of the hardness for a carbon nanotube (see Section 8.6.2).

2.112. Without using the words “stress” or “strain,” define elastic modulus.

2.113. We know that it is relatively easy to subject a specimen to hydrostatic compression, such as by using a chamber filled with a liquid. Devise a means whereby the specimen (say, in the shape of a cube or a round disk) can be subjected to hydrostatic tension, or one approaching this state of stress. (Note that a thin-walled, internally pressurized spherical shell is not a correct answer, because it is subjected only to a state of plane stress.)

2.114. Assume that you are running four-point bending tests on a number of identical specimens of the same length and cross section, but with increasing distance between the upper points of loading. (See Fig. 2.19b.) What changes, if any, would you expect in the test results? Explain.

2.115. Describe a test protocol, complete with forces and geometry, that you execute to determine the quality of food. Consider the crispness of an apple or snack chip, and then consider the softness of cake.

Chapter 3

Physical Properties of Materials

- Physical properties have several significant roles in the selection, processing, and use of materials. They can also be key factors in determining a material's suitability for specific applications, especially when considered simultaneously with mechanical properties.
- Strength-to-weight and stiffness-to-weight ratios, as examples, are described in the context of lightweight designs, an important consideration especially in aerospace and automotive industries.
- Thermal, electrical, magnetic, and optical properties are presented.
- The importance of corrosion and corrosion-resistant materials are then described.
- Design and manufacturing implications of each physical property is considered, with specific examples given.

3.1 Introduction

Why is electrical wiring generally made of copper? Why are stainless steel, aluminum, and copper commonly used in cookware? Why are the handles of cookware usually made of plastic, while other types of handles are made of metal? What kind of material should be chosen for the heating elements in toasters and toaster ovens? Why does metal feel colder to the touch than plastic even though both are at room temperature? Why are the metallic components in some machines being replaced with ceramics? Why are commercial airplane bodies generally made of aluminum while others are made of reinforced plastics?

From these questions it is apparent that a major criterion in material selection requires consideration of physical properties: density, melting point, specific heat, thermal conductivity, thermal expansion, electrical and magnetic properties, and resistance to oxidation and corrosion. Combinations of mechanical and physical properties, such as the strength-to-weight and stiffness-to-weight ratios of materials, are equally important, particularly for aircraft and aerospace structures. Also, high-speed equipment, such as textile and printing machinery, and forming and cutting machines for high-speed operations, require lightweight components to reduce inertial forces and prevent the machines from being subjected to excessive vibration. Several other examples of the importance of physical properties are described in this chapter.

3.2 Density

The **density** of a material is its mass per unit volume; another term is **specific gravity**, which expresses a material's density relative to that of water; specific gravity has therefore no units. The range of densities for a variety of materials at room temperature, along with other properties, is given in Tables 3.1 and 3.2.

Weight saving is particularly important for aircraft, aerospace, automotive structures, sports equipment, and for various other products where energy consumption and power limitations are significant concerns. Substitution of materials for weight savings and fuel economy is a major factor in the design both of advanced equipment and machinery and of consumer products, such as sporting goods, portable computers, and bicycles.

A significant role that density plays is in the **strength-to-weight ratio** (**specific strength**) and **stiffness-to-weight ratio** (**specific stiffness**) of materials. Figure 3.1 shows the ratio of maximum yield stress to density for a variety of metal alloys. Note that titanium and aluminum are at the top of the list; consequently, they are among the most commonly used metals for aircraft and aerospace applications.

The ranges for specific tensile strength and specific stiffness at room temperature for a variety of metallic and nonmetallic materials are given in Fig. 3.2. Note the positions of composite materials, as compared to those of metals, with respect to these properties. These advantages have led composites to become among the most important materials, as described in Chapter 9. At elevated temperatures, specific strength and specific stiffness are likewise important considerations, especially for components operating at these temperatures, such as automotive and jet engines, gas turbines, and furnaces. Typical ranges for a variety of materials are given in Fig. 3.3.

Density is an important factor in the selection of materials for high-speed equipment, such as magnesium in printing and textile machinery, many components of which typically operate at very high speeds. Aluminum is used with some digital cameras for better performance in cold weather. Because of their low density, ceramics (Chapter 8) are being used for components in high-speed automated machinery and machine tools.

On the other hand, there are applications where weight is desirable. Examples are counterweights for various mechanisms (using lead or steel), flywheels, ballasts on ships and aircraft, and weights on golf clubs (using high-density materials such as tungsten).

3.3 Melting Point

The temperature range within which a component or structure is designed to function is an important consideration in selection of materials. Plastics, for example, have the lowest useful temperature range (Table 7.2), while ceramics, graphite, and refractory-metal alloys have the highest useful range. Pure metals have a definite **melting point**, whereas the melting temperature of a metal alloy can have a wide range (Table 3.1), depending on its composition.

The melting point has several indirect effects on manufacturing operations. Because the recrystallization temperature of a metal is related to its melting point (Section 1.7), operations such as annealing and

Table 3.1: Physical Properties of Selected Materials at Room Temperature

Material	Density (kg/m^3)	Melting point (°C)	Specific heat (J/kg K)	Thermal conductivity (W/m-K)	Coefficient of thermal expansion (μm/m-°C)	Electrical resistivity (Ω-m)
Metallic						
Aluminum	2700	660	900	222	23.6	$2.8\ 10^{-8}$
Aluminum alloys	2630–2820	476–654	880–920	121–239	23.0–23.6	$2.8–4.0 \times 10^{-8}$
Beryllium	1854	1278	1884	146	8.5	4.0×10^{-8}
Niobium(columbium)	8580	2468	272	52	7.1	15×10^{-8}
Copper	8970	1082	385	393	16.5	1.7×10^{-8}
Copper alloys	7470–8940	885–1260	377–435	219–234	16.5–20	$1.7–5.9 \times 10^{-8}$
Gold	19,300	1063	129	317	19.3	2.4×10^{-8}
Iron	7860	1537	460	74	11.5	9.5×10^{-8}
Steels	6920–9130	1371–1532	448–502	15–52	11.7–17.3	17.0×10^{-8}
Lead	11,350	327	130	35	29.4	20.6×10^{-8}
Lead alloys	8850–11,350	182–326	126–188	24–46	27.1–31.1	$20.6–24 \times 10^{-8}$
Magnesium	1745	650	1025	154	26.0	4.5×10^{-8}
Magnesium alloys	1770	610–621	1046	75–138	26.0	$4.5–15.9 \times 10^{-8}$
Molybdenum alloys	10,210	2610	276	142	5.1	5.3×10^{-8}
Nickel	8910	1453	440	92	13.3	6.2×10^{-8}
Nickel alloys	7750–8850	1110–1454	381–544	12–63	12.7–18.4	$6.2–110 \times 10^{-8}$
Platinum	21,450	1768	133	71.6	8.8	10.5×10^{-8}
Silicon	2330	1423	712	148	7.63	1.0×10^{-3}
Silver	10,500	961	235	429	19.3	1.6×10^{-8}
Tantalum alloys	16,600	2996	142	54	6.5	13.5×10^{-8}
Tin	7310	232	217	67	22	11.5×10^{-8}
Titanium	4510	1668	519	17	8.35	42×10^{-8}
Titanium alloys	4430–4700	1549–1649	502–544	8–12	8.1–9.5	$40–171 \times 10^{-8}$
Tungsten	19,290	3410	138	166	4.5	5.0×10^{-8}
Zinc	7140	419	385	113	32.5	5.45×10^{-8}
Zinc alloys	6640–7200	386–525	402	105–113	32.5–35	$6.06–6.89 \times 10^{-8}$
Nonmetallic						
Ceramics	2300–5500	—	750–950	10–17	5.5–13.5	—
Glasses	2400–2700	580–1540	500–850	0.6–1.7	4.6–70	—
Graphite	1900–2200	—	840	5–10	7.86	—
Plastics	900–2000	110–330	1000–2000	0.1–0.4	72–200	—
Wood	400–700	—	2400–2800	0.1–0.4	2–60	—

heat treating (Chapter 4) and hot working (Part III) require knowledge of the melting points of the materials involved. These considerations are also important in the selection of tool and die materials. In casting operations (Part II), melting point plays a major role in the selection of the equipment and the melting practice employed. In the electrical-discharge machining process (Section 27.5), the melting points of metals are related to the rate of material removal and of electrode wear.

3.4 Specific Heat

A material's **specific heat** is the energy required to raise the temperature of a unit mass by one degree. Alloying elements have a relatively minor effect on the specific heat of metals. The temperature rise in a workpiece, such as those in forming or machining operations (Parts III and IV, respectively), is a function of

Table 3.2: Physical Properties of Materials, in Descending Order

Density	Melting point	Specific heat	Thermal conductivity	Thermal expansion	Electrical conductivity
Platinum	Tungsten	Wood	Silver	Plastics	Silver
Gold	Tantalum	Beryllium	Copper	Lead	Copper
Tungsten	Molybdenum	Porcelain	Gold	Tin	Gold
Tantalum	Niobium	Aluminum	Aluminum	Magnesium	Aluminum
Lead	Titanium	Graphite	Magnesium	Aluminum	Magnesium
Silver	Iron	Glass	Graphite	Copper	Tungsten
Molybdenum	Beryllium	Titanium	Tungsten	Steel	Beryllium
Copper	Copper	Iron	Beryllium	Gold	Steel
Steel	Gold	Copper	Zinc	Ceramics	Tin
Titanium	Silver	Molybdenum	Steel	Glass	Graphite
Aluminum	Aluminum	Tungsten	Tantalum	Tungsten	Ceramics
Beryllium	Magnesium	Lead	Ceramics		Glass
Glass	Lead		Titanium		Plastics
Magnesium	Tin		Glass		Quartz
Plastics	Plastics		Plastics		

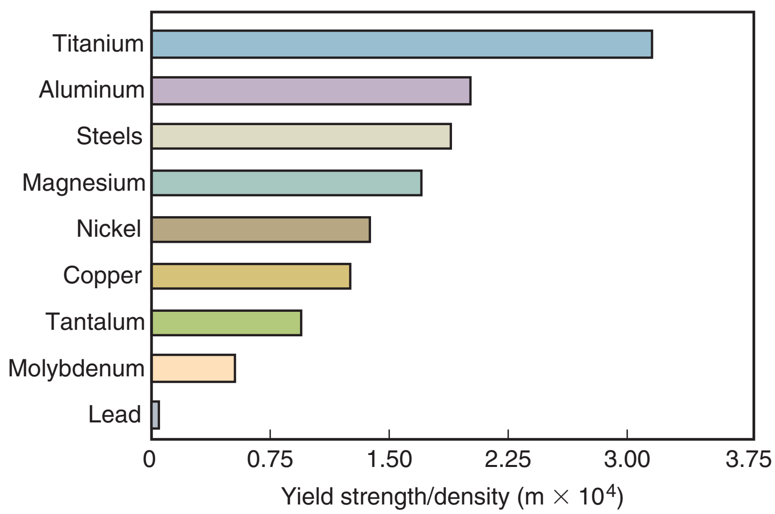


Figure 3.1: Ratio of maximum yield stress to density for selected metals.

the work done and of the specific heat of the workpiece material (Section 2.12). An excessive temperature rise in a workpiece can

a) decrease product quality by adversely affecting its surface finish and dimensional accuracy,

b) cause excessive tool and die wear, and

c) result in undesirable metallurgical changes in the material.

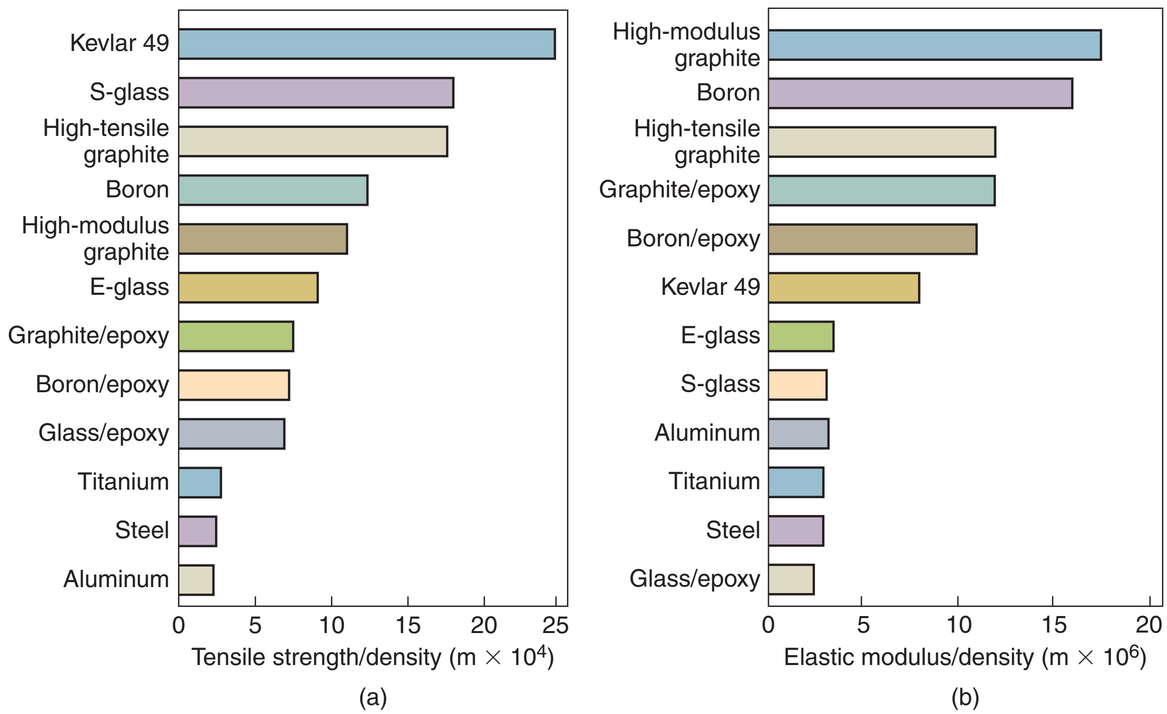


Figure 3.2: Specific strength (tensile strength/density) and specific stiffness (elastic modulus/density) for various materials at room temperature (see also Chapter 9).

3.5 Thermal Conductivity

Thermal conductivity indicates the rate at which heat flows within and through a material. Metallically bonded materials (metals) generally have high thermal conductivity, while ionically or covalently bonded materials (ceramics and plastics) have poor conductivity (Table 3.2). Alloying elements can have a significant effect on the thermal conductivity of alloys, as can be seen by comparing pure metals with their alloys in Table 3.1. In general, materials with high electrical conductivity also have high thermal conductivity.

Thermal conductivity is an important consideration in numerous applications. For example, high thermal conductivity is desirable in cooling fins, cutting tools, and die-casting molds to extract heat quickly. In contrast, materials with low thermal conductivity are used, for instance, in furnace linings, insulation, coffee cups, and handles for pots and pans. One function of a lubricant (Section 33.7) in hot metalworking is to act as an insulator to keep workpieces hot and formable.

3.6 Thermal Expansion

The **thermal expansion** of materials can have several significant effects, particularly the relative expansion or contraction of different materials in assemblies, such as electronic and computer components, glass-to-metal seals, struts on jet engines, coatings on cutting tools (Section 22.5), and moving parts in machinery that require certain clearances for proper functioning. The use of ceramic components in cast-iron engines, for example, also requires consideration of their relative expansion and contraction during their operation.

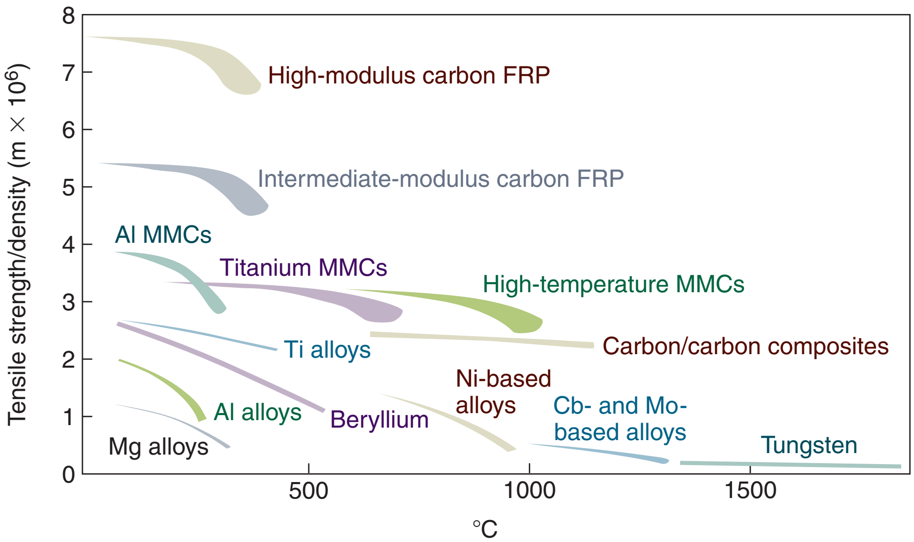


Figure 3.3: Specific strength (tensile strength/density) for a variety of materials as a function of temperature. Note the useful temperature range for these materials and the high values for composite materials. MMC = metal-matrix composite; FRP = fiber-reinforced plastic.

Shrink fits utilize thermal expansion and contraction. A shrink fit is a part, often a sleeve or a hub, that is to be installed over a shaft. The part is first heated and then slipped over the shaft or spindle; when allowed to cool, the hub shrinks and the assembly becomes an integral component.

Typical coefficients of thermal expansion are given in Table 3.1 (see also *Invar* below). Generally, the coefficient of thermal expansion is inversely proportional to the melting point of the material. Alloying elements have a relatively minor effect on the thermal expansion of metals.

Thermal expansion in conjunction with thermal conductivity plays the most significant role in the development of **thermal stresses** (due to *temperature gradients*), both in manufactured components and in tools and dies, and molds for casting operations. This consideration is particularly important in, for example, a forging operation during which hot workpieces are repeatedly placed over a relatively cool die, thus subjecting the die surfaces to thermal cycling. To reduce thermal stresses, a combination of high thermal conductivity and low thermal expansion is desirable. Thermal stresses also can be caused by **anisotropy of thermal expansion**; that is, the material expands differently in different directions, a property generally observed in hexagonal close-packed metals, ceramics, and composite materials.

Thermal expansion and contraction can lead to cracking, warping, or loosening of assembled components during their service life, as well as cracking of ceramic parts and in tools and dies made of relatively brittle materials. **Thermal fatigue** results from thermal cycling and causes a number of surface cracks, especially in tools and dies for casting and in metalworking operations (**heat checking**). **Thermal shock** is the term generally used to describe the development of a crack or cracks after being subjected to a single thermal cycle.

To alleviate some of the problems caused by thermal expansion a family of iron-nickel alloys with very low thermal-expansion coefficients are available, called **low-expansion alloys**. The low thermal expansion characteristic of these alloys is often referred to as the **Invar effect**, after the metal *Invar*. Their thermal coefficient of expansion is typically in the range of from 2×10^{-6} to 9×10^{-6} per °C (compare with those given in Table 3.1). Typical compositions are 64% Fe–36% Ni for Invar and 54% Fe–28% Ni–18% Co for Kovar.

Low-expansion alloys also have good thermal-fatigue resistance, and because of their good ductility they can easily be formed into various shapes. Applications include (a) bimetallic strips, which consist of a low-expansion alloy bonded to a high-expansion alloy (thus, the strip develops a curvature when subjected to temperature variations), and (b) glass-to-metal seals, in which the thermal expansions of the two materials are matched.

3.7 Electrical, Magnetic, and Optical Properties

Electrical conductivity and the *dielectric* properties of materials are important not only in various electrical equipment and machinery, but also in such manufacturing processes as magnetic-pulse forming (Section 16.12), resistance welding (Section 31.5), and electrical-discharge machining and electrochemical grinding of hard and brittle materials (Chapter 27). The units of electrical conductivity are mho/m, where mho is the reciprocal of ohm, the unit of electrical resistance. Alloying elements have a major effect on the electrical conductivity of metals: The higher the conductivity of the alloying element, the higher is the electrical conductivity of the alloy.

Dielectric Strength. An electrically insulating material's *dielectric strength* is the largest electric field to which it can be subjected without degrading or losing its insulating properties. This property is defined as the voltage required per unit distance for electrical breakdown, and has the units of V/m.

Conductors. Materials with high electrical conductivity, such as metals, are generally referred to as *conductors*. **Electrical resistivity** is the inverse of electrical conductivity. Materials with high electrical resistivity are referred to as **dielectrics** or **insulators**.

Superconductors. *Superconductivity* is the phenomenon of near-zero electrical resistivity that occurs in some metals and alloys below a critical temperature. The temperatures involved often are near absolute zero (0 K, –273°C). The highest temperature at which superconductivity has to date been exhibited, at about –123°C, is with an alloy of thallium, barium, calcium, copper, and oxygen; other material compositions are continuously being investigated.

The main application of superconductors is largely for high-power magnets. Superconductors are the enabling technology for magnetic resonance imaging (MRI), used for medical imaging. Other applications envisioned for superconductors include magnetic levitation (maglev) trains, efficient power transmission lines, and extremely fast computer components.

Semiconductors. The electrical properties of semiconductors, such as single-crystal silicon, germanium, and gallium arsenide, are extremely sensitive to temperature and to the presence and type of minute impurities. Thus, by controlling the concentration and type of impurities (called **dopants**), such as phosphorus and boron in silicon, electrical conductivity can be controlled. This property is utilized in semiconductor (solid-state) devices, used extensively in miniaturized electronic circuitry (Chapter 28).

Ferromagnetism and Ferrimagnetism. *Ferromagnetism* is a phenomenon characterized by high permeability and permanent magnetization that are due to the alignment of iron, nickel, and cobalt atoms into domains. It is important in such applications as electric motors, electric generators, electric transformers, and microwave devices. *Ferrimagnetism* is a permanent and large magnetization exhibited by some ceramic materials such as cubic ferrites.

Piezoelectric Effect. The *piezoelectric effect* (*piezo* from Greek, meaning to press) is exhibited by **smart materials**. Two basic behaviors are involved: (a) When subjected to an electric current, they undergo a reversible change in shape, by as much as 4% and (b) when deformed by an external force, they emit a small electric current.

Piezoelectric materials include quartz crystals and some ceramics and polymers. The piezoelectric effect is utilized in making transducers, which are devices that convert the strain from an external force into electrical energy. Typical applications are sensors, force or pressure transducers, inkjet printers, strain gages, sonar detectors, and microphones. As an example, an air bag in an automobile has a sensor that, when subjected to an impact force, sends an electric charge that then deploys the bag.

Magnetostriction. The phenomenon of expansion or contraction of a material when subjected to a magnetic field is called *magnetostriction*. Pure nickel and some iron–nickel alloys exhibit this behavior. Magnetostriction is the principle behind ultrasonic machining equipment (Section 26.6).

Magnetorheostatic and Electrorheostatic Effects. When subjected to magnetic or electric fields, some fluids undergo a major and reversible change in their viscosity within a fraction of a second, turning from a liquid to an almost solid state. For example, magnetorheostatic behavior is attained by mixing very fine iron filings with oil. Called **smart fluids**, they are being developed for such applications as vibration dampeners, engine mounts, prosthetic devices, clutches, and valves.

Optical Properties. Among various other properties, color and opacity are particularly relevant to polymers and glasses (see Sections 7.2.2 and 8.4.3, respectively).

3.8 Corrosion Resistance

Corrosion not only leads to surface deterioration of components and structures, such as bridges and ships, but also reduces their strength and structural integrity. The direct cost of corrosion to the U.S. economy alone has been estimated to be over $400 billion per year, approximately 2% of the gross domestic product; indirect costs of corrosion are estimated at twice this amount. Metals, ceramics, and plastics are all subject to forms of **corrosion**. The word *corrosion* itself usually refers to the deterioration of metals and ceramics, while similar phenomena in plastics (Chapter 7) are generally called **degradation**.

Corrosion resistance is an important aspect of material selection for applications in the chemical, petroleum, and food industries, as well as in manufacturing operations. In addition to various possible chemical reactions from the elements and compounds present, environmental oxidation and corrosion of a wide range of components and structures is a major concern, particularly at elevated temperatures.

Resistance to corrosion depends on the composition of the material and on its particular environment. Corrosive media may consist of various chemicals (acids, alkalis, and salts) and the environment (oxygen, moisture, pollution, and acid rain), including water (fresh or salt water). Nonferrous metals, stainless steels, and nonmetallic materials generally have high corrosion resistance. Steels and some cast irons generally have poor resistance and must be protected by a variety of coatings and surface treatments (Chapter 34).

Corrosion can occur over an entire surface or it can be *localized*, called **pitting**. Pitting is a term that is also used for fatigue wear or failure of gears and in forging (see Section 33.5). Corrosion can also occur along grain boundaries of metals as **intergranular corrosion**, and at the interface of bolted or riveted joints as **crevice corrosion**.

Two dissimilar metals may form a **galvanic cell** (after L. Galvani, 1737–1798); that is, two electrodes in an electrolyte in a corrosive environment that includes moisture and cause **galvanic corrosion**. Two-phase alloys (Section 4.2) are more susceptible to galvanic corrosion (because of the physical separation of the two different metals involved) than are single-phase alloys or pure metals. As a result, heat treatment can have a significant influence on corrosion resistance.

Stress-corrosion cracking (Section 2.10.2) is an example of the effect of a corrosive environment on the integrity of a product that, as manufactured, contained residual stresses. Likewise, cold-worked metals are likely to have residual stresses, thus making them more susceptible to corrosion than are hot-worked or annealed metals.

Tool and die materials also can be susceptible to chemical attack by lubricants and coolants. The chemical reaction alters their surface finish and adversely influences the metalworking operation. One example is carbide tools and dies with cobalt as a binder (Section 22.4); the cobalt is attacked by elements in the

Figure 3.4: A selection of coins, manufactured from different metal alloys of copper, nickel, tin, zinc, and aluminum. Valuable metals such as gold and silver are used for coins, but are not used for general currency. *Source:* Scott Gibson/Corbis RF/Alamy Stock Photo.

metalworking fluid, called **selective leaching**. The compatibility of the tool, die, and workpiece materials with the metalworking fluid, under actual operating conditions, is thus an important consideration.

Chemical reactions should not always be regarded as having only adverse effects. Advanced machining processes, such as chemical and electrochemical machining (Chapter 27), are indeed based on controlled chemical reactions. These processes remove material by chemical action in a manner similar to etching of metallurgical specimens. The usefulness of some level of **oxidation** is demonstrated also by the corrosion resistance of aluminum, titanium, and stainless steel. Aluminum, for example, develops a thin (a few atomic layers), strong and adherent hard-oxide film (Al_2O_3) that better protects the surface from further environmental corrosion. Titanium develops a film of titanium oxide (TiO_2); a similar phenomenon occurs in stainless steels which, because of the chromium present in the alloy, develop a protective film. These processes are known as **passivation**. When the protective film is scratched and exposes the metal underneath, a new oxide film begins to form.

Case Study 3.1 Selection of Materials for Coins

There are six general criteria in the selection of materials for coins (Fig. 3.4).

1. The *subjective factors*, such as the *appearance* of the coin, its color, weight, and its ring (the sound made when striking). Also included in this criterion is the feel of the coin. This term is similar in effect to the feel of a fine piece of wood, polished stone, or tableware; It is difficult to quantify because it combines several human factors.

2. The intended *life* of the coin is also a consideration; this duration will reflect resistance to corrosion and to wear (Chapter 33) while the coin is in circulation. These two factors basically determine the span over which the surface imprint of the coin will remain identifiable as well as the ability of the coin to retain its original luster.

3. The *manufacturing* of the coin includes factors such as the formability of the candidate coin materials, the life of the dies used in the coining operation (Section 14.4), and the capability of the materials and processes to resist counterfeiting.

4. Another consideration is the *suitability for use* in coin-operated devices, such as vending machines and turnstiles. These machines are generally equipped with detection devices that test the coins – first, for proper diameter, thickness, and surface condition, and second, for electrical conductivity and density. The coin is rejected if it fails any of these tests.

5. Health issues must be considered. For example, given the large number of population with nickel allergies, Euro coins are minted from nickel-free alloys.

6. A final consideration is the *cost* of raw materials and processing, and whether there is a sufficient *supply* of the coin materials. For example, Canada recently decided it would eliminate the penny because of the high cost of production and its limited currency value. The United States has similar concerns, since a penny (one cent) costs around 1.6 cents to manufacture.

Summary

- Physical properties can have several important influences on materials selection, manufacturing, and on the service life of components. These properties and other relevant characteristics should be considered, because of their possible effects on product design, service requirements, and compatibility with other materials, including tools, dies, and workpieces.
- Combined properties, such as strength-to-weight and stiffness-to-weight ratios, are important factors in selecting materials for lightweight and high-performance structures.
- Thermal conductivity and thermal expansion are major factors in the development of thermal stresses and thermal fatigue and shock, effects that are important in tool and die life in manufacturing operations.
- Chemical reactions, including oxidation and corrosion, are important factors in material selection, design, and manufacturing, as well as in the service life of components. Passivation and stress-corrosion cracking are additional phenomena to be considered.
- Some physical properties are utilized in manufacturing processes and their control, such as the magnetostriction effect (in ultrasonic machining of materials) and the piezoelectric effect (in force transducers and various other sensors).

Key Terms

Conductors
Corrosion
Degradation
Density
Dielectric
Electrical conductivity
Electrical resistivity
Electrorheostatic
Ferromagnetism
Galvanic corrosion
Heat checking
Invar effect
Magnetorheostatic
Magnetostriction
Melting point
Oxidation
Passivation
Piezoelectric effect
Selective leaching
Semiconductors
Smart fluids
Smart materials
Specific heat
Specific stiffness
Specific strength
Stress-corrosion cracking

Superconductivity
Thermal conductivity
Thermal expansion
Thermal fatigue
Thermal stresses

Bibliography

Ashby, M.F., **Materials Selection in Mechanical Design**, 4th ed., Pergamon, 2010.
ASM Handbook, Vol. 13A: **Corrosion: Fundamentals, Testing, and Protection**, 2003.
—, Vol. 13B: *Corrosion Materials*, 2005.
—, Vol. 13B: *Corrosion: Environments and Industries*, 2006, ASM International.
Budinski, K.G., **Engineering Materials: Properties and Selection**, 9th ed., Prentice Hall, 2009.
Fox, M., **Optical Properties of Solids**, Oxford, 2010.
Fulay, P., and Lee, J.-K., **Electronic, Magnetic, and Optical Materials**, CRC Press, 2016.
Hummel, R.E., **Electronic Properties of Materials**, 4th ed., Springer, 2011
Revie, R.W. (ed.), **Uhlig's Corrosion Handbook**, 3rd ed., Wiley-Interscience, 2011.
Shackelford, J. F., **Introduction to Materials Science for Engineers**, 8th ed., Macmillan, 2014.
Solymar, L., Walsh, D., and R.R.A. Syms, **Electrical Properties of Materials,** 9th ed., Oxford, 2014.
White, M.A., **Physical Properties of Materials**, 2nd ed., CRC Press, 2011.

Review Questions

3.1. List several reasons that density is an important material property.

3.2. Explain why the melting point of a material can be an important factor in material selection.

3.3. What adverse effects can be caused by thermal expansion of materials? Give some examples.

3.4. Is thermal cracking the same as thermal shock? Why or why not?

3.5. What is the piezoelectric effect?

3.6. Describe the factors that can lead to the corrosion of a metal.

3.7. What is a superconductor? Describe two applications of superconducting materials.

3.8. What is the difference between thermal conductivity and thermal expansion?

3.9. What is corrosion? How can it be prevented or accelerated?

3.10. What is specific heat? Why is it important in manufacturing?

3.11. Explain stress-corrosion cracking. Why is it also called season cracking?

3.12. What is the difference between a superconductor and a semiconductor?

3.13. What are smart materials?

3.14. What is magnetostriction? What are its uses?

Qualitative Problems

3.15. What is the fundamental difference between mechanical properties of materials, discussed in Chapter 2, and physical properties of materials described in this chapter?

3.16. Describe the significance of structures and machine components made of two materials with different coefficients of thermal expansion.

3.17. Which of the properties described in this chapter are important for (a) pots and pans, (b) cookie sheets for baking, (c) rulers, (d) paper clips, (e) music wire, and (f) beverage cans? Explain your answers.

3.18. Note in Table 3.1 that the properties of the alloys of metals have a wide range compared with the properties of the pure metals. Explain why.

3.19. Rank the following in order of increasing thermal conductivity: aluminum, copper, silicon, titanium, ceramics, and plastics. Comment on how this ranking influences applications of these materials.

3.20. Does corrosion have any beneficial effects? Explain.

3.21. Explain how thermal conductivity can play a role in the development of residual stresses in metals.

3.22. List examples of products where materials that are transparent are desired. List applications for opaque materials.

3.23. Refer to Fig. 3.2 and explain why the trends seen are to be expected.

3.24. Two physical properties that have a major influence on the cracking of workpieces, tools, or dies during thermal cycling are thermal conductivity and thermal expansion. Explain why.

3.25. Which of the materials described in this chapter has the highest (a) density, (b) electrical conductivity, (c) thermal conductivity, (d) specific heat, (e) melting point, and (f) cost.

3.26. Is oxidation beneficial with respect to corrosion prevention, or is it part of the corrosion process? Explain.

3.27. Which properties described in this chapter can be affected by applying a coating?

Quantitative Problems

3.28. If we assume that all the work done in plastic deformation is converted into heat, the temperature rise in a workpiece is (1) directly proportional to the work done per unit volume and (2) inversely proportional to the product of the specific heat and the density of the workpiece. Using Fig. 2.5, and letting the areas under the curves be the unit work done, calculate the temperature rise for (a) 8650 steel, (b) 304 stainless steel, and (c) 1100-H14 aluminum.

3.29. The natural frequency, f, of a cantilever beam is given by

$$f = 0.56\sqrt{\frac{EIg}{wL^4}},$$

where E is the modulus of elasticity, I is the moment of inertia, g is the gravitational constant, w is the weight of the beam per unit length, and L is the length of the beam. How does the natural frequency of the beam change, if at all, as its temperature is increased? Assume that the material is steel.

3.30. Plot the following for the materials described in this chapter: elastic modulus versus density, yield stress versus density, thermal conductivity versus density. Comment on the implications of these plots.

3.31. It can be shown that thermal distortion in precision devices is low for high values of thermal conductivity divided by the thermal expansion coefficient. Rank the materials in Table 3.1 according to their ability to resist thermal distortion.

3.32. Add a column to Table 3.1 that lists the volumetric heat capacity of the materials listed, expressed in units of $\mathrm{J/cm^3\ K}$. Compare the results to the value for liquid water ($4.184\ \mathrm{J/cm^3\ K}$). Note that the volumetric heat capacity of a material is the product of its density and specific heat.

3.33. Using strength and density data, determine the minimum weight of a 2 m-long tension member that must support a load of 8 kN, manufactured from (a) annealed 303 stainless steel; (b) normalized 8620 steel; (c) as-rolled 1080 steel; (d) 5052-O aluminum alloy; (e) AZ31B-F magnesium; and (f) pure copper.

3.34. Plot the thermal conductivity against electrical conductivity for the materials in Table 3.1. Is there a correlation? Explain.

Synthesis, Design, and Projects

3.35. Conduct a literature search and add the following materials to Table 3.1: cork, cement, ice, sugar, lithium, graphene, and chromium.

3.36. From your own experience, make a list of parts, components, or products that have corroded and have had to be replaced or discarded.

3.37. List applications where the following properties would be desirable: (a) high density, (b) low density, (c) high melting point, (d) low melting point, (e) high thermal conductivity, and (f) low thermal conductivity.

3.38. Describe several applications in which both specific strength and specific stiffness are important.

3.39. Design several mechanisms or instruments based on utilizing the differences in thermal expansion of materials, such as bimetallic strips that develop a curvature when heated.

3.40. For the materials listed in Table 3.1, determine the specific strength and specific stiffness. Describe your observations.

3.41. The maximum compressive force that a lightweight column can withstand before buckling depends on the ratio of the square root of the stiffness to the density for the material. For the materials listed in Table 2.2, determine (a) the ratio of tensile strength to density and (b) the ratio of elastic modulus to density. Comment on the suitability of each for being made into lightweight columns.

3.42. Describe possible applications and designs using alloys exhibiting the Invar effect of low thermal expansion.

3.43. Collect some pieces of different metallic and nonmetallic materials listed in Table 3.2. Using simple tests and/or instruments, determine the validity of the descending order of the physical properties shown in the table.

3.44. Add the following materials to Table 3.1: (a) uranium; (b) lithium; (c) sodium.

3.45. Design an actuator to turn on a switch when the temperature drops below a certain level. Use two materials with different coefficients of thermal expansion in your design.

3.46. Conduct an Internet and technical literature review and write a one-page paper highlighting applications of piezoelectric materials.

3.47. It has been widely reported that mechanical properties such as strength and ductility can be very different for micro-scale devices than are measured at normal length scales. Explain whether or not you would expect the physical properties described in this chapter to be scale dependent.

3.48. If you were given a metal (not an alloy), and asked to identify it, list (in order) the experiments or measurements you would perform. Explain what influence the shape of the metal would have on your prioritization.

Chapter 4

Metal Alloys: Their Structure and Strengthening by Heat Treatment

- This chapter reviews the structures of metal alloys, including solid solutions, intermetallic compounds, and two-phase systems.
- Phase diagrams show graphically the phases that develop as a function of alloy composition and temperature.
- The system of iron and carbon and the phases involved are described in detail.
- Heat treatment of metals is a common method of improving mechanical properties; it involves controlled heating and cooling, transforming a microstructure into a different phase.
- Some metals, such as aluminum and stainless steels, can be heat treated only by precipitation, hardening, or aging.
- Improving the ductility of a material is at the expense of such properties as strength or hardness.
- The chapter ends with a review of the characteristics of heat-treating equipment.

4.1 Introduction

The properties and behavior of metals and alloys during manufacturing and their performance during their service life depend on their composition, structure, and their processing history, including the heat treatment to which they have been subjected. Important properties, such as strength, hardness, ductility, toughness, and resistance to wear, are greatly influenced by alloying elements and the heat-treatment processes employed.

The most common example of a process that improves properties is *heat treatment* (Sections 4.7–4.10), which modifies microstructures. A variety of mechanical properties important to manufacturing can then develop, such as improved formability, machinability, or increased strength and hardness to improve the performance of tools and dies. These properties also enhance service performance of machine components, such as gears, cams, and shafts (Fig. 4.1).

This chapter follows the outline shown in Fig. 4.2, beginning with the role of various alloying elements, the solubility of one element in another, phases, equilibrium phase diagrams, and the influence of composition, temperature, and time. The chapter also describes methods and techniques of heating, quenching, tempering, and annealing of metals and alloys, and the characteristics of the equipment involved.

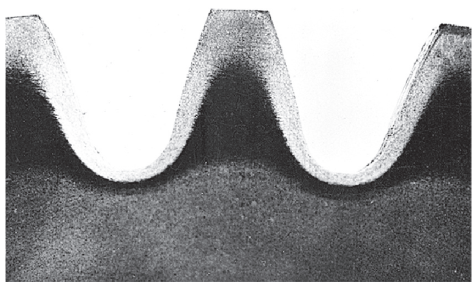


Figure 4.1: Cross-section of gear teeth showing induction-hardened surfaces. *Source:* TOCCO Div., Park-Ohio Industries, Inc.

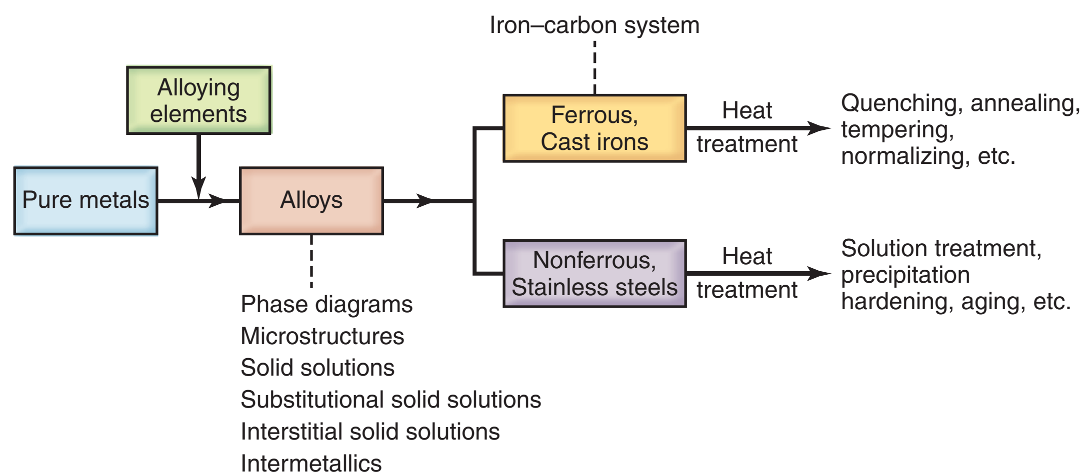


Figure 4.2: Outline of topics described in Chapter 4.

4.2 Structure of Alloys

When describing the basic crystal structure of metals, in Chapter 1, it was noted that the atoms are all of the *same* type, except for the presence of impurity atoms. These metals are known as **pure metals**, even though they may not be completely pure. *Commercially pure* metals are used for various purposes, such as aluminum for foil, copper for electrical conductors, nickel or chromium for plating, and gold for electrical contacts. Pure metals have somewhat limited properties but they can be enhanced or modified by **alloying**. The majority of metals used in engineering applications are some form of alloy. An **alloy** consists of two or more chemical elements, at least one of which is a metal. Alloying consists of two basic forms: *solid solutions* and *intermetallic compounds*.

4.2.1 Solid Solutions

Two terms are essential in describing alloys: **solute** and **solvent**. The solute is the *minor* element (such as salt or sugar) that is added to the solvent, which is the *major* element (such as water). In terms of the elements in a crystal structure, the solute (composed of *solute atoms*) is the element that is added to the solvent (composed of *host atoms*). When the particular crystal structure of the solvent is maintained during alloying, the alloy is called a *solid solution*.

Substitutional Solid Solutions. If the size of the solute atom is similar to that of the solvent atom, the solute atoms can replace solvent atoms and form a *substitutional solid solution* (Fig. 1.8). An example is brass (Section 6.4), which is an alloy of zinc and copper in which zinc (the solute atom) is introduced into the lattice of copper (the solvent atom). The properties of brass can thus be modified by controlling the amount of zinc in copper.

Interstitial Solid Solutions. If the size of the solute atom is much smaller than that of the solvent atom, each solute atom can occupy an *interstitial* position, forming an *interstitial solid solution*. An important family of interstitial solid solutions is **steel** (Chapter 5), which is an alloy of iron and carbon in which the carbon atoms are present in interstitial positions between iron atoms. The atomic radius of carbon is 0.071 nm, which is very small as compared to the 0.124 nm radius of the iron atom. The properties of carbon steels can be varied over a wide range by adjusting the ratio of carbon to iron. The ability to control this ratio is a major reason why steel is such a versatile and useful material with a very wide range of properties and applications.

4.2.2 Intermetallic Compounds

Intermetallic compounds are complex structures, consisting of two metals in which solute atoms are present among solvent atoms in specific proportions. Typical examples are aluminides of titanium (Ti_3Al), nickel (Ni_3Al), and iron (Fe_3Al). Some intermetallic compounds have solid solubility, and the type of their atomic bond may range from metallic to ionic. Intermetallic compounds are strong, hard, and brittle. Because of their high melting points, strength at elevated temperatures, good oxidation resistance, and relatively low density, intermetallic compounds are candidate materials for applications such as advanced gas-turbine engines.

4.2.3 Two-phase Systems

Recall that a solid solution is one in which two or more elements form a single homogeneous solid phase, in which the elements are uniformly distributed throughout the solid mass. Such a system has a maximum concentration of solute atoms in the solvent-atom lattice, just as there is a solubility limit for sugar in water. Most alloys consist of two or more solid phases, and they may be regarded as mechanical mixtures; such a system with two solid phases is known as a *two-phase system*.

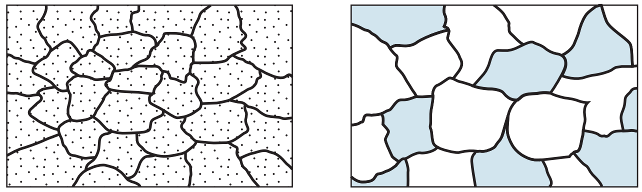


Figure 4.3: (a) Schematic illustration of grains, grain boundaries, and particles dispersed throughout the structure of a two-phase system, such as a lead–copper alloy. The grains represent lead in solid solution in copper, and the particles are lead as a second phase. (b) Schematic illustration of a two-phase system consisting of two sets of grains: dark and light. The blue grains and the white grains have separate compositions and properties.

A **phase** is defined as a physically distinct and homogeneous portion in a material; each phase is a homogeneous part of the total mass, and has its own characteristics and properties. Consider a mixture of sand and water as an example of a two-phase system; these two very different components have their own distinct structures, characteristics, and properties. There is a clear boundary in this mixture between the water (one phase) and the sand particles (the second phase). Another example is ice in water: the two phases have exactly the same chemical elements (hydrogen and oxygen) even though their properties are very different. Note that it is not necessary for one phase to be a liquid; for example, sand suspended in ice is also a two-phase system.

An example of a two-phase system in metals occurs when lead is added to copper in the molten state. After the mixture solidifies, the structure consists of two phases: (a) one having a small amount of lead in solid solution in copper, and (b) the other having lead particles (roughly spherical in shape) *dispersed* throughout the structure (Fig. 4.3a). The lead particles are analogous to sand particles in water, described above. The copper–lead alloy has properties that are different from those of either copper or lead alone.

Alloying with finely dispersed particles (the **second-phase particles**) is an important method of strengthening alloys and controlling their properties. In two-phase alloys, the second-phase particles become obstacles to dislocation movement, and thus increase the strength of the alloy. Figure 4.3b shows another example of a two-phase alloy, which is an aggregate structure where there are two sets of grains, each with its own composition and properties. The darker grains in the figure may, for example, have a different structure than the lighter grains; they may be brittle, while the lighter grains are ductile.

Defects may develop during metalworking operations, such as forging or extrusion (Chapters 14 and 15). Such flaws may be due to the lack of ductility of one of the phases in the alloy. In general, two-phase alloys are stronger and less ductile than solid solutions.

4.3 Phase Diagrams

Pure metals have clearly defined melting or freezing points, and solidification takes place at a *constant temperature*. When the temperature of a molten metal is reduced to the freezing point, the energy of the *latent heat of solidification* is given off while the temperature remains constant. Eventually, solidification is complete and the solid metal continues cooling to ambient (room) temperature.

Unlike pure metals, alloys solidify over a *range* of temperatures (Fig. 4.4). Solidification begins when the temperature of the molten metal drops below the liquidus temperature, and is completed when the temperature reaches the **solidus**. Within this temperature range, the alloy is in a *mushy* or *pasty* state; its composition and state are then described by the particular alloys phase diagram.

A phase diagram, also called an **equilibrium** or constitutional diagram, shows the relationships among temperature, the composition, and the phases present in a particular alloy system at equilibrium.

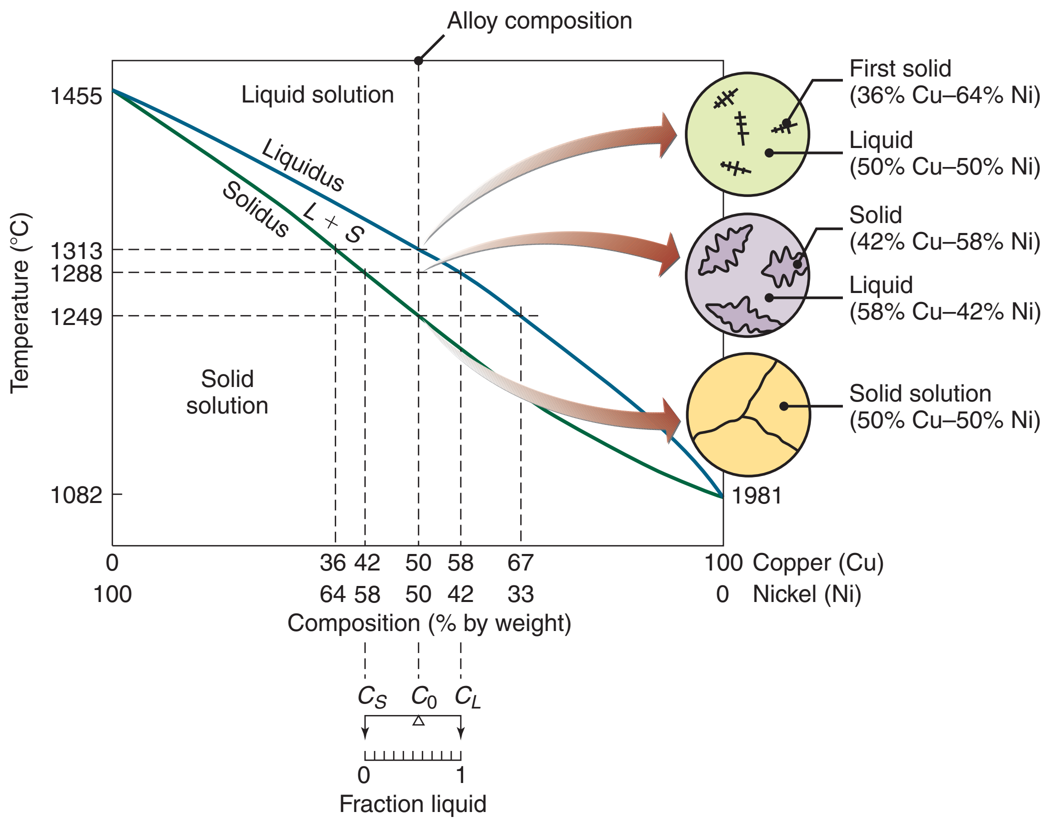


Figure 4.4: Phase diagram for nickel–copper alloy system obtained at a slow rate of solidification. Note that pure nickel and pure copper each have a specific freezing or melting temperature. The top circle on the right depicts the nucleation of crystals; the second circle shows the formation of dendrites (see Section 10.2.2). The bottom circle shows the solidified alloy, with grain boundaries.

Equilibrium means that the state of a system does not vary with time. The word *constitutional* indicates the relationships among the structure, the composition, and the physical makeup of the alloy. As described in detail below, types of phase diagrams include those for (a) complete solid solutions; (b) **eutectics**, such as cast irons; and (c) eutectoids, such as steels.

An example of a phase diagram is shown in Fig. 4.4 for the copper–nickel alloy; it is called a **binary phase diagram** because there are two elements (copper and nickel) present in the system. The left boundary of this diagram (100% Ni) indicates the melting point of pure nickel; the right boundary (100% Cu) indicates the melting point of pure copper. (All percentages are by weight, not by number of atoms.)

Lever Rule. The *composition* of various phases in a phase diagram can be determined by a procedure called the *lever rule*. As shown in the lower portion of Fig. 4.4, the procedure is to first construct a lever between the solidus and liquidus lines (called *tie line*), balanced (on the triangular support) at the nominal weight composition C_o of the alloy. The left end of the lever represents the composition C_S of the *solid* phase and the right end of the composition C_L of the *liquid* phase. Note from the graduated scale that the liquid fraction is also indicated along the tie line, ranging from 0 at the left (fully solid) to 1 at the right (fully liquid).

The lever rule states that the **weight fraction of solid** is proportional to the distance between C_o and C_L:

$$\frac{S}{S+L} = \frac{C_o - C_L}{C_S - C_L} \tag{4.1}$$

Likewise, the **weight fraction of liquid** is proportional to the distance between C_S and C_o, hence

$$\frac{L}{S+L} = \frac{C_S - C_o}{C_S - C_L} \tag{4.2}$$

Note that these quantities are fractions, and they must be multiplied by 100 to determine percentages.

From inspection of the tie line in Fig. 4.4 (and for a nominal alloy composition of $C_o = 50\%$ Cu–50% Ni) it can be noted that, because C_o is closer to C_L than it is to C_S, the solid phase contains less copper than does the liquid phase. By measurement on the phase diagram and using the lever-rule equations, it can be seen that the composition of the solid phase is 42% Cu and of the liquid phase is 58% Cu, as stated in the middle circle at the right in Fig. 4.4. These calculations refer to copper. Reversing the phase diagram in the figure, so that the left boundary is 0% nickel (whereby nickel now becomes the alloying element in copper) will give the compositions of the solid and liquid phases in terms of nickel. The lever rule is also known as the *inverse lever rule* because, as indicated by Eqs. (4.1) and (4.2), the amount of each phase is proportional to the length of the opposite end of the lever.

The *completely solidified* alloy in the phase diagram shown in Fig. 4.4 is a *solid solution*, because the alloying element, Cu (the solute atom), is completely dissolved in the host metal, Ni (the solvent atom), and each grain has the same composition. The atomic radius of copper is 0.128 nm and that of nickel is 0.125 nm, and both elements have a face-centered cubic structure; thus, they readily form solid solutions.

The mechanical properties of solid solutions of Cu–Ni depend on their composition (Fig. 4.5). The properties of pure copper are, up to a limit, improved upon by increasing the nickel content; thus, there is an optimal percentage of nickel that gives the highest strength and hardness to the Cu–Ni alloy.

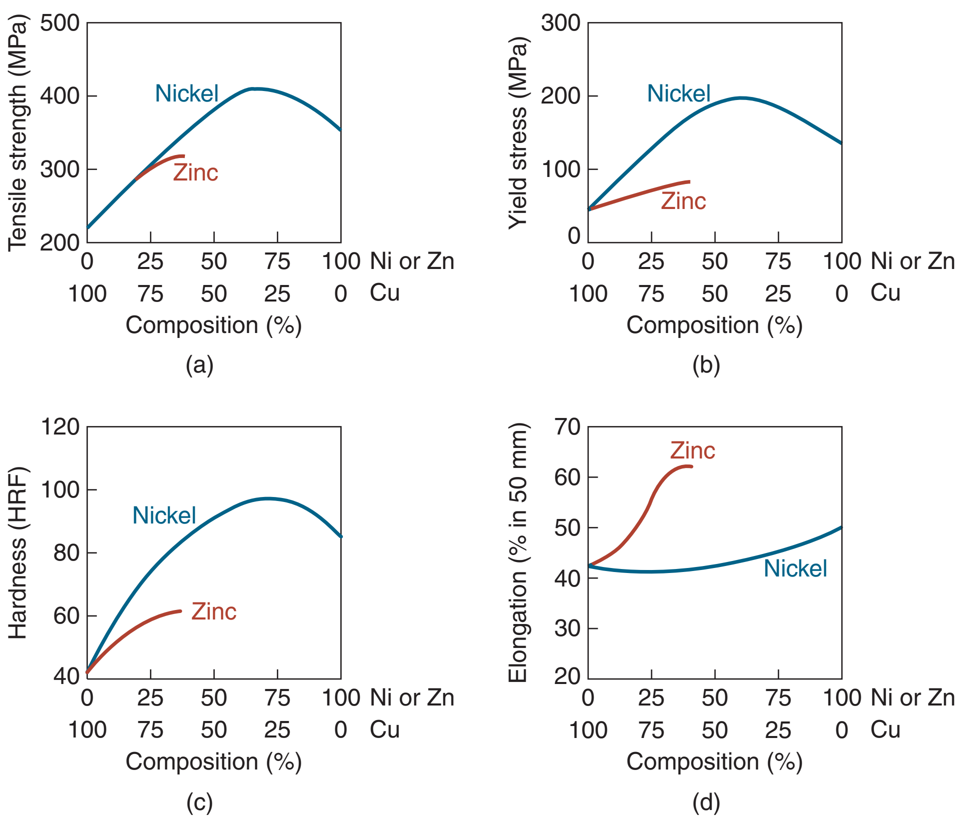


Figure 4.5: Mechanical properties of copper–nickel and copper–zinc alloys as a function of their composition. The curves for zinc are short, because zinc has a maximum solid solubility of 40% in copper.

Figure 4.5 shows how zinc, as an alloying element in copper, affects the mechanical properties of the alloy. Note the maximum of 40% solid solubility for zinc (solute) in copper (solvent), whereas copper and nickel are completely soluble in each other. The improvements in properties are due to *pinning* (blocking) of dislocations (Section 1.4.1) at substitutional nickel or zinc atoms, which may also be regarded as impurity atoms. As a result, dislocations cannot move as freely, and thus the strength of the alloy increases.

4.4 The Iron–Carbon System

Steels and cast irons are represented by the iron–carbon binary system. Commercially pure iron contains up to 0.008% C, steels up to 2.11% C, and cast irons up to 6.67% C, although most cast irons contain less than 4.5% C. The iron–carbon system is described in this section, including the techniques employed to evaluate and modify the properties of these important materials, for specific applications.

The **iron–iron-carbide phase diagram** is shown in Fig. 4.6. Although this diagram can be extended to the right — to 100% C (pure graphite); see Fig. 4.10 — the range that is significant to engineering applications is up to 6.67% C, because Fe_3C is a stable phase. Pure iron melts at a temperature of 1538°C, as shown at the left boundary in Fig. 4.6. As iron cools, it first forms delta ferrite, then austenite, and finally alpha ferrite.

Ferrite. Alpha ferrite, also denoted α-ferrite or simply **ferrite**, is a solid solution of body-centered cubic (bcc) iron; it has a maximum solid solubility of 0.022% C at a temperature of 727°C. Just as there is a solubility limit for salt in water (with any extra amount precipitating as solid salt at the bottom of a container), there is a solid solubility limit for carbon in iron.

Ferrite is relatively soft and ductile; it is magnetic from room temperature to 768°C, the so-called *Curie temperature* (after M. Curie, 1867–1934). Although very little carbon can dissolve interstitially in bcc iron, the amount of carbon can significantly affect the mechanical properties of ferrite. Furthermore, significant amounts of chromium, manganese, nickel, molybdenum, tungsten, and silicon can be contained in iron in solid solution, imparting special properties.

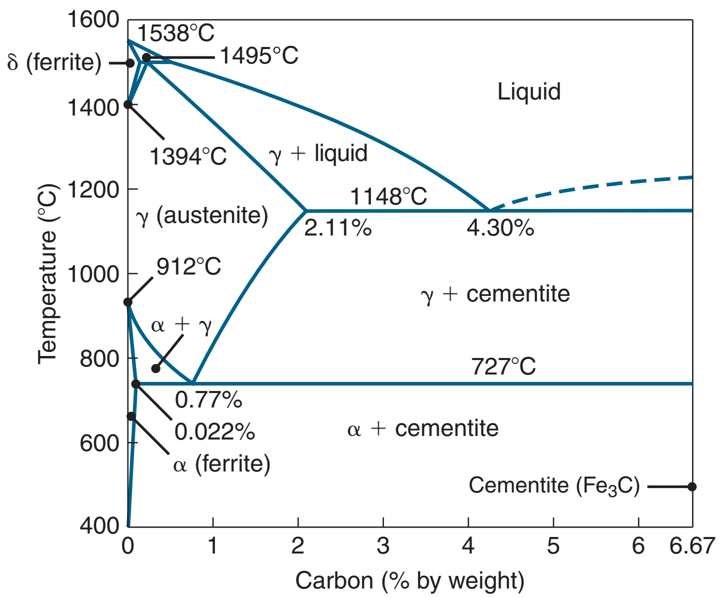


Figure 4.6: The iron–iron-carbide phase diagram. Because of the importance of steel as an engineering material, this diagram is one of the most important of all phase diagrams.

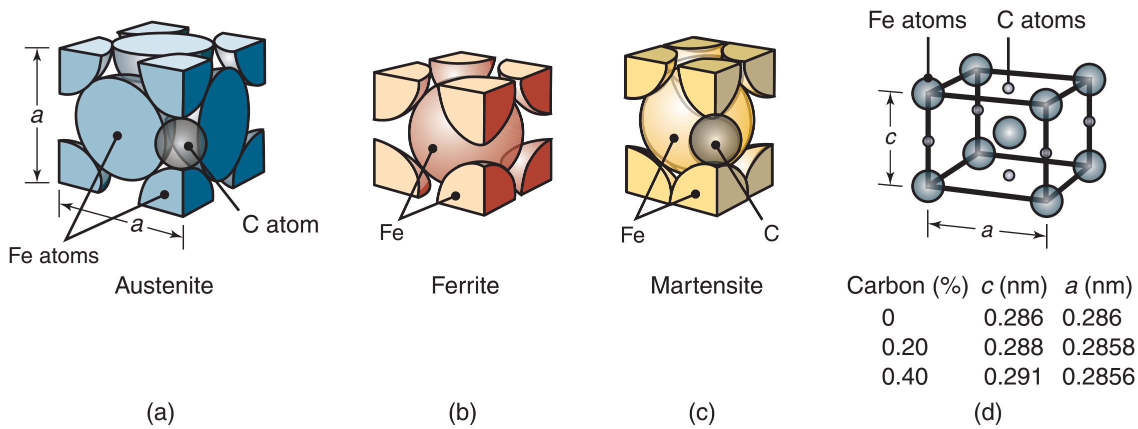


Figure 4.7: The unit cells for (a) austenite, (b) ferrite, and (c) martensite. The effect of percentage of carbon (by weight) on lattice dimensions for martensite is shown in (d). Note the interstitial position of the carbon atoms (see Fig. 1.8). Note also the increase in dimension c with increasing carbon content; this effect causes the unit cell of martensite to be in the shape of a rectangular prism.

Austenite. As shown in Fig. 4.6, within a certain temperature range iron undergoes a **polymorphic transformation** from a bcc to an fcc structure, becoming *gamma iron* (γ-iron) or, more commonly, **austenite** (after W.R. Austen, 1843–1902). This structure has a solid solubility of up to 2.11% C at 1148°C. Because the fcc structure has more interstitial positions, the solid solubility of austenite is about two orders of magnitude higher than that of ferrite, with the carbon occupying the interstitial positions, as shown in Fig. 4.7a.

Austenite is an important phase in heat treatment of steels (Section 4.7). It is denser than ferrite, and its single-phase fcc structure is ductile at elevated temperatures, thus possessing good formability. Large amounts of nickel and manganese can be dissolved in fcc iron, to impart various properties. Steel is nonmagnetic in austenitic form, either at high temperatures or, for austenitic stainless steels, at room temperature.

Cementite. The right boundary of Fig. 4.6 represents **cementite**, which is 100% iron carbide (Fe_3C) with a carbon content of 6.67%. Cementite, from the Latin *caementum* (meaning stone chips), is also called **carbide** (not be confused with other carbides, used as dies, cutting tools, and abrasives, such as tungsten carbide, titanium carbide, and silicon carbide; Chapters 8 and 22). Cementite is a very hard and brittle intermetallic compound and has a significant influence on the properties of steels.

4.5 The Iron–Iron-carbide Phase Diagram and the Development of Microstructures in Steels

The region of the iron–iron-carbide phase diagram that is significant for steels is shown in Fig. 4.8 (an enlargement of the lower left-hand portion of Fig. 4.6). Various microstructures can be developed, depending on (a) carbon content, (b) amount of plastic deformation (working), and (c) method of heat treatment. For example, consider the eutectic point of iron with a 0.77% C content, while it is being cooled very slowly from a temperature of, say, 1100°C in the austenite phase; the reason for very slow cooling is to maintain equilibrium.

At 727°C, a reaction takes place in which austenite is transformed into alpha ferrite (bcc) and cementite. Because the solid solubility of carbon in ferrite is only 0.022%, the extra carbon forms cementite;

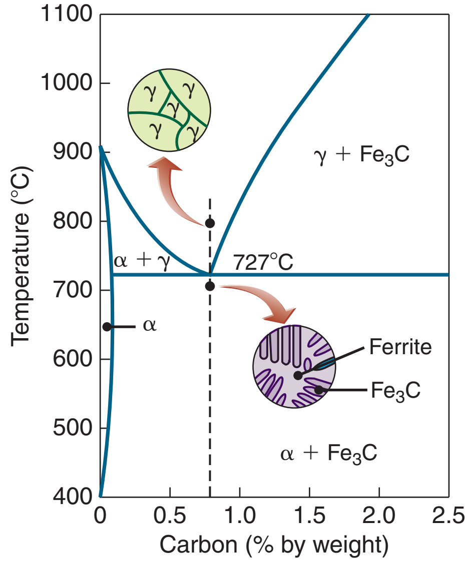


Figure 4.8: Schematic illustration of the microstructures for an iron–carbon alloy of eutectoid composition (0.77% carbon), above and below the eutectoid temperature of 727°C.

this reaction is called a **eutectoid** (meaning *eutecticlike*). This reaction indicates that at a certain temperature, a single solid phase (austenite) is transformed into two other solid phases (ferrite and cementite). The structure of eutectoid steel is called pearlite, because, at low magnifications, it resembles mother-of-pearl (Fig. 4.9). The microstructure of pearlite consists of alternating layers (lamellae) of ferrite and cementite; consequently, the mechanical properties of pearlite are intermediate between those of ferrite (soft and ductile) and cementite (hard and brittle).

4.5.1 Effects of Alloying Elements in Iron

Although carbon is the basic element that transforms iron into steel, other elements are added to impart a variety of desirable properties. The main effect of these alloying elements on the iron–iron-carbide phase diagram is to shift the eutectoid temperature and eutectoid composition (percentage of carbon in steel at the eutectoid point); these elements shift other phase boundaries as well.

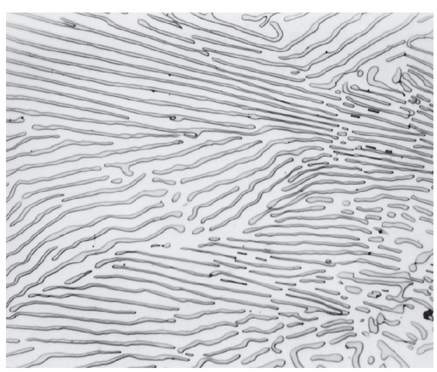


Figure 4.9: Microstructure of pearlite in 1080 steel, formed from austenite of eutectoid composition. In this lamellar structure, the lighter regions are ferrite and the darker regions are carbide. Magnification: $2500\times$.

The eutectoid temperature may be raised or lowered from 727°C, depending on the particular alloying element. Alloying elements always lower the eutectoid composition; that is, its carbon content is lower than 0.77%. Lowering the eutectoid temperature means increasing the austenite range; as a result, an alloying element such as nickel is known as an **austenite former**. Because nickel has an fcc structure, it favors the fcc structure of austenite. Conversely, chromium and molybdenum have a bcc structure, thus favoring the bcc structure of ferrite; these elements are known as **ferrite stabilizers**.

4.6 Cast Irons

The term **cast iron** refers to a family of ferrous alloys composed of iron, carbon (ranging from 2.11% to about 4.5%), and silicon (up to about 3.5%). Cast irons are classified according to their solidification morphology from the eutectic temperature (see also Section 12.3.2):

1. Gray cast iron or gray iron
2. Ductile cast iron, also called nodular cast iron or spheroidal graphite cast iron
3. White cast iron
4. Malleable iron
5. Compacted graphite iron.

Cast irons are also classified by their structure: ferritic, pearlitic, quenched and tempered, or austempered.

The equilibrium phase diagram relevant to cast irons is shown in Fig. 4.10, in which the right boundary is 100% C — that is, pure graphite. Because the eutectic temperature is 1154°C, cast irons are completely liquid at temperatures lower than those required for liquid steels; consequently, iron with high carbon content can be cast (see Part II) at lower temperatures than can steels.

Cementite is **metastable**, not completely stable, with an extremely low rate of decomposition. It can, however, be made to decompose into alpha ferrite and graphite. The formation of graphite (**graphitization**) can be controlled, promoted, and accelerated by modifying the composition and the rate of cooling, and by the addition of silicon.

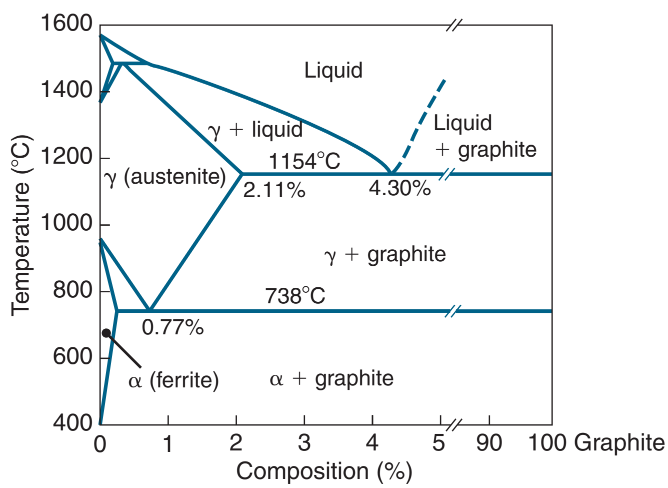


Figure 4.10: Phase diagram for the iron–carbon system with graphite (instead of cementite) as the stable phase. Note that this figure is an extended version of Fig. 4.6.

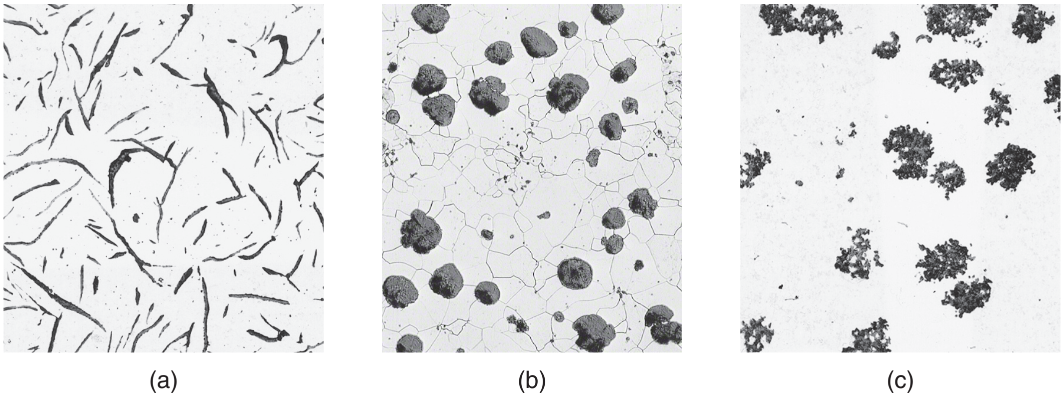


Figure 4.11: Microstructure for cast irons. Magnification: $100\times$. (a) Ferritic gray iron with graphite flakes. (b) Ferritic ductile iron (nodular iron), with graphite in nodular form. (c) Ferritic malleable iron; this cast iron solidified as white cast iron, with the carbon present as cementite, and was heat treated to graphitize the carbon.

Gray Cast Iron. In this structure, graphite exists largely in the form of *flakes* (Fig. 4.11a). It is called **gray cast iron** or **gray iron**, because, when broken, the fracture path is along the graphite flakes, and has a gray, sooty appearance. These flakes act as stress raisers; as a result, gray iron has negligible ductility and is weak in tension, although strong in compression. On the other hand, the presence of graphite flakes gives gray iron the important capacity to dampen vibrations (by internal friction); thus gray cast iron is a suitable and commonly used material for constructing machine-tool bases and machinery structures (Section 25.3).

Three types of gray cast iron are **ferritic, pearlitic,** and **martensitic**. Because of their different structures, each has different properties and applications. In ferritic gray iron (also known as *fully gray iron*), the structure consists of graphite flakes in an alpha-ferrite matrix. Pearlitic gray iron has a structure of graphite in a matrix of pearlite, and although still brittle, it is stronger than fully gray iron. Martensitic gray iron is obtained by austenitizing a pearlitic gray iron and then quenching it rapidly to produce a structure of graphite in a martensite matrix; as a result, it is very hard.

Ductile (Nodular) Iron. In this structure, graphite is in a **nodular** or **spheroid** form (Fig. 4.11b), which permits the material to be somewhat ductile and shock resistant. The shape of the graphite flakes can be modified into nodules (spheres) by small additions of magnesium and/or cerium to the molten metal prior to pouring. Ductile iron can be made ferritic or pearlitic by heat treatment; it can also be heat treated to obtain a structure of tempered martensite (Section 4.7).

White Cast Iron. *White cast iron* is obtained either by cooling gray iron rapidly or by adjusting the composition by keeping the carbon and silicon content low. It is also called *white iron* because of the white crystalline appearance of the fracture surface. The structure is very hard, wear resistant, and brittle, because of the presence of large amounts of iron carbide, instead of graphite.

Malleable Iron. *Malleable iron* is obtained by annealing white cast iron in an atmosphere of carbon monoxide and carbon dioxide, at between 800° and 900°C, for up to several hours, depending on the size of the part. During this process, the cementite decomposes (*dissociates*) into iron and graphite. The graphite exists as *clusters* or *rosettes* (Fig. 4.11c) in a ferrite or pearlite matrix. Consequently, malleable iron has a structure similar to that of nodular iron, promoting good ductility, strength, and shock resistance — hence, the term *malleable* (from the Latin *malleus* meaning it can be hammered).

Compacted-graphite Iron. The graphite in this structure is in the form of short, thick, interconnected flakes with undulating surfaces and rounded extremities. The mechanical and physical properties of this cast iron are intermediate between those of flake-graphite and nodular-graphite cast irons.

4.7 Heat Treatment of Ferrous Alloys

The microstructures described thus far can be modified by **heat-treatment—**that is, by controlled heating and cooling of the alloys at various rates. These treatments induce **phase transformations**, which greatly influence such mechanical properties as strength, hardness, ductility, toughness, and wear resistance. The specific effects of thermal treatment depend on the particular alloy, its composition and microstructure, the degree of prior cold work, and the rates of heating and cooling during heat treatment.

This section focuses on the microstructural changes in the iron–carbon system. Because of their technological significance, the structures considered are pearlite, spheroidite, bainite, martensite, and tempered martensite. The heat-treatment processes described are annealing, quenching, and tempering.

Pearlite. If the ferrite and cementite lamellae in the pearlite structure of the eutectoid steel, shown in Fig. 4.9, are thin and closely packed, the microstructure is called **fine pearlite**; if they are thick and widely spaced, it is called **coarse pearlite**. The difference between the two depends on the rate of cooling through the eutectoid temperature, which is the site of a reaction in which austenite is transformed into pearlite. If the rate of cooling is relatively high (as in air), the structure is fine pearlite; if cooling is slow (as in a furnace), coarse pearlite is produced.

Spheroidite. When pearlite is heated to just below the eutectoid temperature and then held at that temperature for a period of time (called **subcritical annealing**, Section 4.11), such as at 700°C for a day, the cementite lamellae transform to roughly spherical shapes (Fig. 4.12). Unlike the lamellar shapes of cementite, which act as stress raisers, **spheroidites** (spherical particles) have smaller stress concentrations because of their rounded shapes. Consequently, this structure has higher toughness and lower hardness than the pearlite structure. It can be cold worked, because the ductile ferrite has high toughness and the spheroidal carbide particles prevent the initiation of cracks within the the structure.

Bainite. Visible only through electron microscopy, *bainite* is a very fine microstructure, consisting of ferrite and cementite, similar to pearlite, but with a different morphology. Bainite can be produced in steels adding alloying elements and at cooling rates higher than those required for pearlite. This structure, called **bainitic**

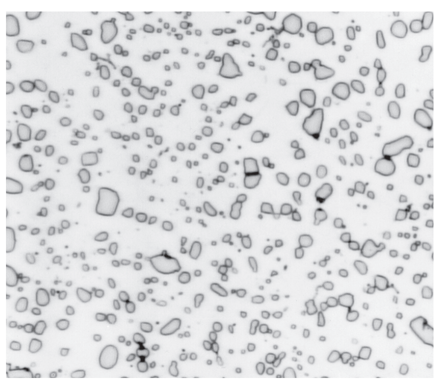


Figure 4.12: Microstructure of eutectoid steel. Spheroidite is formed by tempering the steel at 700°C. Magnification: $1000\times$.

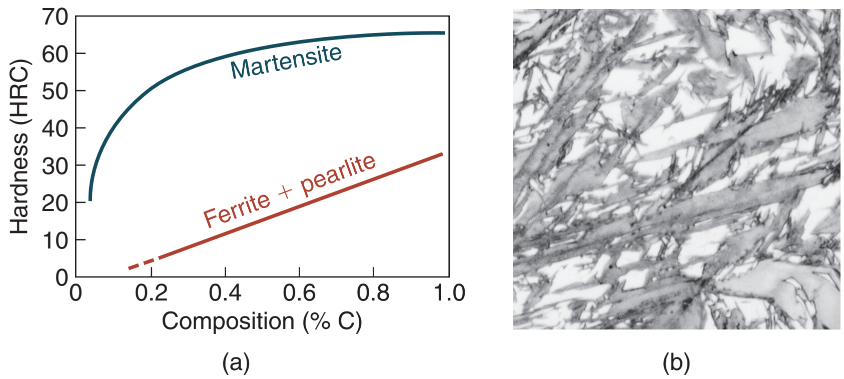


Figure 4.13: (a) Hardness of martensite as a function of carbon content. (b) Micrograph of martensite containing 0.8% carbon. The gray platelike regions are martensite; they have the same composition as the original austenite (white regions). Magnification: 1000×.

steel (after E.C. Bain, 1891–1971), is generally stronger and more ductile than pearlitic steels at the same hardness level.

Martensite. When austenite is cooled at a high rate, such as by quenching in water, its fcc structure is transformed into a **body-centered tetragonal** (bct) structure. This structure can be described as a body-centered rectangular prism that is slightly elongated along one of its principal axes (see Fig. 4.7d). This microstructure is called *martensite* (after A. Martens, 1850–1914). Because martensite does not have as many slip systems as a bcc structure, and the carbon is in interstitial positions, it is extremely hard and brittle (Fig. 4.13). Martensite transformation takes place almost instantaneously, because it involves not the diffusion process but a slip mechanism, and thus allowing plastic deformation. This is a time-dependent phenomenon that is the mechanism in other transformations as well.

Retained Austenite. If the temperature to which the alloy is quenched is not sufficiently low, only a portion of the structure is transformed to martensite. The rest is *retained austenite*, which is visible as white areas in the structure, along with the dark, needlelike martensite. Retained austenite can cause dimensional instability and cracking, and lower the hardness and strength of the alloy.

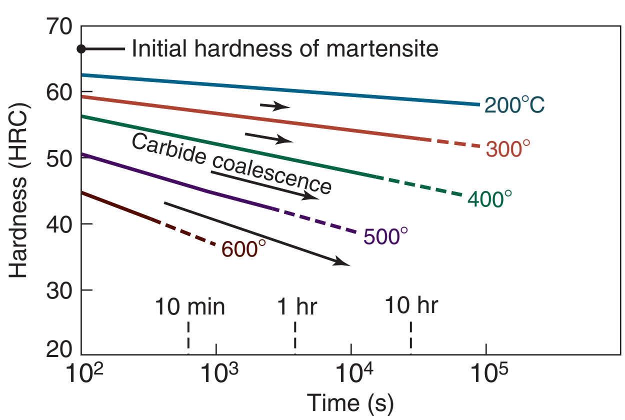


Figure 4.14: Hardness of tempered martensite as a function of tempering time for 1080 steel quenched to 65 HRC. Hardness decreases because the carbide particles coalesce and grow in size, thereby increasing the interparticle distance of the softer ferrite.

Tempered Martensite. Martensite is tempered to improve its mechanical properties. *Tempering* is a process by which hardness is reduced and toughness is increased. The body-centered tetragonal martensite is heated to an intermediate temperature, typically 150° to 650°C, where it decomposes to a two-phase microstructure, consisting of body-centered cubic alpha ferrite and small particles of cementite. With increasing tempering time and temperature, the hardness of tempered martensite decreases (Fig. 4.14). The reason is that the cementite particles coalesce and grow, and the distance between the particles in the soft ferrite matrix increases as the less stable and smaller carbide particles dissolve.

4.7.1 Time–Temperature–Transformation Diagrams

The percentage of austenite transformed into pearlite as a function of temperature and time (Fig. 4.15a). This transformation is best illustrated by Fig. 4.15b and c in diagrams called **isothermal transformation (IT) diagrams**, or *time–temperature–transformation* (TTT) *diagrams*, constructed from the data given in Fig. 4.15a. The higher the temperature or the longer the time, the more austenite is transformed into pearlite. Note that, for each temperature, there is a minimum time for the transformation to begin; this time period defines the *critical cooling rate*; with longer times, austenite begins to transform into pearlite, as can be traced in Figs. 4.15b and c.

The TTT diagrams shown allow the design of heat treatment schedules to obtain desirable microstructures. Consider, for example, the TTT curves shown in Fig. 4.15c. The steel can be raised to a very high temperature (above the eutectic temperature) to start with a state of austenite. If the steel is cooled rapidly, it can follow the 140°C/s cooling rate trajectory shown, resulting in complete martensite. On the other hand, it can be more slowly cooled (in a molten salt bath) to develop pearlite- or bainite-containing steels. For tempered martensite, the heat treatment and quench stages will be followed by a tempering process.

The differences in hardness and toughness of the various structures developed are shown in Fig. 4.16; note that fine pearlite is harder and less ductile than coarse pearlite. The effects of various percentages of carbon, cementite, and pearlite on other mechanical properties of steels are shown in Fig. 4.17.

4.8 Hardenability of Ferrous Alloys

The capability of an alloy to be hardened by heat treatment is called its **hardenability**, and is a measure of the *depth* of hardness that can be obtained by heating and subsequent quenching. Note that the term hardenability should not be confused with hardness, which is the resistance of a material to indentation or scratching (Section 2.6). It can be seen that hardenability of ferrous alloys depends on their (a) carbon content, (b) grain size of the austenite, (c) alloying elements present in the material, and (d) the cooling rate.

4.8.1 The End-quench Hardenability Test

In this commonly used **Jominy test** (after W.E. Jominy, 1893–1976), a round test bar 100 mm long, made from the particular alloy, is **austenitized** — that is, heated to the proper temperature to form 100% austenite. It is then quenched directly at one end (Fig. 4.18a) with a stream of water at 24°C. The cooling rate thus varies throughout the length of the bar, the rate being highest at the lower end, being in direct contact with the water. The hardness along the length of the bar is then measured at various distances from the quenched end.

As expected from the description of the effects of cooling rates in Section 4.7, hardness decreases away from the quenched end of the bar (Fig. 4.18b). The greater the depth to which hardness increases, the greater is the hardenability of the alloy. Each composition of an alloy has its particular **hardenability band**. Note that the hardness at the quenched end increases with increasing carbon content, and that 1040, 4140, and 4340 steels have the same carbon content (0.40%), and thus they have the same hardness (57 HRC) at the quenched end.

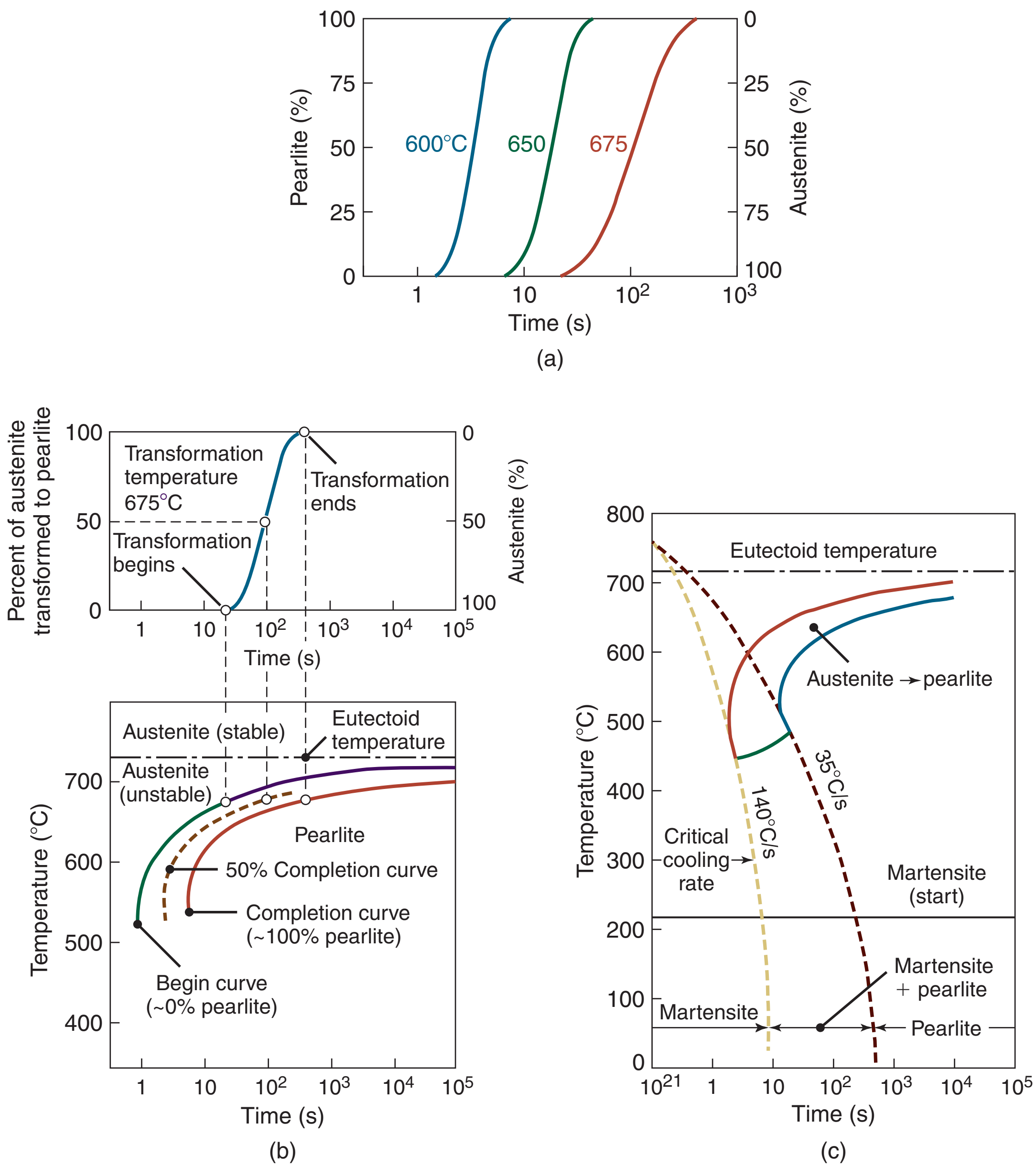


Figure 4.15: (a) Austenite-to-pearlite transformation of iron–carbon alloy as a function of time and temperature. (b) Isothermal transformation diagram obtained from (a) for a transformation temperature of 675°C. (c) Microstructures obtained for a eutectoid iron–carbon alloy as a function of cooling rate.

4.8.2 Quenching Media

The fluid used for quenching the heated specimen also has an effect on hardenability. Quenching may be carried out in water, brine (salt water), oil, molten salt, or air; caustic solutions, polymer solutions, and gases may also be used. Because of the differences in thermal conductivity, specific heat, and heat of vaporization of these media, the rate of cooling of the specimen (**severity of quench**) is also different. In relative terms and in decreasing order, the cooling capacities of various quenching media are: agitated brine, 5; still water, 1; still oil, 0.3; cold gas, 0.1; and still air, 0.02.

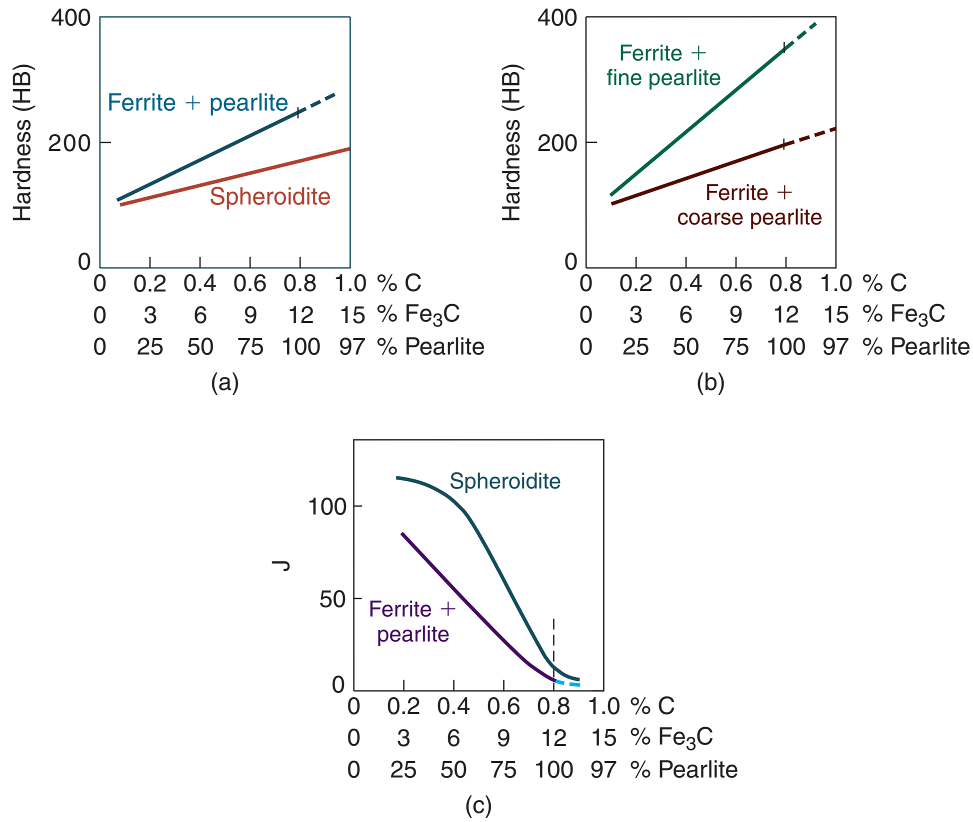


Figure 4.16: (a) and (b) Hardness and (c) toughness for annealed plain-carbon steels as a function of carbide shape. Carbides in the pearlite are lamellar. Fine pearlite is obtained by increasing the cooling rate. The spheroidite structure has sphere-like carbide particles.

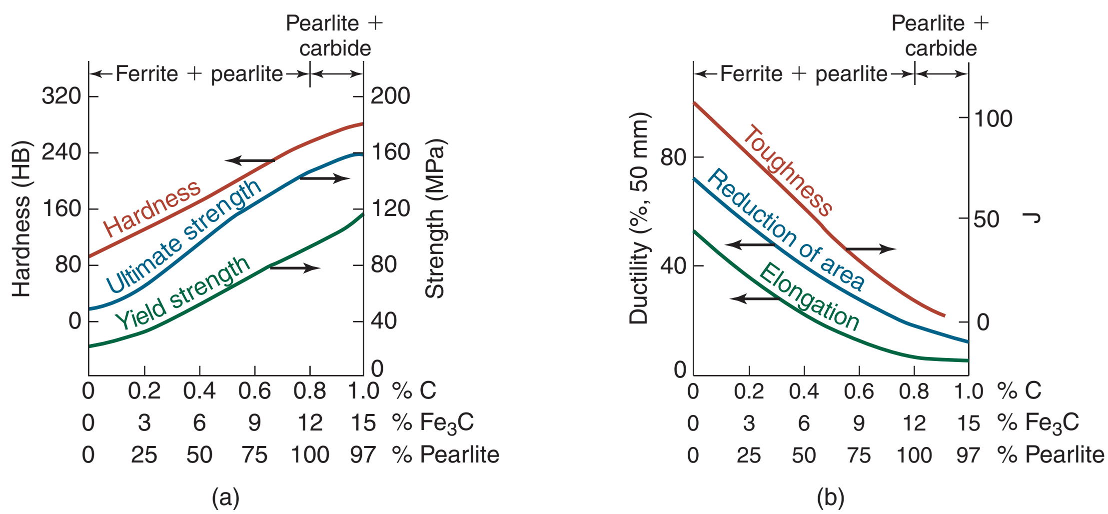


Figure 4.17: Mechanical properties of annealed steels as a function of composition and microstructure. Note in (a) the increase in hardness and strength, and in (b), the decrease in ductility and toughness, with increasing amounts of pearlite and iron carbide.

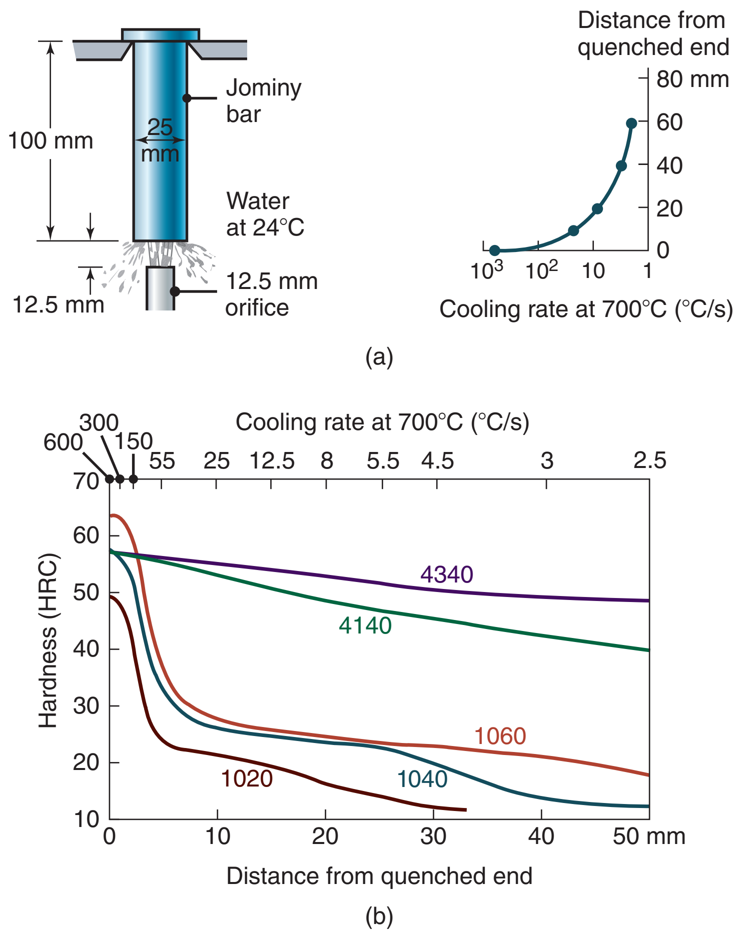


Figure 4.18: (a) End-quench test and cooling rate. (b) Hardenability curves for five different steels, as obtained from the end-quench test. Small variations in composition can change the shape of these curves. Each curve is actually a band, and its exact determination is important in the heat treatment of metals, for better control of properties.

Agitation is a significant factor in the rate of cooling; the more vigorous the agitation, the higher is the rate of cooling. In tool steels, the quenching medium is specified by a letter (see Table 5.7), such as W for water hardening, O for oil hardening, and A for air hardening. The cooling rate also depends on the surface-area-to-thickness or surface-area-to-volume ratio of the part; the higher this ratio, the higher is the cooling rate. For example, a thick plate cools more slowly than a thin plate with the same surface area. These considerations are also significant in the cooling of metals and of plastics in casting and in molding processes (see Sections 10.5.1 and 19.3).

Water is a common medium for rapid cooling; however, the heated specimen may form a **vapor blanket** along its surfaces, due to the water-vapor bubbles that form when water boils at the metal–water interfaces. This blanket creates a barrier to heat conduction, because of the lower thermal conductivity of the vapor. Agitating the fluid or the part helps to reduce or eliminate the blanket; also, water may be sprayed onto the part under the high pressure. Brine is an effective quenching medium, because salt helps to nucleate bubbles at the interfaces, which improves agitation; note, however, that brine can corrode the part.

Polymer quenchants can be used for ferrous as well as for nonferrous alloys, and have cooling characteristics that generally are between those of water and petroleum oils. Typical polymer quenchants are polyvinyl alcohol, polyalkaline oxide, polyvinyl pyrrolidone, and polyethyl oxazoline. These quenchants have such advantages as better control of hardness, elimination of fumes and fire (as may occur when oils are used as a quenchant), and reduction of corrosion (as may occur when water is used).

4.9 Heat Treatment of Nonferrous Alloys and Stainless Steels

Nonferrous alloys and some stainless steels cannot be heat treated by the techniques described for ferrous alloys. The reason is that nonferrous alloys do not undergo phase transformations as do steels. The hardening and strengthening mechanisms for these alloys are therefore fundamentally different. Heat-treatable aluminum alloys, copper alloys, martensitic, and some stainless steels are hardened and strengthened by **precipitation hardening**. In this process, small particles of a different phase, called **precipitates**, are uniformly dispersed in the matrix of the original phase (Fig. 4.3a). Precipitates form because the solid solubility of one element (one component of the alloy) in the other is exceeded.

Three stages are involved in precipitation hardening, which can best be described by reference to the phase diagram for the aluminum–copper system (Fig. 4.19a). For a composition of 95.5% Al–4.5% Cu, a single-phase (kappa phase) substitutional solid solution of copper (solute) in aluminum (solvent) exists between 500° and 570°C. The *kappa phase* is aluminum rich, has an fcc structure, and is ductile. Below the lower temperature (that is, below the lower solubility curve) there are two phases: *kappa* (κ) and *theta* (θ), which is a hard intermetallic compound of $CuAl_2$. This alloy can be heat treated, and its properties modified by two different methods: *solution treatment* and *precipitation hardening*.

4.9.1 Solution Treatment

In **solution treatment**, the alloy is heated to within the solid-solution kappa phase (say, 540°C), and then cooled rapidly such as by quenching it in water. The structure developed soon after quenching

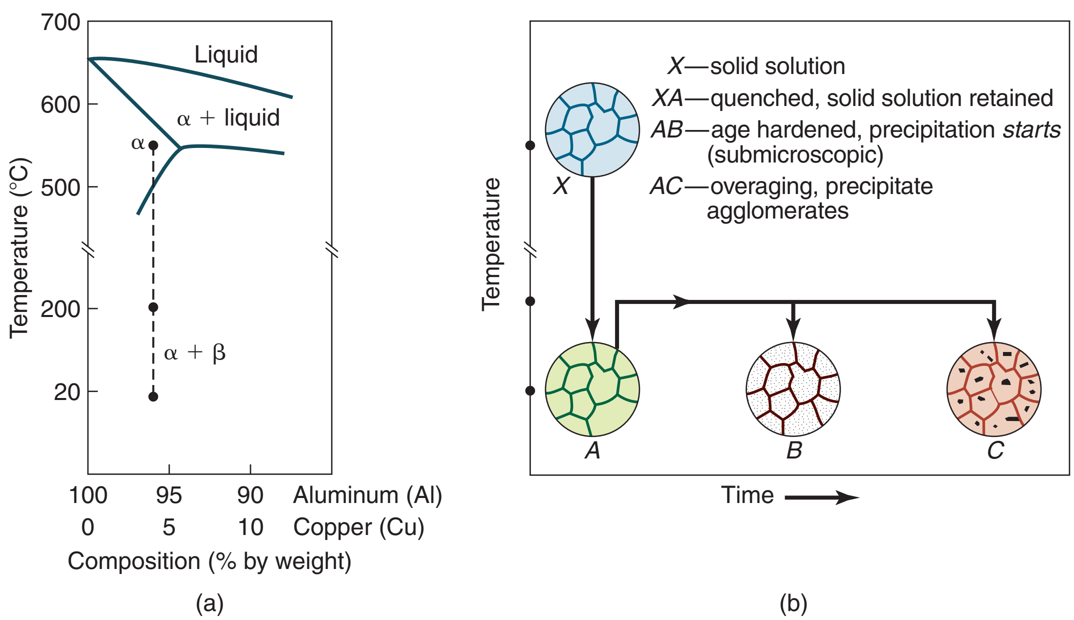


Figure 4.19: (a) Phase diagram for the aluminum–copper alloy system. (b) Various microstructures obtained during the age-hardening process.

(A in Fig. 4.19b) consists only of the single phase kappa; this alloy has moderate strength and considerable ductility.

4.9.2 Precipitation Hardening

The structure developed in A in Fig. 4.19b can be made stronger by **precipitation hardening**. In this process, the alloy is first reheated to an intermediate temperature, and then held for a period of time, during which precipitation takes place; the copper atoms diffuse to nucleation sites and combine with aluminum atoms. This process develops the theta phase, which forms as submicroscopic precipitates (shown in B by the small dots within the grains of the kappa phase). The resulting structure is stronger than that in A, although it is less ductile; the increase in strength is due to the increased resistance to dislocation movement in the region of the precipitates.

Aging. Because the precipitation process is one of time and temperature, it is also called aging, and the property improvement is known as **age hardening**. If carried out above room temperature, the process is called **artificial aging**. However, several aluminum alloys harden and become stronger over a period of time and at room temperature; this process is then called **natural aging**. Such alloys are first quenched and then, if desired, they are shaped by plastic deformation at room temperature; finally, they are allowed to develop strength and hardness by aging naturally. The rate of natural aging can be slowed by refrigerating the quenched alloy (**cryogenic treatment**).

In the precipitation process, if the reheated alloy is held at elevated temperature for an extended period of time, the precipitates begin to coalesce and grow. They become larger, but fewer in number, as shown by the larger dots in C in Fig. 4.19b. This process is called **over-aging**, whereby the alloy becomes softer and less strong.

There is an optimal time–temperature relationship in the aging process that must be observed in order to obtain the desired properties (Fig. 4.20). It is apparent that an aged alloy can be used only up to a certain maximum temperature in service, as otherwise it will over-age and lose some of its strength and hardness, although it will have better dimensional stability.

Maraging. This is a precipitation-hardening treatment for a special group of high strength, iron-base alloys. The term *maraging* is derived from the words *martensite age hardening*, a process in which one or more intermetallic compounds are precipitated in a matrix of low-carbon martensite. A typical maraging steel may contain 18% Ni, in addition to other elements, and aging takes place at 480°C. Because hardening by maraging does not depend on the cooling rate; uniform and full hardness can be developed throughout large parts, and with minimal distortion. Typical uses of maraging steels are in dies and tooling for casting, molding, forging, and extrusion (Parts II and III).

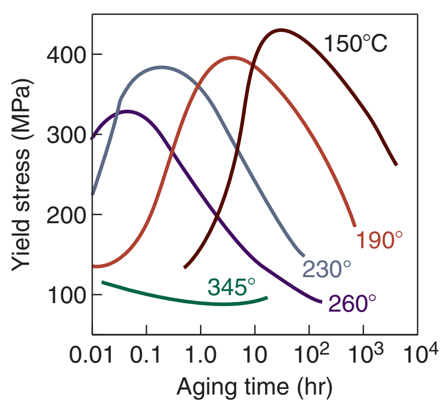


Figure 4.20: The effect of aging time and temperature on the yield stress of 2014-T4 aluminum alloy. Note that, for each temperature, there is an optimal aging time for maximum strength.

4.10 Case Hardening

The heat-treatment processes described thus far involve microstructural alterations and property changes in the bulk of the component by means of *through hardening*. It is not desirable to through harden parts, because a hard part lacks the required toughness for some applications. For example, a small surface crack could propagate rapidly through a part and cause sudden and total failure. In many cases, modification of only the *surface properties* of a part is desirable (hence, the term *surface* or *case hardening*). This widely used method is particularly useful for improving resistance to surface indentation, fatigue, and wear; typical applications are gear teeth, cams, shafts, bearings, fasteners, pins, automotive clutch plates, tools, and dies.

Several case-hardening processes are available (Table 4.1):

1. **Carburizing** (gas, liquid, and pack carburizing)
2. **Carbonitriding**
3. **Cyaniding**
4. **Nitriding**
5. **Boronizing**
6. **Flame hardening**
7. **Induction hardening**
8. **Laser-beam hardening.**

Basically, these are operations where the component is heated in an atmosphere containing such elements as carbon, nitrogen, or boron, which modify the composition, microstructure, and properties of surfaces. For steels with sufficiently high carbon content, surface hardening takes place without the use of any of these additional elements. Only the heat-treatment processes described in Section 4.7 are required to modify the microstructures, usually by either flame hardening or induction hardening, as outlined in Table 4.1.

Laser beams and **electron beams** (Sections 27.6 and 27.7) are used effectively to harden small and as well as large surfaces, such as gears, valves, punches, and engine cylinders. The depth of the case-hardened layer is typically less than 2.5 mm; these methods are also used for through hardening of relatively small parts. The main advantages of laser surface hardening are close control of power input, low part distortion, and the ability to reach areas that would otherwise be inaccessible by other means.

Because case hardening involves a localized surface layer, the parts have a hardness gradient. Typically, the hardness is a maximum at the surface and decreases inward, the rate of decrease depending on the composition and physical properties of the metal and processing variables. Surface-hardening techniques can also be used for *tempering* (Section 4.11), for modifying the properties of surfaces that have been subjected to heat treatment. Several other processes and techniques for surface hardening, such as shot peening and surface rolling to improve wear resistance and other characteristics, are described in Section 34.2.

Decarburization is the phenomenon in which alloys lose carbon from their surfaces, as a result of heat treatment or of hot working in a medium, usually oxygen, which reacts with the carbon. Decarburization is undesirable because it affects the hardenability of surfaces, by lowering their carbon content; it also adversely affects the hardness, strength, and fatigue life of steels, significantly lowering their endurance limit.

4.11 Annealing

Annealing is a general term used to describe the restoration of a cold-worked or heat-treated alloy to its original properties–for instance, to increase ductility (and hence formability) and reduce hardness and

Table 4.1: Outline of Heat-treatment Processes for Surface Hardening

Process	Metals hardened	Element added to surface	Procedure	General characteristics	Typical applications
Carburizing	Low-carbon steel (0.2% C), alloy steels (0.08–0.2% C)	C	Heat steel at 870°–950°C in an atmosphere of carbonaceous gases (gas carburizing) or carbon-containing solids (pack carburizing). Then quench.	A hard, high-carbon surface is produced. Hardness 55 to 65 HRC. Case depth < 0.5 to 1.5 mm. Some distortion of part during heat treatment.	Gears, cams, shafts, bearings, piston pins, sprockets, clutch plates
Carbo-nitriding	Low-carbon steel	C and N	Heat steel at 700°–800°C in an atmosphere of carbonaceous gas and ammonia. Then quench in oil.	Surface hardness 55 to 62 HRC. Case depth 0.07 to 0.5 mm. Less distortion than in carburizing.	Bolts, nuts, gears
Cyaniding	Low-carbon steel (0.2% C), alloy steels (0.08–0.2% C)	C and N	Heat steel at 760°–845°C in a molten bath of solutions of cyanide (e.g., 30% sodium cyanide) and other salts.	Surface hardness up to 65 HRC. Case depth 0.025 to 0.25 mm. Some distortion.	Bolts, nuts, screws, small gears
Nitriding	Steels (1% Al, 1.5% Cr, 0.3% Mo), alloy steels (Cr, Mo), stainless steels, high-speed tool steels	N	Heat steel at 500°–600°C in an atmosphere of ammonia gas or mixtures of molten cyanide salts. No further treatment.	Surface hardness up to 1100 HV. Case depth 0.1 to 0.6 mm and 0.02 to 0.07 mm for high speed steel.	Gears, shafts, sprockets, valves, cutters, boring bars, fuel-injection pump parts
Boronizing	Steels	B	Part is heated using boron-containing gas or solid in contact with part.	Extremely hard and wear resistant surface. Case depth 0.025 to 0.075 mm.	Tool and die steels
Flame hardening	Medium-carbon steels, cast irons	None	Surface is heated with an oxyacetylene torch, then quenched with water spray or other quenching methods.	Surface hardness 50 to 60 HRC. Case depth 0.7 to 6 mm. Little distortion.	Gear and sprocket teeth, axles, crankshafts, piston rods, lathe beds and centers
Induction hardening	Same as above	None	Metal part is placed in copper induction coils and is heated by high frequency current, then quenched.	Same as above	Same as above

strength, or to modify the microstructure of an alloy. The annealing process is also used to relieve residual stresses in a manufactured part, as well as to improve machinability and dimensional stability. The term annealing also applies to the thermal treatment of glass and similar products (Section 18.4), and for castings and weldments.

The annealing process typically consists of the following steps:

1. Heating the workpiece to a specific temperature range in a furnace

2. Holding it at that temperature for a period of time (soaking)

3. Cooling the workpiece, in air or in a furnace

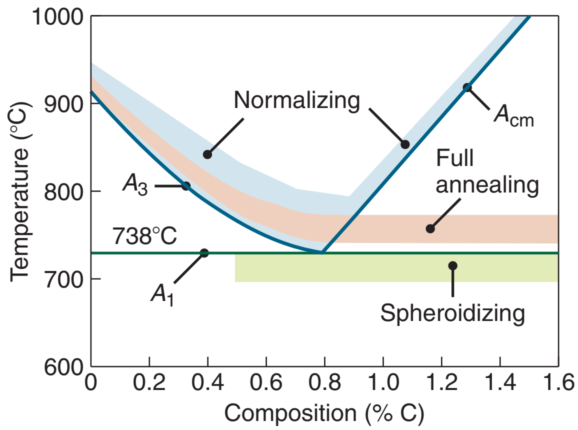


Figure 4.21: Heat-treating temperature ranges for plain-carbon steels, as indicated on the iron–iron carbide phase diagram.

Annealing may be carried out in an inert or a controlled atmosphere, or at lower temperatures to minimize or prevent surface oxidation.

The *annealing temperature* may be higher than the metal's recrystallization temperature, depending on the degree of cold work. For example, the recrystallization temperature for copper ranges between 200° and 300°C, whereas the temperature required to fully recover the original properties ranges from 260° to 650°C, depending on the degree of prior cold work (see also Section 1.7).

Full annealing is a term applied to annealing of ferrous alloys. The steel is heated to above A_1 or A_3 (Fig. 4.21), then cooling takes place slowly [typically at 10°C per hour], in a furnace, after which it is turned off. The structure developed through full annealing is coarse pearlite, which is soft and ductile and has small, uniform grains.

To avoid excessive softness, the cooling cycle may be done completely in still air. In **normalizing**, the part is heated to a temperature above A_3 or $A_{\rm cm}$ in order to transform the structure to austenite. The process results in somewhat higher strength and hardness, and lower ductility than in full annealing (Fig. 4.22). The structure developed is fine pearlite, with small, uniform grains. Normalizing is generally carried out to refine the grain structure, obtain uniform structure (*homogenization*), decrease residual stresses, and improve machinability. The structure of spheroidites and the procedure for obtaining them are described in Section 4.7 and shown in Figs. 4.12 and 4.21. *Spheroidizing annealing* improves the cold workability (Section 14.5) and the machinability of steels (Section 21.7).

Stress-relief Annealing. To reduce or eliminate residual stresses, the part is typically subjected to *stress-relief annealing* or **stress relieving**. The temperature and the time required for this process depend on the material and on the magnitude of the residual stresses present. Residual stresses may have been induced during forming, machining, or other shaping processes, or they may have been caused by volume changes during phase transformations.

Tempering. If steels are hardened by heat treatment, *tempering* or **drawing** is used in order to reduce brittleness, increase ductility and toughness, and reduce residual stresses. The term tempering is also used for glasses (Section 18.4). In tempering, the steel is heated to a specific temperature, depending on its composition, and then cooled at a prescribed rate. The results of tempering for an oil-quenched AISI 4340 steel are shown in Fig. 4.23. Alloy steels may undergo **temper embrittlement**, which is caused by the segregation of impurities along grain boundaries, at temperatures between 480° and 590°C.

Austempering. In this process, the heated steel is rapidly quenched from the austenitizing temperature, to avoid formation of ferrite or pearlite. The part is held at a certain specific temperature until isothermal

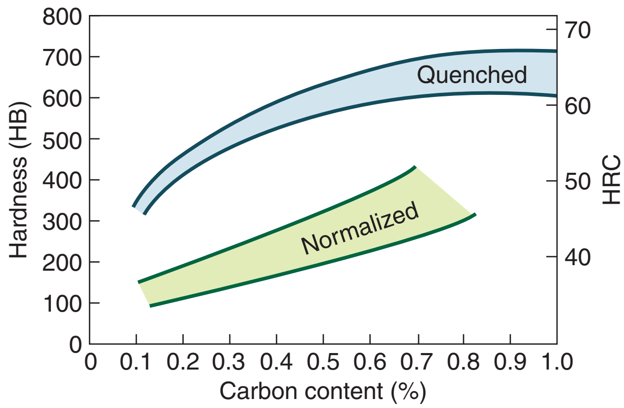


Figure 4.22: Hardness of steels in the quenched and normalized conditions as a function of carbon content.

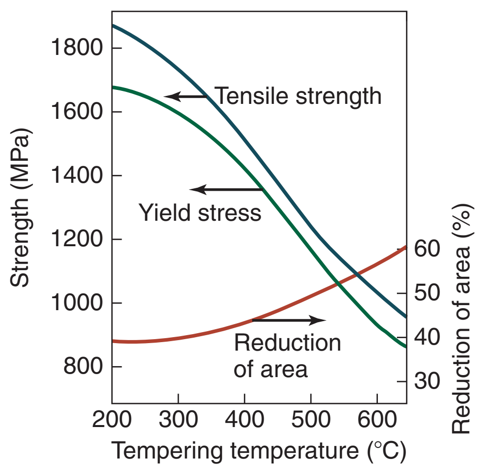


Figure 4.23: Mechanical properties of oil-quenched 4340 steel as a function of tempering temperature.

transformation from austenite to bainite is complete. It is then cooled to room temperature, usually in still air and at a moderate rate, in order to avoid thermal gradients within the part. The quenching medium most commonly used is molten salt, at temperatures ranging from 160° to 750°C.

Austempering is often substituted for conventional quenching and tempering, either to reduce the tendency for cracking and distortion during quenching or to improve ductility and toughness while maintaining hardness. Because of the shorter cycle time involved, this process is economical.

Martempering (Marquenching). In *martempering*, steel or cast iron is first quenched from the austenitizing temperature in a hot-fluid medium, such as hot oil or molten salt. It is then held at that temperature until the temperature is uniform throughout the part. It is cooled at a moderate rate, such as in air in order to avoid excessive temperature gradients within the part. The part is subsequently tempered, because the structure developed is otherwise primarily untempered martensite, thus not suitable for most applications. Martempered steels have lower tendency to crack, distort, or develop residual stresses during heat treatment. In **modified martempering**, the quenching temperature is lower, and the cooling rate is higher; the process is suitable for steels with lower hardenability.

Ausforming. In this process, also called **thermomechanical processing**, the steel is formed into desired shapes within controlled ranges of temperature and time to avoid the formation of nonmartensitic transformation products. The part is then cooled at various rates to develop the desired microstructures.

4.12 Heat Treating Furnaces and Equipment

Two basic types of furnaces are used for heat treating: **batch furnaces** and **continuous furnaces**. Because they consume much energy, their insulation and efficiency are important design considerations, as are their initial cost, the personnel needed for their operation and for maintenance, and their safe use.

Uniform temperature and accurate control of temperature–time cycles are important. Modern furnaces are equipped with various electronic controls, including computer-controlled systems, programmed to run through a complete heat-treating cycle repeatedly and with reproducible accuracy. The fuels used are usually natural gas, oil, or electricity (for resistance or induction heating); the type of fuel affects the furnaces atmosphere: Unlike electric heating, gas or oil introduces combustion products into the furnace (a disadvantage). Electrical heating, however, has a slower start-up time and is more difficult to adjust and control.

Batch Furnaces. In a batch furnace, the parts to be heat treated are loaded into and unloaded from the furnace in individual batches. The furnace basically consists of an insulated chamber, a heating system, and an access door or doors. Batch furnaces are of the following basic types:

1. A **box furnace** is a horizontal rectangular chamber with one or two access doors through which parts are loaded.
2. A **pit furnace** is a vertical pit below ground level into which the parts are lowered.
3. A **bell furnace** is a round or rectangular box furnace without a bottom, and is lowered over stacked parts that are to be heat treated. This type of furnace is particularly suitable for coils of wire, rods, and sheet metal.
4. In an **elevator furnace**, the parts are loaded onto a car platform, rolled into position, and then raised into the furnace.

Continuous Furnaces. In this type of furnace, the parts to be heat treated move continuously through the furnace on conveyors of various designs.

Salt-bath Furnaces. Because of their high heating rates and better control of uniformity of temperature, *salt baths* are commonly used in various heat-treating operations, particularly for nonferrous strip or wire. Heating rates are high because of the higher thermal conductivity of liquid salts compared with that of air or gases.

Fluidized Beds. Dry, fine, and loose solid particles, usually aluminum oxide, are heated and suspended in a chamber by the upward flow of hot gas at various speeds. The parts to be heat treated are then placed within the floating particles, hence the term *fluidized bed*.

Induction Heating. In this method, the part is heated rapidly by the electromagnetic field generated by an *induction coil* carrying alternating current, which induces eddy currents in the part. The coil, which can be shaped to fit the contour of the part to be heat treated (Fig. 4.24), an important consideration, is made of copper or a copper-base alloy. The coil, which is usually water cooled, may be designed to quench the part as well.

Furnace Atmospheres. The atmospheres in furnaces can be controlled so as to avoid oxidation, tarnishing, and decarburization of ferrous alloys heated to elevated temperatures. Oxygen causes corrosion, rusting,

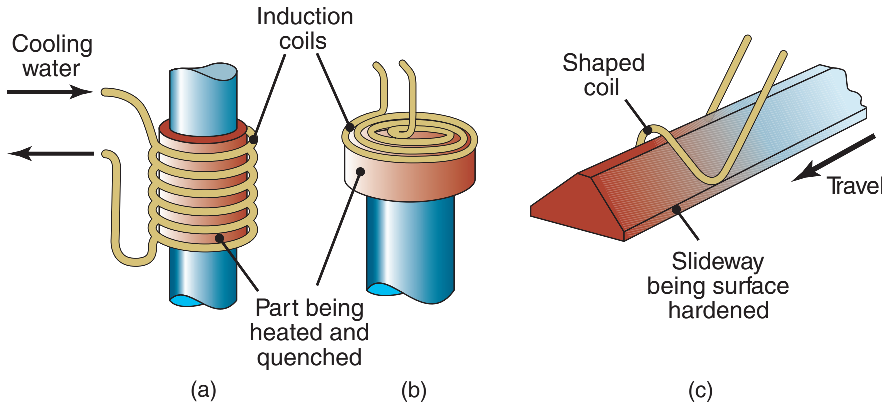


Figure 4.24: Types of coils used in induction heating of various surfaces of parts.

and scaling. Carbon dioxide, which has various effects, may be neutral or decarburizing (see Fig. 30.1), depending on its concentration in the furnace atmosphere. Nitrogen is a common neutral atmosphere, and a vacuum provides a completely neutral atmosphere. Water vapor in the furnace causes oxidation, resulting in a blue color. The term **bluing** is used to describe formation of a thin, blue film of oxide on finished parts, for the purpose of improving their appearance and their resistance to oxidation.

4.13 Design Considerations for Heat Treating

In addition to metallurgical factors, successful heat treating involves several design considerations for avoiding problems such as cracking, distortion, and nonuniformity of properties throughout and among heat-treated parts. The rate of cooling during quenching may not be uniform, particularly in parts of complex shapes with varying cross-sections and thicknesses; the nonuniformity may produce severe temperature gradients in the part. Lack of uniformity can lead to variations in part contraction, resulting in thermal stresses that may cause warping or cracking. Nonuniform cooling also causes residual stresses in the part, which then can lead to stress-corrosion cracking (Section 2.10.2). The quenching method selected, the care taken during the process, and the selection of a proper quenching medium and temperature are important considerations.

As a general guideline for part design for heat treating: (a) Sharp internal or external corners should be avoided, as otherwise stress concentrations at these corners may raise the level of stresses high enough to cause cracking. (b) The part should have its thicknesses as nearly uniform as possible. (c) The transition between regions of different thicknesses should be made smooth. (d) Parts with holes, grooves, keyways, splines, and asymmetrical shapes may be difficult to heat treat, because they may crack during quenching. (e) Large surfaces with thin cross-sections are likely to warp. (f) Hot forgings and hot steel-mill products may have a *decarburized skin* (a layer that has lost its carbon, Section 4.10) that may not respond successfully to heat treatment.

Summary

- Commercially pure metals generally do not have sufficient strength for most engineering applications; consequently, they are alloyed with various elements to alter their structures and properties. Important concepts in alloying are the solubility of alloying elements in a host metal and the phases present at various ranges of temperature and composition.

- Alloys basically have two forms: solid solutions and intermetallic compounds; solid solutions may be substitutional or interstitial. There are certain conditions pertaining to the crystal structure and atomic radii that have to be met to develop these structures.
- Phase diagrams show the relationships among the temperature, composition, and phases present in a particular alloy system. As temperature is decreased at various rates, correspondingly various transformations take place, resulting in microstructures that have widely different characteristics and properties.
- Among binary systems, the most important is the iron–carbon system, which includes a wide range of steels and cast irons. Important components in this system are ferrite, austenite, and cementite. The basic types of cast irons are gray iron, ductile (nodular) iron, white iron, malleable iron, and compacted-graphite iron.
- The mechanisms for hardening and strengthening metal alloys basically involve heating the alloy and subsequently quenching it at varying cooling rates. As a result, important phase transformations take place, producing various structures such as pearlite (fine or coarse), spheroidite, bainite, and martensite. Heat treating of nonferrous alloys and stainless steels involves solution treatment and precipitation hardening.
- Furnace atmosphere, the quenchants used, the control and characteristics of the equipment, and the shape of the parts to be heat treated are important considerations.
- Hardenability is the capability of an alloy to be hardened by heat treatment. The Jominy end-quench test is a method commonly used to determine hardenability bands for alloys.
- Case hardening is an important process for improving the wear and fatigue resistance of parts. Several methods are available, such as carburizing, nitriding, induction hardening, and laser hardening.
- Annealing includes normalizing, process annealing, stress relieving, tempering, austempering, and martempering, each with the purpose of enhancing the ductility and toughness of heat-treated parts.

Key Terms

Age hardening
Aging
Alloy
Annealing
Austenite
Austempering
Bainite
Case hardening
Cast iron
Cementite
Curie temperature
Decarburization
End-quench test
Equilibrium diagram
Eutectic
Eutectoid reaction
Ferrite
Furnaces
Hardenability
Heat treatment
Intermetallic compounds
Iron–Carbon system
Jominy test
Maraging
Martempering
Martensite
Normalizing
Pearlite

Phase diagram
Phase transformations
Precipitation hardening
Pure metals
Retained austenite
Solid solution
Solute
Solution treatment
Solvent
Spheroidite
Stress relieving
Tempered martensite
Tempering
Time–Temperature–transformation diagrams
Two-phase systems

Bibliography

ASM Handbook, Vol. 3: **Alloy Phase Diagrams**, ASM International, 2016.
ASM Handbook, Vol. 4A: **Steel Heat Treating Fundamentals and Processes**, ASM International, 2013.
—, Vol. 4B: **Steel Heat Treating Technologies**, ASM International, 2014.
—, Vol. 4C: **Induction Heating and Heat Treatment**, ASM International, 2014.
—, Vol. 4D: **Heat Treating of Irons and Steels**, ASM International, 2014.
—, Vol. 4E: **Heat Treating of Nonferrous Alloys**, ASM International, 2016.
—, Vol. 9: **Metallography and Microstructures**, ASM International, 2004.
Brandt, D.A., and Warner, J.C., **Metallurgy Fundamentals**, 5th ed., Goodheart-Wilcox, 2009.
Bryson, W.E., **Heat Treatment, Selection, and Application of Tool Steels**, 2nd ed., Hanser Gardner, 2005.
Callister, W.D., Jr., and Rethwisch, D.G., **Materials Science and Engineering**, 9th ed., Wiley, 2013.
Campbell, F,C., (ed.), **Phase Diagrams: Understanding the Basics,** ASM International, 2012.
Hosford, W.F., **Physical Metallurgy**, 2nd ed., Taylor & Francis, 2010.
Krauss, G., **Steels: Processing, Structure, and Performance**, ASM International, 2005.
Shackelford, J. F., **Introduction to Materials Science for Engineers**, 8th ed., Macmillan, 2014.

Review Questions

4.1. Describe the difference between a solute and a solvent.

4.2. What is a solid solution?

4.3. What are the conditions for obtaining (a) substitutional and (b) interstitial solid solutions?

4.4. Describe the difference between a single-phase and a two-phase system.

4.5. What is an induction heater? What kind of part shapes can be heated by induction heating?

4.6. Describe the major features of a phase diagram.

4.7. What do the terms "equilibrium" and "constitutional," as applied to phase diagrams, indicate?

4.8. What is the difference between "eutectic" and "eutectoid"?

4.9. What is tempering? Why is it performed?

4.10. Explain what is meant by "severity of quenching."

4.11. What are precipitates? Why are they significant in precipitation hardening?

4.12. What is the difference between natural and artificial aging?

4.13. Describe the characteristics of ferrite, austenite, and cementite.

4.14. What is the purpose of annealing?

4.15. What is a Time–Temperature–Transformation diagram? How is it used?

Qualitative Problems

4.16. You may have seen some technical literature on products stating that certain parts in those products are "heat treated." Describe briefly your understanding of this term and why the manufacturer includes it.

4.17. Describe the engineering significance of the existence of a eutectic point in phase diagrams.

4.18. What is the difference between hardness and hardenability?

4.19. Referring to Table 4.1, explain why the items listed under typical applications are suitable for surface hardening.

4.20. It generally is not desirable to use steels in their as-quenched condition. Explain why.

4.21. Describe the differences between case hardening and through hardening, insofar as engineering applications of metals are concerned.

4.22. Describe the characteristics of (a) an alloy, (b) pearlite, (c) austenite, (d) martensite, and (e) cementite.

4.23. Explain why carbon, among all elements, is so effective in imparting strength to iron in the form of steel.

4.24. How does the shape of graphite in cast iron affect its properties?

4.25. In Section 4.8.2, several fluids are listed in terms of their cooling capacity in quenching. Which physical properties of these fluids influence their cooling capacity?

4.26. Why is it important to know the characteristics of heat-treating furnaces? Explain.

4.27. Explain why, in the abscissa of Fig. 4.16c, the percentage of pearlite begins to decrease after 0.8% carbon content is reached.

4.28. What is the significance of decarburization? Give some examples.

4.29. Explain your understanding of size distortion and shape distortion in heat-treated parts, and describe their causes.

4.30. Comment on your observations regarding Fig. 4.18b.

4.31. In Fig. 4.1, the hardened surface at the tip of the gear teeth is much higher than at the root. Explain why.

4.32. List the methods by which (a) steel and (b) aluminum can be hardened. Indicate the methods that are common to both materials.

Quantitative Problems

4.33. Design a heat-treating cycle for carbon steel, including temperature and exposure times, to produce (a) pearlite–martensite steels and (b) bainite–martensite steels.

4.34. Using Fig. 4.4, estimate the following quantities for a 75% Cu–25% Ni alloy: (a) the liquidus temperature, (b) the solidus temperature, (c) the percentage of nickel in the liquid at 1150°C, (d) the major phase at 1150°C, and (e) the ratio of solid to liquid at 1150°C.

4.35. Extrapolating the curves in Fig. 4.14, estimate the time that it would take for 1080 steel to soften to 40 HRC at (a) 300°C and (b) 400°C.

4.36. A typical steel for tubing is AISI 1040, and one for music wire is 1085. Considering their applications, explain the reason for the difference in carbon content.

Synthesis, Design, and Projects

4.37. It was stated in this chapter that, in parts design, sharp corners should be avoided in order to reduce the tendency toward cracking during heat treatment. If it is essential for a part to have sharp corners for functional purposes, and it still requires heat treatment, what method would you recommend for manufacturing this part?

4.38. The heat-treatment processes for surface hardening are summarized in Table 4.1. Each of these processes involves different equipment, procedures, and cycle times; as a result, each incurs different costs. Review the available literature, contact various companies, and then make a similar table outlining the costs involved in each process.

4.39. It can be seen that, as a result of heat treatment, parts can undergo size distortion and shape distortion to various degrees. By referring to the Bibliography at the end of this chapter, make a survey of the technical literature, and report quantitative data regarding the distortions of parts having different shapes.

4.40. Figure 4.18b shows hardness distributions in end-quench tests, as measured along the length of the round bar. Make a simple qualitative sketch showing the hardness distribution across the diameter of the bar. Would the shape of the curve depend on the bar's carbon content? Explain.

4.41. Throughout this chapter, you have seen specific examples of the importance and the benefits of heat treating parts or certain regions of parts. Refer to the Bibliography in this chapter, make a survey of the heat-treating literature, and then compile several examples and illustrations of parts that have been heat treated.

4.42. Refer to Fig. 4.24, think of a variety of other part shapes to be heat treated, and design coils that are appropriate for these shapes. Describe how your designs would change if the parts have varying shapes along their length (such as from a square at one end to a round shape at the other end).

4.43. Inspect various parts in your car or home, and identify those that are likely to have been case hardened. Explain your reasons.

4.44. A vendor provides you heat-treated steel parts, but tension test data suggests the strength is not as high as desired. List investigations you can perform to determine if the low strength is caused by improper heat treating.

Chapter 5

Ferrous Metals and Alloys: Production, General Properties, and Applications

- Ferrous metals and alloys are the most widely used structural materials, generally because of their properties and performance, ease of manufacturing, and low cost.
- The chapter opens with a brief outline of iron and steel production from ore, including descriptions of furnaces and foundry operations.
- The casting of ingots is described, followed by continuous casting operations that are now into wide use for producing high-quality metals in large volumes.
- The chapter then discusses in detail the properties and applications of ferrous metals, including cast irons, carbon and alloy steels, high strength steels, stainless steels, and tool and die steels.

5.1 Introduction

By virtue of their relatively low cost and wide range of mechanical, physical, and chemical properties, **ferrous metals and alloys** are among the most useful of all metals. They contain iron as their base metal and are generally classified as *carbon and alloy steels*, *stainless steels*, *tool and die steels*, *cast irons*, and *cast* *steels*. **Steel** refers to a ferrous alloy, which can be as simple as a mixture of iron and carbon, but also often containing a number of alloying elements to impart various properties. Ferrous alloys are produced as

- Sheet steel for automobiles, appliances, and containers
- Plates for boilers, ships, and bridges
- Structural members such as I-beams, bar products, axles, crankshafts, and railroad rails
- Tools, dies, and molds
- Rods and wire for fasteners such as bolts, rivets, nuts, and staples.

Carbon steels are the least expensive of all structural metals. As an example of their widespread use, ferrous metals make up 70–85% by weight of structural members and mechanical components. The average U.S. passenger vehicle (including trucks and sport utility vehicles) contains about 1000 kg of steel, accounting for about 60% of its total weight.

The use of iron and steel as structural materials has been one of the most important technological developments. Primitive ferrous tools, which first appeared about 4000 to 3000 B.C., were made from meteoritic iron, obtained from meteorites that had struck the earth. True ironworking began in Asia Minor in about 1100 B.C. and signaled the advent of the *Iron Age*. Invention of the blast furnace in about 1340 A.D. made possible the production of large quantities of high-quality iron and steel. (See Table I.2.)

5.2 Production of Iron and Steel

5.2.1 Raw Materials

The three basic materials used in iron- and steelmaking are **iron ore**, **limestone**, and **coke**. Although it does not occur in a free state in nature, *iron* is one of the most abundant elements (in the form of various ores) in the world, making up about 5% of the earth's crust. The principal iron ores are *taconite* (a black flintlike rock), *hematite* (an iron-oxide mineral), and *limonite* (an iron oxide containing water). After it is mined, the ore is crushed into fine particles, the impurities are removed (by various means, such as magnetic separation), and the ore is formed into pellets, balls, or briquettes, using water and various binders. Typically, pellets are about 65% pure iron and about 25 mm in diameter. The concentrated iron ore is referred to as *beneficiated* (as are other concentrated ores). Some iron-rich ores are used directly, without pelletizing.

Coke is obtained from special grades of bituminous coal (a soft coal rich in volatile hydrocarbons and tars) that are heated in vertical ovens, to temperatures of up to 1150°C, and then cooled with water in quenching towers. Coke has several functions in steelmaking, including (a) generating the high level of heat required for the chemical reactions in ironmaking to take place and (b) producing carbon monoxide (a reducing gas, meaning that it removes oxygen, thus reducing iron oxide to iron. The chemical by-products of coke are used in the synthesis of plastics and of chemical compounds.

The function of *limestone* (calcium carbonate) is to remove impurities from molten iron. The limestone reacts chemically with impurities, acting like a **flux** (meaning to flow as a fluid) that causes the impurities to melt at a low temperature. The limestone combines with the impurities and forms a **slag** (which, being light, floats over the molten metal, and, subsequently, is removed). *Dolomite* (an ore of calcium magnesium carbonate) also is used as a flux. The slag is later used in making cement, fertilizers, glass, building materials, rock-wool insulation, and road ballast.

5.2.2 Ironmaking

The three raw materials described above are first dumped into the top of a **blast furnace** (Fig. 5.1), an operation called *charging the furnace*. A blast furnace is basically a large steel cylinder lined with refractory (heat-resistant) brick; it has a height of about a 10-story building. The charge mixture is then melted in a

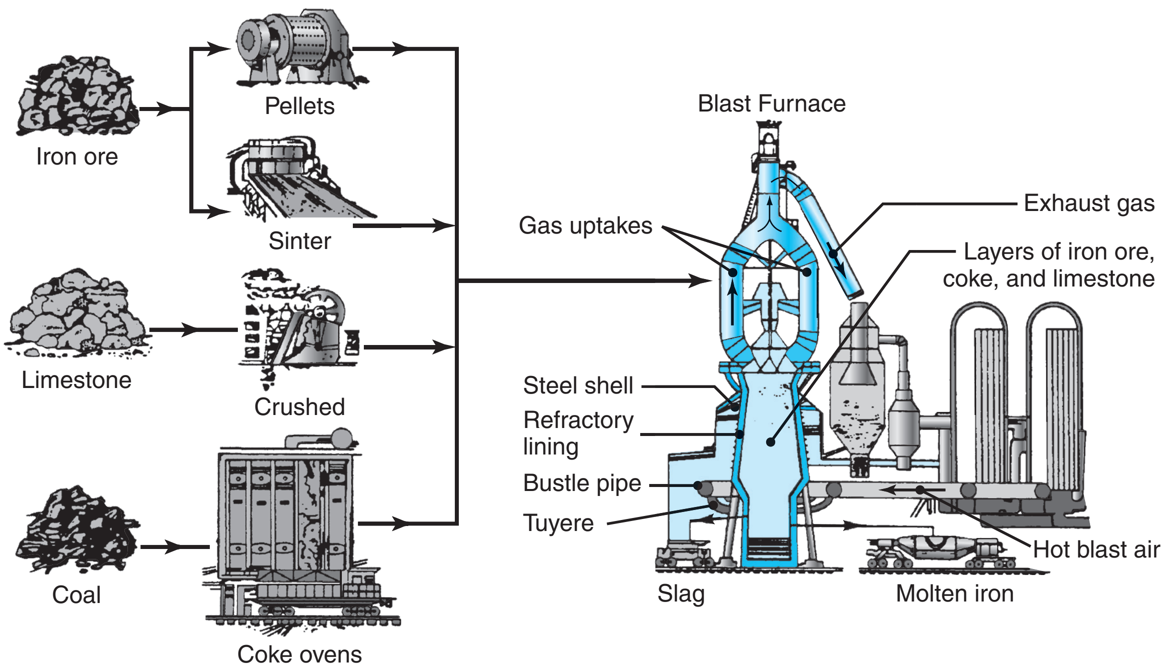


Figure 5.1: Schematic illustration of a blast furnace.

reaction at 1650°C, with the air preheated to about 1100°C and *blasted* into the furnace (hence the term "blast furnace"), through nozzles called *tuyeres*. Although a number of other reactions may take place, the basic reaction is that of oxygen combining with carbon to produce carbon monoxide, which, in turn, reacts with the iron oxide, reducing it to **iron**. Preheating the incoming air is necessary because the burning coke alone does not produce sufficiently high temperatures for these reactions to take place.

The molten metal accumulates at the bottom of the blast furnace, while the impurities float to the top. At intervals of four to five hours, the molten metal is drawn off (*tapped*) into ladle cars, each holding as much as 145 metric tons of molten iron. The molten metal at this stage is called **pig iron**, or simply **hot metal**, and has a typical composition of 4% C, 1.5% Si, 1% Mn, 0.04% S, 0.4% P, the rest being iron. The word **pig** comes from the early practice of pouring the molten iron into small sand molds, arranged around a main channel. These closely packed molds reminded early ironworkers of a litter of small pigs crowding against their mother sow. The solidified metal is later used in making iron and steels.

5.2.3 Steelmaking

Steel was first produced in China and Japan about 600 to 800 A.D. The steelmaking process is essentially one of refining the pig iron by (a) reducing the percentages of manganese, silicon, carbon, and other elements and (b) controlling the composition of the output through the addition of various elements. The molten metal from the blast furnace is then transported into one of four types of furnaces: **open-hearth**, **electric**, **vacuum**, or **basic-oxygen**. The name "open-hearth" is derived from the shallow hearth shape open directly to the flames that melt the metal. Developed in the 1860s, the open-hearth furnace has now been replaced by electric furnaces and by the basic-oxygen process, because they are more efficient and produce steels of better quality.

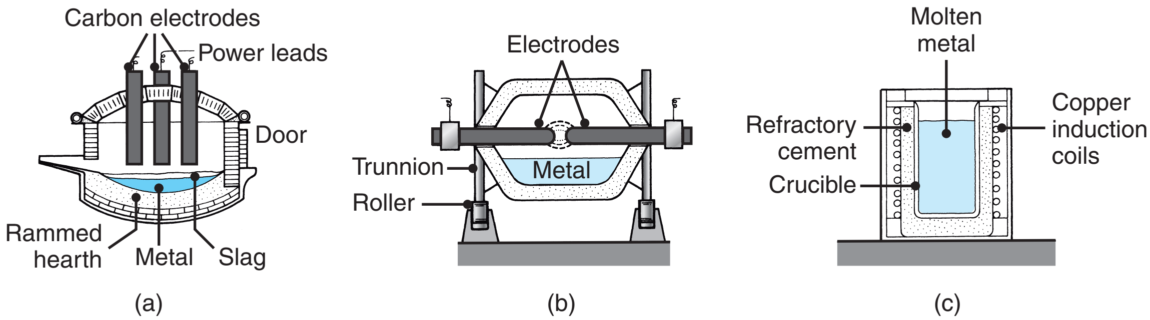


Figure 5.2: Schematic illustration of types of electric furnaces: (a) direct arc, (b) indirect arc, and (c) induction.

Electric Furnace. The source of heat in this type of furnace is a continuous electric arc formed between the electrodes and the charged metal (Fig. 5.2a and b). Temperatures as high as 1925°C are generated in this type of furnace. There are usually three graphite electrodes that can be as large as 750 mm in diameter and 1.5 to 2.5 m long. Their height in the furnace can be adjusted to the amount of metal present and of wear of the electrodes.

Steel scrap and small amounts of carbon and limestone are first dropped into the electric furnace through the open roof. The roof is then closed and the electrodes are lowered. The power is turned on, and within about two hours, the temperature increases sufficiently to melt the metal. The current is then shut off, the electrodes are raised, the furnace is tilted, and the molten metal is poured into a *ladle* (a receptacle used for transferring and pouring molten metal). Furnace capacities range from 55 to 80 metric tons of steel per day; the quality of steel produced is better than that from either open-hearth or basic-oxygen process.

For smaller quantities, electric furnaces can be of the **induction** type (Fig. 5.2c). The metal is placed in a **crucible** – a large pot made of refractory material and surrounded with a copper coil through which alternating current is passed. The induced current in the charge generates heat and melts the metal.

Basic-oxygen Furnace. The basic-oxygen furnace (BOF) is the fastest and by far the most common steelmaking furnace. Typically, 180 metric tons of molten pig iron and 80 metric tons of scrap are charged into a vessel (Fig. 5.3a); some units can hold as much as 360 metric tons. Pure oxygen is then blown into the furnace, for about 20 minutes, through a water-cooled *lance* (a long tube) and under a pressure of about 1250 kPa, as shown in Fig. 5.3b. Fluxing agents (such as calcium or magnesium oxide) are added through a chute. The process is known as *basic* because of the pH of these fluxing agents.

The vigorous agitation of the oxygen refines the molten metal by an oxidation process, in which iron oxide is produced. The oxide reacts with the carbon in the molten metal, producing carbon monoxide and carbon dioxide. The lance is then retracted, and the furnace is tapped by tilting it (note the opening in Fig. 5.3c for the molten metal). The slag is removed by tilting the furnace in the opposite direction. The BOF process is capable of refining 225 metric tons of steel in 35 to 50 minutes. Most BOF steels have low impurity levels and are of better quality than open-hearth furnace steels. They are then processed into plates, sheets, and various structural shapes, such as I-beams and channels (see Fig. 13.1).

Vacuum Furnace. Steel may also be melted in induction furnaces from which the air has been removed (hence the vacuum), similar to the one shown in Fig. 5.2c. Cooling is accomplished by injecting an inert gas, typically argon, at high pressure into the furnace. Because the process removes gaseous impurities from the molten metal and prevents oxidation, vacuum furnaces produce high-quality steels. Vacuum furnaces are also commonly used for heat treating (Section 4.7) and brazing (Section 32.2).

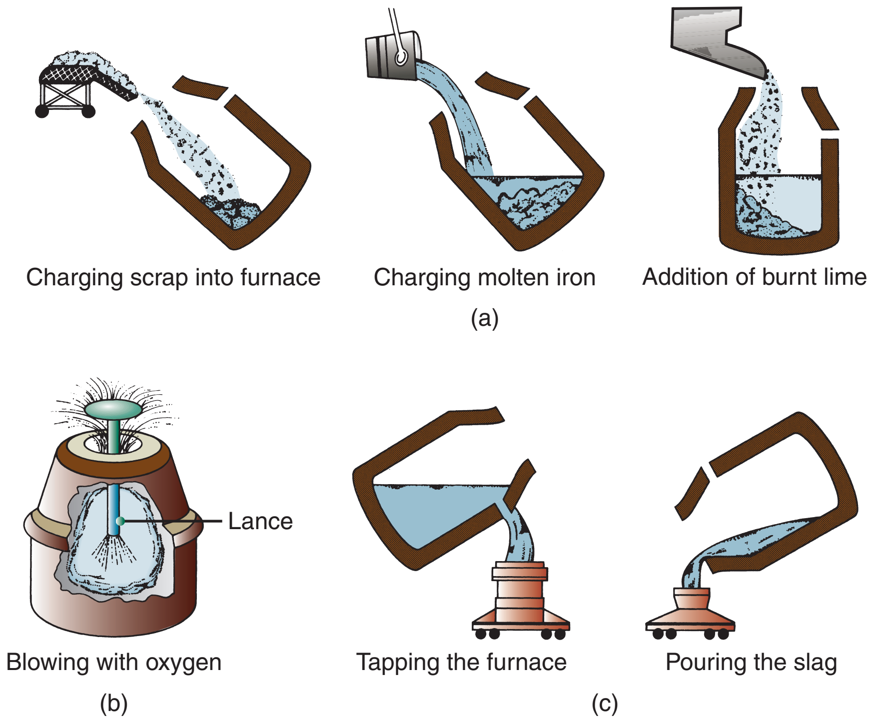


Figure 5.3: Schematic illustrations showing charging, melting, and pouring of molten iron in a basic-oxygen process.

5.3 Casting of Ingots

Traditionally, the steelmaking process involves shaping of the molten steel into a solid form (*ingot*) for further processing, such as rolling it into shapes, casting into semifinished forms, or forging. The molten metal is poured (*teemed*) from the ladle into ingot molds, in which the metal solidifies. Molds usually are made of cupola iron or blast-furnace iron with 3.5% carbon. They are tapered in order to facilitate the removal of the solidified metal from the mold. The bottoms of the molds may be closed or open; if they are open, they are placed on a flat surface. The cooled ingots are stripped from the molds and then lowered into **soaking pits**, where they are reheated to a uniform temperature of about 1200°C for subsequent processing.

Certain important reactions take place during the solidification of an ingot that influence the quality of the steel produced. For example, significant amounts of oxygen and other gases can dissolve in the molten metal during steelmaking. Most of these gases are rejected during the solidification of the metal, because the solubility limit of the gases in the metal decreases sharply as its temperature decreases (see Fig. 10.17). Rejected oxygen combines with carbon and forms carbon monoxide, which causes porosity in the solidified ingot.

Depending on the amount of gas evolved during solidification, three types of steel ingots can be produced:

1. **Killed Steel.** The term *killed* comes from the fact that the steel lies quietly after being poured into the mold. Killed steel is fully deoxidized; that is, oxygen is removed and the associated porosity is thus eliminated. In the deoxidation process, the oxygen dissolved in the molten metal is reacts with elements such as aluminum, silicon, manganese, and vanadium that have been added to the melt.

These elements have an affinity for oxygen, forming metallic oxides. If aluminum is used, the product is called *aluminum-killed steel*.

If they are sufficiently large, the oxide inclusions in the molten bath float out and adhere to, or are dissolved in, the slag. A fully killed steel is thus free of any porosity caused by gases; it also is free of any **blowholes** (large spherical holes near the surfaces of the ingot). Consequently, the chemical and mechanical properties of a killed-steel ingot are relatively uniform throughout. However, because of shrinkage during solidification, an ingot of this type develops a **pipe** at the top (also called a **shrinkage cavity**), with the appearance of a funnel-like shape.

2. **Semi-killed Steel.** Semi-killed steel is a *partially deoxidized steel*. It contains some porosity (generally in the upper central section of the ingot), but it has little or no pipe. Although the piping is less, this advantage is offset by the presence of porosity in the upper region. Semi-killed steels are economical to produce.

3. **Rimmed Steel.** In rimmed steel, which generally has a carbon content of less than 0.15%, the evolved gases are only partially killed (controlled) by adding other elements, such as aluminum. The gases produce blowholes along the outer rim of the ingot—hence the term *rimmed*. Rimmed steels have little or no piping and they have a ductile skin, with good surface finish; however, if not controlled properly, the blowholes may break through the skin. Also, impurities and inclusions tend to segregate toward the center of the ingot.

Refining. The properties and manufacturing characteristics of ferrous alloys are affected adversely by the amount of impurities, inclusions, and other elements present (see Section 2.10). The removal of impurities is known as *refining*; most refining is done in melting furnaces or in ladles, by the addition of various elements.

Refining is particularly important in producing high-grade steels and alloys for high-performance and critical applications, such as aircraft components, automobile structural elements, medical devices, and cutlery. Moreover, warranty periods on shafts, camshafts, crankshafts, and similar parts can be increased significantly by using higher quality steels. Such steels are then subjected to **secondary refining** in ladles (**ladle metallurgy**) and ladle refining (**injection refining**), which generally consists of melting and processing the steel in a vacuum. The examples are electron-beam melting, vacuum-arc remelting, argon-oxygen decarburization, and vacuum-arc double-electrode remelting.

5.4 Continuous Casting

Conceived in the 1860s, **continuous** or **strand casting** was first developed for casting nonferrous metal strips. The process is now used widely for steel, aluminum, and copper production, with major productivity improvements and cost reductions. One system for continuous casting is shown schematically in Fig. 5.4a. The molten metal in the ladle is equalized in temperature, by blowing nitrogen gas through it for 5 to 10 minutes. It is then poured into a refractory-lined intermediate pouring vessel (**tundish**), where impurities are skimmed off; the tundish holds as much as 2.7 metric tons of metal. The molten metal is then tapped from the tundish, travels downward through water-cooled copper molds, and begins to solidify, and is drawn through the molds at a constant velocity by rollers (called *pinch rolls*).

Prior to starting the casting operation, a solid *starter bar* (*dummy bar*) is inserted into the bottom of the mold. As the molten metal is first poured, it solidifies onto the dummy bar. The bar is withdrawn at the same rate at which the metal is poured. The cooling rate is such that the metal develops a solidified skin (*shell*), so as to support itself during its travel downward, typically at speeds of about 25 mm/s. The shell thickness at the exit of the mold is about 12 to 18 mm. Additional cooling is provided by water sprays along the travel path of the solidifying metal. The molds are typically coated with graphite

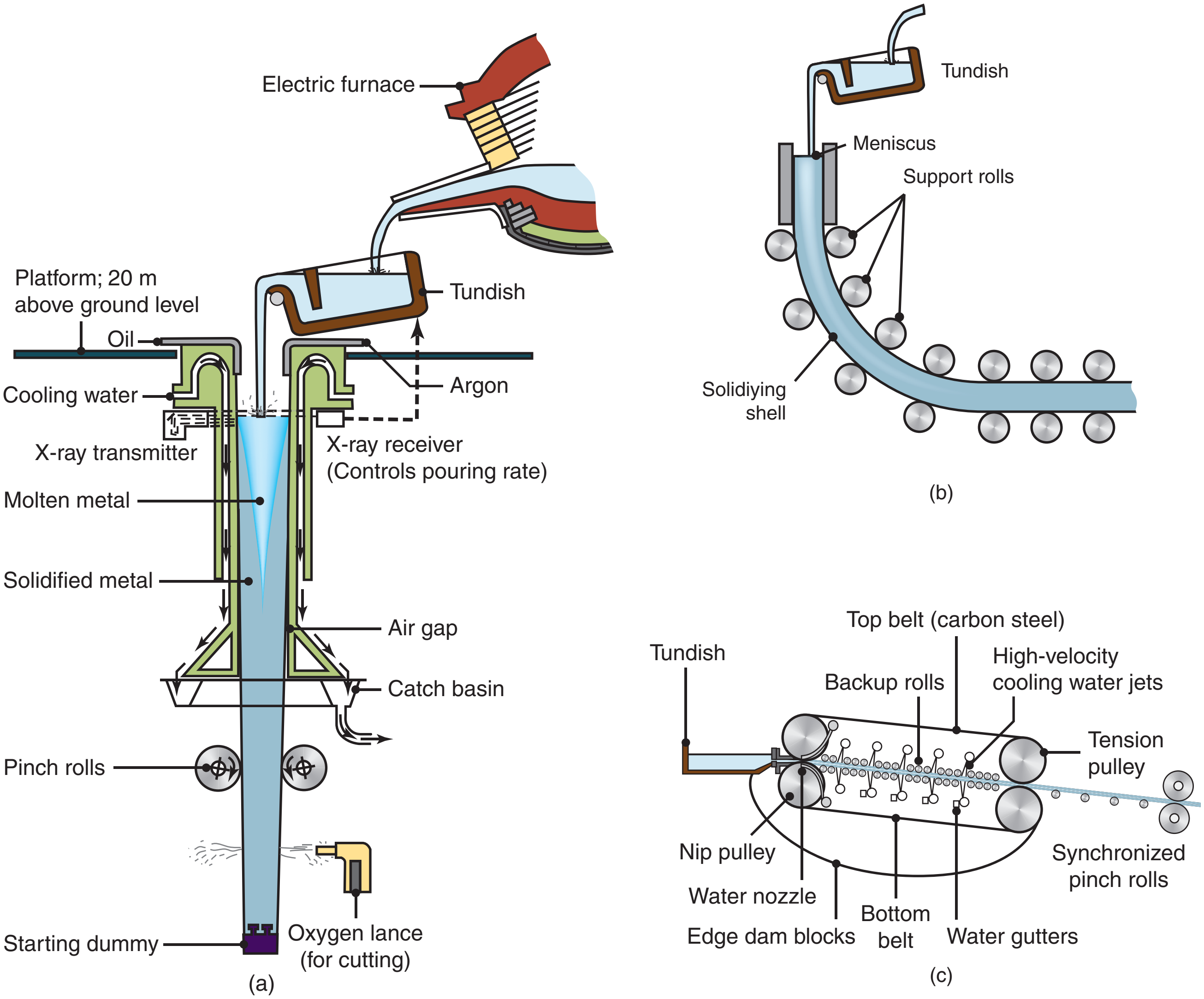


Figure 5.4: (a) The continuous-casting process for steel. Typically, the solidified metal descends at a speed of 25 mm/s. Note that the platform is about 20 m above ground level. *Source: Metalcaster's Reference and Guide*, American Foundrymen's Society. (b) Continuous casting using support or guide rollers to allow transition from a vertical pour zone to horizontal conveyors. (c) Continuous strip casting of nonferrous metal strip. *Source:* (a) Courtesy of American Foundry Society and (b) Courtesy of Hazelett.

or similar solid lubricants, in order to reduce both friction and adhesion at the mold–metal interfaces. The molds may also be vibrated to reduce friction and sticking.

The continuously cast metal may then be cut into specific lengths by shearing or computer-controlled torch cutting, or it may be fed directly into a rolling mill for further reduction in thickness and for the shaping into specific shapes, such as channels and I-beams. In addition to lower cost, continuously cast metals have more uniform compositions and properties than those obtained by traditional ingot casting.

Modern facilities are computer-controlled, on continuously cast strands, with final sheet thicknesses on the order of 2–6 mm by 750–1250 mm wide, for carbon and stainless steels. They have capabilities for a rapid switchover from one type of steel to another. Steel plates or other shapes undergo one or more further processing, such as (a) cleaning and pickling by chemicals to remove surface oxides, (b) cold rolling to improve strength and surface finish, (c) annealing, and (d) galvanizing or aluminizing to improve resistance to corrosion.

In **strip casting**, thin slabs or strips are produced directly from molten metal. The metal solidifies in similar manner to strand casting, but the hot solid is then rolled to form the final shape (Fig. 5.4b). The compressive stresses in rolling (see Section 13.2) serve to reduce porosity and to provide better material properties. Thus, strip casting eliminates a hot-rolling operation in the production of metal strips or slabs. In modern facilities, final thicknesses on the order of 2 to 6 mm can be obtained, for carbon, stainless, and electrical steels as well as other metals.

5.5 Carbon and Alloy Steels

Steel is an alloy that consists primarily of iron with a carbon content between 0.2 and 2.1% by weight. Alloys with higher than 2.1% carbon are known as *cast irons* (described in Section 12.3.2), and have a lower melting point than other steels and good castability. Carbon and alloy steels are among the most commonly used metals and have a wide variety of compositions, processing options, and applications (Table 5.1). They are available in variety of basic product shapes: plate, sheet, strip, bar, wire, tube, castings, and forgings.

Table 5.1: Applications for Selected Carbon and Alloy Steels.

Product	Steel
Aircraft forgings, tubing, fittings	4140, 8740
Automobile bodies	1010
Axles	1040, 4140
Ball bearings and races	52100
Bolts	1035, 4042, 4815
Camshafts	1020, 1040
Chains (transmission)	3135, 3140
Coil springs	4063
Connecting rods	1040, 3141, 4340
Crankshafts (forged)	1045, 1145, 3135, 3140
Differential gears	4023
Gears (car and truck)	4027, 4032
Landing gear	4140, 4340, 8740
Lock washers	1060
Nuts	3130
Railroad rails and wheels	1080
Springs (coil)	1095, 4063, 6150
Springs (leaf)	1085, 4063, 9260, 6150
Tubing	1040
Wire	1045, 1055
Wire (music)	1085

5.5.1 Effects of Various Elements in Steels

Various elements are added to steels in order to impart specific properties, such as hardenability, strength, hardness, toughness, wear resistance, workability, weldability, and machinability. These elements are listed in Table 5.2, with summaries of their beneficial and detrimental effects. Generally, the higher the percentages of these elements, the greater are the particular properties that they impart. For example, the higher the carbon content, the greater the hardenability of the steel and the greater its strength, hardness, and wear resistance. On the other hand, ductility, weldability, and toughness are reduced with increasing carbon content.

Some *residual elements*, called **trace elements**, may remain after production, refining, and processing of steels. Although the elements in Table 5.2 may also be considered as residuals, the following generally are considered unwanted residual elements:

Table 5.2: Effect of Various Elements in Steels.

Element	Effect
Aluminum	Deoxidizes nitriding steels, limits austenite grain growth, increases hardness of nitriding steels
Bismuth	Improves machinability
Boron	Improves hardness without loss of (and perhaps some improvement in) machinability and formability.
Calcium	Deoxidizes steel; improves toughness; may improve formability and machinability.
Carbon	Improves hardenability, strength, hardness, and wear resistance; reduces ductility, weldability and toughness.
Cerium, magnesium, zirconium	Deoxidizes steel, improves toughness in HSLA steels; controls shape of inclusions.
Chromium	Improves toughness, hardenability, wear and corrosion resistance, and high-temperature strength; promotes carburization and depth of hardening in heat treatment.
Cobalt	Improves strength and hardness at elevated temperatures.
Copper	Improves resistance to atmospheric corrosion; can increase strength without loss in ductility; adversely affects hot workability and surface quality.
Lead	Improves machinability; can cause liquid metal embrittlement.
Manganese	Deoxidizes steel, improves hardenability, strength, abrasion resistance, and machinability; reduces hot shortness, and decreases weldability.
Molybdenum	Improves hardenability, wear resistance, toughness, elevated-temperature strength, creep resistance, and hardness; it minimizes temper embrittlement.
Nickel	Improves strength, toughness, corrosion resistance and hardenability.
Niobium, tantalum	Improves strength and impact toughness; it lowers transition temperature and may decrease hardenability.
Phosphorus	Improves strength, hardenability, corrosion resistance, and machinability; it severely reduces ductility and toughness.
Selenium	Improves machinability.
Silicon	Improves strength, hardness, corrosion resistance, and electrical conductivity; decreases machinability and cold formability.
Sulfur	Improves machinability when combined with manganese; decreases impact strength, ductility and weldability.
Tellurium	Improves machinability, formability and toughness.
Titanium	Deoxidizes steel; improves hardenability.
Tungsten	Improves hardness, especially at elevated temperature.
Vanadium	Improves strength, toughness, abrasion resistance, and hardness at elevated temperatures; it inhibits grain growth during heat treatment.

Antimony and **arsenic** cause temper embrittlement.

Hydrogen severely embrittles steels; however, heating during processing drives out most of the hydrogen.

Nitrogen improves strength, hardness, and machinability; in aluminum-deoxidized steels, it controls the size of inclusions. Nitrogen can increase or decrease strength, ductility, and toughness, depending on the presence of other elements.

Oxygen slightly increases the strength of rimmed steels; it severely reduces toughness.

Tin causes hot shortness and temper embrittlement.

5.5.2 Designations for Steels

Traditionally, the American Iron and Steel Institute (AISI) and the Society of Automotive Engineers (SAE) have designated carbon and alloy steels by four digits. The first two digits indicate the alloying elements and their percentages, and the last two digits indicate the carbon content by weight.

The American Society for Testing and Materials (ASTM) has a designation system that incorporates the AISI and SAE designations and includes standard specifications for steel products. For ferrous metals, the designation consists of the letter A followed by numbers (generally three). The current standard numbering system is known as the *Unified Numbering System* (UNS) and has been widely adopted by the ferrous and nonferrous industries. It consists of a letter, indicating the general class of the alloy, followed by five digits, designating its chemical composition. Typical letter designations are:

G–AISI and SAE carbon and alloy steels

J–cast steels

K–miscellaneous steels and ferrous alloys

S–stainless steels and superalloys

T–tool steels

Two examples are: G41300 for AISI 4130 alloy steel, and T30108 for AISI A-8 tool steel.

5.5.3 Carbon Steels

Carbon steels generally are classified by their proportion, by weight, of carbon content. The general mechanical properties of carbon and alloy steels are given in Table 5.3, and the effect of carbon on the properties of steel is shown in Fig. 5.5 and summarized as:

- **Low-carbon steel**, also called **mild steel**, has less than 0.30% C. It often is used for common industrial products (such as bolts, nuts, sheets, plates, and tubes) and for machine components that do not require high strength.

- **Medium-carbon steel** has 0.30–0.60% C. It generally is used in applications requiring higher strength than is available in low-carbon steels, such as in machinery, automotive and agricultural parts (gears, axles, connecting rods, and crankshafts), railroad equipment, and parts for metalworking machinery.

- **High-carbon steel** has more than 0.60% C. Generally, high-carbon steel is used for applications requiring strength, hardness, and wear resistance, such as cutting tools, cable, music wire, springs, and cutlery. After being manufactured into shapes, the parts usually are heat treated and tempered (Chapter 4). The higher the carbon content of the steel, the higher is its hardness, strength, and wear resistance after heat treatment.

Table 5.3: Typical Mechanical Properties of Selected Carbon and Alloy Steels.

AISI	Condition	Ultimate tensile strength (MPa)	Yield strength (MPa)	Elongation in 50 mm (%)	Reduction of area (%)	Typical hardness (HB)
1020	As-rolled	448	346	36	59	143
	Normalized	441	330	35	67	131
	Annealed	393	294	36	66	111
1080	As-rolled	1010	586	12	17	293
	Normalized	965	524	11	20	293
	Annealed	615	375	24	45	174
3140	Normalized	891	599	19	57	262
	Annealed	689	422	24	50	197
4340	Normalized	1279	861	12	36	363
	Annealed	744	472	22	49	217
8620	Normalized	632	385	26	59	183
	Annealed	536	357	31	62	149

- Carbon steels containing sulfur and phosphorus are known as **resulfurized** carbon steels (11xx series) and **rephosphorized and resulfurized** carbon steels (12xx series). For example, 1112 steel is a resulfurized steel with a carbon content of 0.12%. These steels have improved machinability, as described in Section 21.7.

5.5.4 Alloy Steels

Steels containing significant amounts of alloying elements are called **alloy steels**. **Structural-grade alloy steels** are used mainly in the construction and transportation industries, because of their high strength. Other types of alloy steels are used in applications where strength, hardness, creep and fatigue resistance, and toughness are required. They can be heat treated to obtain the specific desired properties.

5.5.5 High-strength Low-alloy Steels

In order to improve the strength-to-weight ratio of steels, several **high-strength, low-alloy steels** (HSLA) have been developed. These steels have low carbon content (usually less than 0.30%) and are characterized by a microstructure consisting of fine-grain ferrite as one phase and a hard second phase of martensite and austenite. The mechanical properties for selected HSLA steels are given in Table 5.4. These steels have high strength and energy-absorption capabilities as compared to conventional steels. The ductility, formability, and weldability of HSLA steels are, however, generally inferior to those of conventional low-alloy steels (see Fig. 5.6). To improve these properties, several ultra-high-strength steels have been developed, as described in Section 5.5.6.

Sheet products of HSLA steels are used typically for parts of truck bodies and other transportation equipment (in order to reduce weight and hence fuel consumption) and in mining, agricultural, and various other industrial applications. Plates are used in ships, bridges, building construction, and for shapes such as I-beams, channels, and angles used in buildings and in various structures.

Designations. Three steel categories comprise the system of AISI designations for high-strength sheet steel (Table 5.5). *Structural quality* (S) includes C, Mn, P, and N. *Low alloys* (X) contain Nb, Cr, Cu, Mo, Ni, Si, Ti, V, and Zr, either singly or in combination. *Weathering steels* (W) have environmental-corrosion resistance approximately four times higher than that of conventional low-carbon steels and contain Si, P, Cu, Ni, and

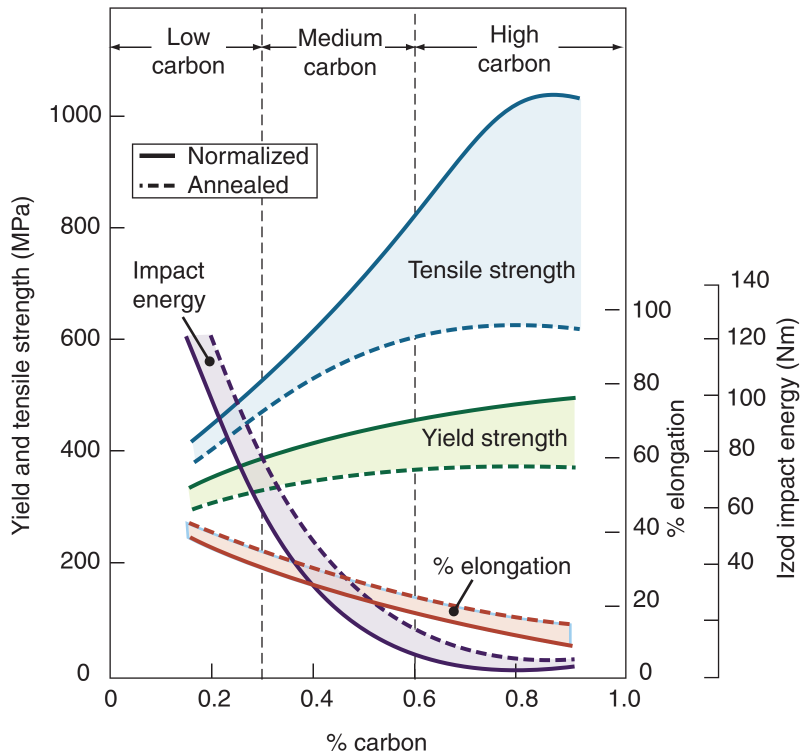


Figure 5.5: Effect of carbon content on the mechanical properties of carbon steel.

Cr in various combinations. The formability of these sheet steels is graded by the letters F (excellent), K (good), and O (fair).

Another designation scheme in wide use is that defined by the Ultralight Steel Auto Body Consortium (ULSAB). The ULSAB practice is to define both the type of steel and its yield and tensile strengths in a compact designation, in the form XX aaa/bbb, where XX is the type of steel, aaa is the yield strength in MPa, and bbb is the ultimate tensile strength in MPa. These types of steel are

BH–Bake-hardenable

HSLA–High-strength low-alloy

DP–Dual-phase

TRIP–Transformation-induced plasticity

TWIP–Twinning-induced plasticity

MART–Martensitic

CP–Complex phase.

Thus, HSLA 350/450 would be a high-strength low-alloy steel with a minimum yield strength of 350 MPa and a minimum ultimate tensile strength of 450 MPa.

Microalloyed Steels. These steels provide superior properties and can eliminate the need for heat treatment. They have a ferrite–pearlite microstructure, with fine dispersed particles of carbonitride. When

Table 5.4: Mechanical Properties of Selected Advanced High-strength Steels.

Steel	Minimum ultimate strength (MPa)	Minimum yield strength (MPa)	Elongation in 50 mm (%)	Strain-hardening exponent, *n*
BH 210/340	340	210	36	0.18
BH 260/370	370	260	32	0.13
HSLA 350/450	450	350	25	0.14
DP 350/600	600	350	27	0.14
DP 500/800	800	500	17	0.14
DP 700/1000	1000	700	15	0.13
DP 1180	1180	1000	5	—
TRIP 450/800	800	450	29	0.24
TRIP 400/600	600	400	30	0.23
CP 700/800	800	700	12	0.13
CP 1000	950	875	10	—
MART 950/1200	1200	950	6	0.07
MART 1250/1520	1520	1250	5	0.065
22MnB5, hot stamped	1500	1100	3	—
27MnCrB5, as rolled	967	478	12	0.06
27MnCrB5, hot stamped	1350	1097	5	0.06
37MnB4, as rolled	810	580	12	0.06
37MnB4, hot stamped	2040	1378	4	0.06

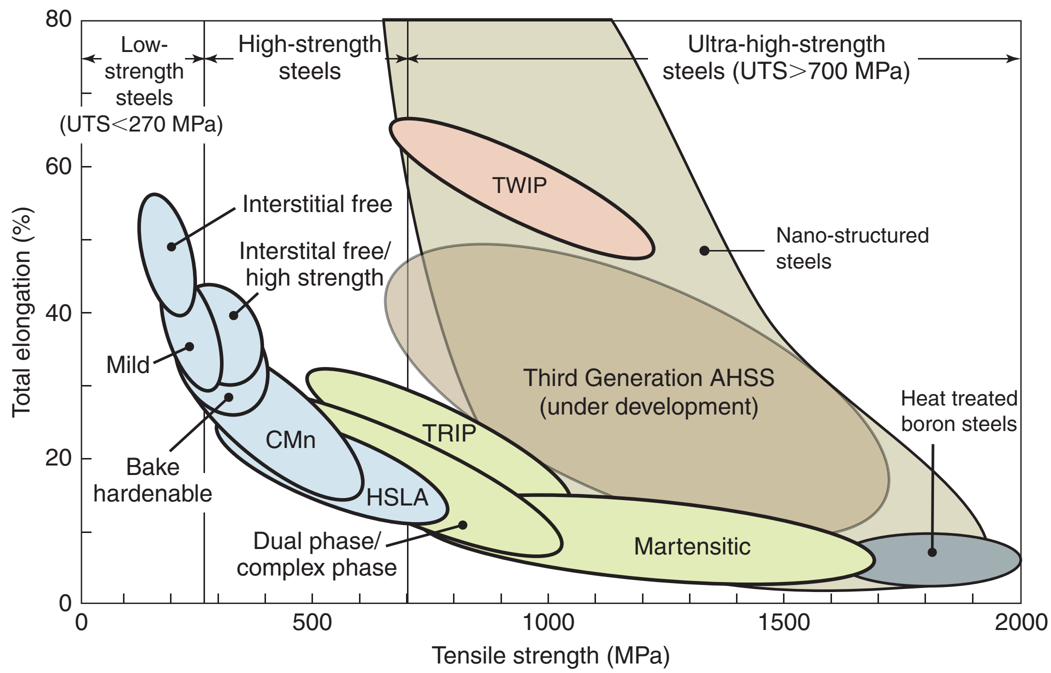


Figure 5.6: Comparison of advanced high-strength steels.

Table 5.5: AISI Designations for High-strength Sheet Steel.

Yield strength MPa	Chemical composition	Deoxidation practice
240	S = structural quality	F = killed plus sulfide inclusion control
275		
310		
350	X = low alloy	
415		K = killed
485	W = weathering	
550		O = nonkilled
690	D = dual-phase	
830		
970		
Example: 350XF		
350	X	F
350 MPa yield strength	Low alloy	Killed plus sulfide inclusion control.

subjected to controlled cooling (usually in air), these steels develop improved and uniform strength. Compared to medium-carbon steels, microalloyed steels also can provide cost savings of as much as 10%, since the steps of quenching, tempering, and stress relieving are not required.

Nanoalloyed Steels. These steels have extremely small grain sizes (10–100 nm), and are produced using metallic glasses (Section 6.15) as a precursor or with special alloying-element combinations that result in a nanostructured combination of constituents. The latter is a leading approach for the production of third generation high-strength steels (see below).

5.5.6 Ultra-high-strength Steels

These steels are defined by AISI as those with an ultimate tensile strength higher than 700 MPa. There are five important types of ultra-high-strength steel: dual-phase, TRIP, TWIP, complex phase, and martensitic. The main application of these steels is for crashworthy design of automobiles. The use of stronger steels allows for smaller cross sections in structural components, thus resulting in weight savings and fuel economy increases without compromising safety. The significant drawbacks of all these steels are higher cost, higher tool and die wear, higher forming loads, and more springback.

Dual-phase steels are processed specially for a mixed ferrite and martensite structure. They have a high work-hardening exponent [n in Eq. (2.8)], which improves their ductility and formability.

TRIP steels consist of a ferrite–bainite matrix and 5–20% retained austenite. During forming, the austenite progressively transforms into martensite. Thus, TRIP steels have both excellent ductility because of the austenite and high strength after forming. As a result, these steels can be used to produce more complex parts than other high-strength steels.

TWIP steels (from *TW*inning-*I*nduced *P*lasticity) are austenitic and have high manganese content (17–20%). These steels derive their properties from the generation of twins during deformation (see Section 1.4) without a phase change, resulting in very high strain hardening and avoiding necking during forming. As can be seen in Fig. 5.6, TWIP steels combine high strength and high formability.

Complex-phase grades (CP grades) are very fine-grained microstructures of ferrite and a high volume fraction of hard phases (martensite and bainite). These steels can have ultimate tensile strengths as high as 800 MPa, and are therefore of interest for automotive crash considerations, such in bumpers and roof supports. *Martensitic grades* are also available, consisting of high fractions of martensite and attaining tensile strengths as high as 1500 MPa.

Terminology has been developed to refer to the generations of advanced high strength steels (AHSS), differentiated by their color shading in Fig. 5.6:

- Conventional high-strength steels include the traditional mild grades, bake hardenable, and HSLA grades.
- **First Generation AHSS** refer to dual phase, complex phase, TRIP, and martensitic steels.
- TWIP steels are part of a class of materials referred to as **Second Generation AHSS**. Because of their very high strength, springback (see Section 16.5) is a concern, hence these materials must be hot stamped (see Section 16.11).
- **Third Generation AHSS** are now becoming commercially available. They combine the high strength of Second Generation AHSS with the improved formability of First Generation AHSS, through careful control of microstructures and phases. For example, a high strength phase such as martensite or ultra fine-grained ferrite may be mixed with a constituent that is highly formable and ductile, such as austenite. These materials can be cold-formed, thereby eliminating the challenges introduced by hot stamping (see Section 16.11).

Case Study 5.1 Advanced High-strength Steels in Automobiles

Increasing fuel economy in automobiles has received considerable attention in recent years for both environmental and economic reasons. Regulatory requirements call for automobile manufacturers to achieve corporate average fuel economy (CAFE) standards. To achieve higher fuel economy without compromising performance or safety, manufacturers have increasingly used advanced high-strength steels in structural elements of automobiles. For example, note the application of steel in the 2016 Honda Civic automobile shown in Fig. 5.7. Compared to the 2013 model, the use of hot stamped steel has increased from 1 to 14% of the body weight, and high strength steel makes up 58% of the body weight.

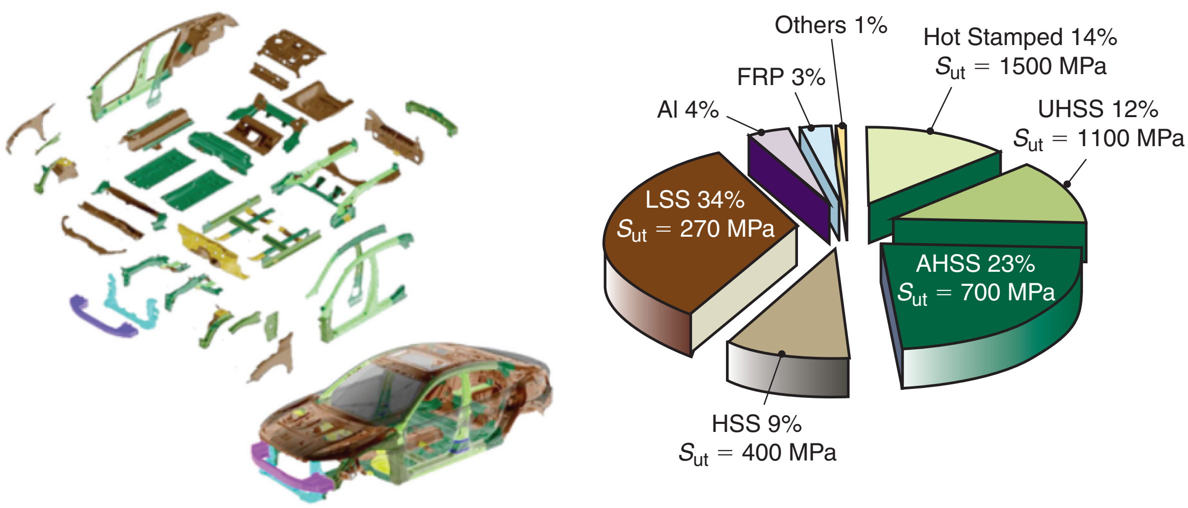
5.6 Stainless Steels

Stainless steels are characterized primarily by their corrosion resistance and high strength and ductility. They are called *stainless* because, in the presence of oxygen (air), they develop a thin, hard, adherent film of chromium oxide that protects the metal from corrosion (*passivation*; see Section 3.8). This protective film builds up again in the event that the surface is scratched. For passivation to occur, the minimum chromium content should be 10–12% by weight. In addition to chromium and carbon, other alloying elements in stainless steels are nickel, molybdenum, copper, titanium, silicon, manganese, columbium, aluminum, nitrogen, and sulfur.

The higher the carbon content, the lower is the corrosion resistance of stainless steels. The reason is that the carbon combines with the chromium in the steel and forms chromium carbide; the reduced presence of chromium oxide lowers the passivity of the steel. In addition, the chromium carbide introduces a second phase, thereby promoting galvanic corrosion.

Developed in the early 1900s, stainless steels are produced in electric furnaces or by the basic-oxygen process, and by techniques similar to those used in other types of steelmaking processes. The level of purity is controlled through various refining techniques. Stainless steels are available in a wide variety of shapes. Typical applications include cutlery, kitchen equipment, health care and surgical equipment, and applications in the chemical, food-processing, and petroleum industries.

Stainless steels generally are divided into five types (see also Table 5.6).



(a) (b)

Figure 5.7: Advanced high-strength steel applications in the 2016 Honda Civic. (a) Structural components and alloy used; (b) Classes of material and weight contribution, with typical ultimate strength range. The color schemes are the same in both images; note: UHSS = ultra-high–strength steel; AHSS = Advanced high-strength steel; HSS = high-strength steel; LSS = low-strength steel; Al is 6000 series aluminum components; FRP = fiber-reinforced composite. *Source:* Courtesy of the American Iron and Steel Institute.

Table 5.6: Mechanical Properties and Typical Applications of Selected Annealed Stainless Steels at Room Temperature.

AISI (UNS) designation	Ultimate tensile strength (MPa)	Yield strength (MPa)	Elongation in 50 mm (%)	Characteristics and typical applications
303 (S30300)	550–620	240–260	53–50	Screw machine products (shafts, valves, bolts, bushings, and nuts) and aircraft fittings (bolts, nuts, rivets, screws, and studs)
304 (S30400)	5–620	240–290	60–55	Chemical and food-processing equipment, brewing equipment, cryogenic vessels, gutters, downspouts, and flashings
316 (S31600)	50–590	210–290	60–55	High corrosion resistance and high creep strength, chemical- and pulp-handling equipment, photographic equipment, brandy vats, fertilizer parts, ketchup-cooking kettles, and yeast tubs
410 (S41000)	480–520	240–310	35–25	Machine parts, pump shafts, bolts, bushings, coal chutes, cutlery, tackle, hardware, jet engine parts, mining machinery, rifle barrels, screws, and valves
416 (S41600)	480–520	275	30–20	Aircraft fittings, bolts, nuts, fire extinguisher inserts, rivets, and screws

Austenitic (200 and 300 series). These steels generally are composed of chromium, nickel, and manganese in iron. They are nonmagnetic and have excellent corrosion resistance, but they are susceptible to stress-corrosion cracking (Section 3.8). Austenitic stainless steels, which are hardenable by cold working, are the most ductile of all stainless steels and can be easily formed . These steels are used in a wide variety of applications, such as kitchenware, fittings, welded construction, lightweight transportation equipment, furnace and heat-exchanger parts, and as components for severe chemical environments.

Ferritic (400 series). These steels have high chromium content. They are magnetic and have good corrosion resistance, but have lower ductility than austenitic stainless steels. Ferritic stainless steels are hardenable by cold working. They generally are used for nonstructural applications, such as kitchen equipment and automotive trim.

Martensitic (400 and 500 series). Most martensitic stainless steels do not contain nickel and are hardenable only by heat treatment. These steels are magnetic, and have high strength, hardness, and fatigue resistance, good ductility, and moderate corrosion resistance. They typically are used for cutlery, surgical tools, instruments, valves, and springs.

Precipitation-hardening (PH). These stainless steels contain chromium and nickel, along with copper, aluminum, titanium, or molybdenum. They have good corrosion resistance and ductility, and have high strength at elevated temperatures. Their main applications are in aircraft and aerospace structural components.

Duplex Structure. These stainless steels have a mixture of austenite and ferrite. They have good strength and higher resistance to both corrosion (in most environments) and stress-corrosion cracking than do the 300 series of austenitic steels. Typical applications are in water-treatment plants and for heat-exchanger components.

Case Study 5.2 Stainless Steels in Automobiles

The types of stainless steel usually selected by materials engineers for use in automobile parts are 301, 409, 430, and 434. Because of its good corrosion resistance and mechanical properties, type 301 is used for wheel covers. Cold working during the forming process increases its yield strength and gives the wheel cover a springlike action. Type 409 is used extensively for catalytic converters. In addition to being corrosion resistant, type 434 closely resembles the color of chromium plating, thus an attractive alternative to 430.

Stainless steels are also well suited for use in various automobile components, such as exhaust manifolds (replacing cast-iron manifolds to reduce weight, and increasing durability, providing higher thermal conductivity, and reduced emissions), mufflers, tailpipes, and brake tubing.

5.7 Tool and Die Steels

Tool and die steels are specially alloyed steels (Tables 5.7 and 5.8), designed for tool and die requirements such as high strength, impact toughness, and wear resistance at room and elevated temperatures. They commonly are used in the forming and machining of metals (Parts III and IV).

5.7.1 High-speed Steels

High-speed steels (HSS) are the most highly alloyed tool and die steels. First developed in the early 1900s, they maintain their hardness and strength at elevated operating temperatures. There are two basic types of high-speed steels: the **molybdenum type** (M-series) and the **tungsten type** (T-series).

The **M-series** steels contain up to about 10% molybdenum, with chromium, vanadium, tungsten, and cobalt as other alloying elements. The **T-series** steels contain 12–18% tungsten, with chromium, vanadium, and cobalt as other alloying elements. The M-series steels generally have higher abrasion resistance than

Table 5.7: Basic Types of Tool and Die Steels.

Type	AISI
High speed	M (molybdenum base)
	T (tungsten base)
Hot work	H1 to H19 (chromium base)
	H20 to H39 (tungsten base)
	H40 to H59 (molybdenum base)
Cold work	D (high carbon, high chromium)
	A (medium alloy, air hardening)
	O (oil hardening)
Shock resisting	S
Mold steels	P1 to P19 (low carbon)
	P20 to P39 (others)
Special purpose	L (low alloy)
	F (carbon–tungsten)
Water hardening	W

Table 5.8: Processing and Service Characteristics of Common Tool and Die Steels.

AISI designation	Resistance to decarb-urization	Resistance to cracking	Approx. hardness (HRC)	Machinability	Toughness	Resistance to softening	Resistance to wear
M2	Medium	Medium	60–65	Medium	Low	Very high	Very high
H11, 12, 13	Medium	Highest	38–55	Medium to high	Very high	High	Medium
A2	Medium	Highest	57–62	Medium	Medium	High	High
A9	Medium	Highest	35–56	Medium	High	High	Medium to high
D2	Medium	Highest	54–61	Low	Low	High	High to very high
D3	Medium	High	54–61	Low	Low	High	Very high
H21	Medium	High	36–54	Medium	High	High	Medium to high
P20	High	High	28–37	Medium to high	High	Low	Low to medium
P21	High	Highest	30–40	Medium	Medium	Medium	Medium
W1, W2	Highest	Medium	50–64	Highest	High	Low	Low to medium

T-series, undergo less distortion in heat treatment, and are less expensive. High-speed steel tools can be coated with titanium nitride and titanium carbide for improved wear resistance (see Chapter 34).

5.7.2 Die Steels

Hot-work steels (H-series) are designed for use at elevated temperatures; they have high toughness and high resistance to wear and cracking. The alloying elements are generally tungsten, molybdenum, chromium, and vanadium. **Cold-work steels** (A-, D-, and O-series) are used for cold-working operations. They generally have high resistance to wear and cracking, and are available as oil-hardening or air-hardening types. **Shock-resisting steels** (S-series) have impact toughness and are used in applications such as header dies, punches, and chisels. Various tool and die materials for a variety of manufacturing applications are given in Table 5.9.

Table 5.9: Typical Tool and Die Materials for Metalworking Processes.

Process	Material
Die casting	H13, P20
Powder metallurgy	
Punches	A2, S7, D2, D3, M2
Dies	WC, D2, M2
Molds for plastics and rubber	S1, O1, A2, D2, 6F5, 6F6, P6, P20, P21, H13
Hot forging	6F2, 6G, H11, H12
Hot extrusion	H11, H12, H13, H21
Cold heading	W1, W2, M1, M2, D2, WC
Cold extrusion	
Punches	A2, D2, M2, M4
Dies	O1, W1, A2, D2
Coining	52100, W1, O1, A2, D2, D3, D4, H11, H12, H13
Drawing	
Wire	WC, diamond
Shapes	WC, D2, M2
Bar and tubing	WC, W1, D2
Rolls	
Rolling	Cast iron, cast steel, forged steel, WC
Thread rolling	A2, D2, M2
Shear spinning	A2, D2, D3
Sheet metals	
Cold Shearing	D2, A2, A9, S2, S5, S7
Hot Shearing	H11, H12, H13
Pressworking	Zinc alloys, 4140 steel, cast iron, epoxy composites, A2, D2, O1
Deep drawing	W1, O1, cast iron, A2, D2
Machining	Carbides, high-speed steels, ceramics, diamond, cubic boron nitride

Notes: Tool and die materials usually are hardened 55 to 65 HRC for cold working and 30 to 55 HRC for hot working. Tool and die steels contain one or more of the following major alloying elements: chromium, molybdenum, tungsten, and vanadium. (For further details, see the bibliography at the end of this chapter.)

Summary

- The major categories of ferrous metals and alloys are carbon steels, alloy steels, stainless steels, and tool and die steels. Their wide range of properties, availability, and their generally low cost have made them among the most useful of all metallic materials.

- Steelmaking processes increasingly involve continuous-casting and secondary-refining techniques, resulting in higher quality steels and higher productivity.

- Carbon steels are generally classified as low-carbon (mild steel), medium-carbon, and high-carbon steels. Alloy steels contain several alloying elements, particularly chromium, nickel, and molybdenum.

- High-strength low-alloy (HSLA) steels have a low carbon content and consist of fine-grained ferrite as one phase and a second phase of martensite and austenite. Micro- and nanoalloyed steels are fine-grained, high-strength low-alloy steels that provide superior properties without the need for heat treatment.

- Second-generation high-strength steels have been developed; they provide exceptional strength and are used for applications where a combination of high strength and stiffness are required, such as in structural components protecting operators of automobiles. Third-generation steels are becoming available, combining the high strength of second generation steels with higher ductility; they are cold formable, whereas second-generation steels are mainly hot stamped.

- Stainless steels have chromium as the major alloying element; they are called stainless because they form a passivating chromium-oxide layer on their surface. These steels are generally classified as austenitic, ferritic, martensitic, and precipitation-hardening steels.
- Tool and die steels are among the most important metallic materials, and are used widely in casting, forming, and machining operations. They generally consist of high-speed steels, hot- and cold-work steels, and shock-resisting steels.

Key Terms

Alloy steels
Basic-oxygen furnace
Blast furnace
Carbon steels
Complex-phase steels
Continuous casting
Dual-phase steels
Electric furnace
High speed steel
High-strength low-alloy steels
Ingot
Killed steel
Martensitic steel
Microalloyed steels
Nanoalloyed steels
Open-hearth furnace
Pig iron
Refining
Rimmed steel
Semi-killed steel
Stainless steels
Steel
Strand casting
Third generation steel
Tool and die steels
Trace elements
TRIP steels
TWIP steels
Vacuum furnace

Bibliography

ASM Handbook, Vol. 1: **Properties and Selection: Iron, Steels, and High-Performance Alloys**, ASM International, 1990.
ASM Specialty Handbook: **Carbon and Alloy Steels**, ASM International, 1995.
ASM Specialty Handbook: **Stainless Steels**, ASM International, 1994.
ASM Specialty Handbook: **Tool Materials**, ASM International, 1995.
Beddoes, J., and Parr, J.G., **Introduction to Stainless Steels**, 3rd ed., ASM International, 1999.
Bhadeshia, H., and Honeycombe, R., **Steels: Microstructure and Properties**, 4th ed., Butterworth-Heinemann, 2017.
Demeri, M.Y., **Advanced High-Strength Steels: Science, Technology and Applicstions,** ASM International, 2013.
Krauss, G., **Steels: Processing, Structure, and Performance**, 2nd ed., ASM International, 2015.
Mandal, S.K., **Steel Metallurgy: Properties, Specifications and Applications**, McGraw-Hill, 2015.
Reed, C., **The Superalloys: Fundamentals and Applications**, Cambridge University Press, 2008.
Roberts, G.A., Krauss, G., and Kennedy, R., **Tool Steels**, 5th ed., ASM International, 1998.

Review Questions

5.1. What are the major categories of ferrous alloys?

5.2. Why is steel so commonly used?

5.3. List the basic raw materials used in making iron and steel, and explain their functions.

5.4. List the types of furnaces commonly used in steelmaking, and describe their characteristics.

5.5. List and explain the characteristics of the types of steel ingots.

5.6. What does refining mean? How is it done?

5.7. What is continuous casting? What advantages does continuous casting have over casting into ingots?

5.8. What is the role of a tundish in continuous casting?

5.9. Name the four alloying elements that have the greatest effect on the properties of steels.

5.10. What are trace elements?

5.11. What are the percentage carbon contents of low-carbon, medium-carbon, and high-carbon steels?

5.12. How do stainless steels become stainless?

5.13. What are the major alloying elements in tool and die steels and in high-speed steels?

5.14. How does chromium affect the surface characteristics of stainless steels?

5.15. What kinds of furnaces are used to refine steels?

5.16. What is high-speed steel?

5.17. What is TRIP? TWIP?

5.18. What are the applications of advanced high-strength steels?

5.19. What characteristics are common among die steels?

5.20. What effect does carbon content have on mechanical properties of steel? What effects does it have on physical properties?

5.21. What is killed steel? In this context, what does ‘killed’ refer to?

Qualitative Problems

5.22. Identify several different products that are made of stainless steel, and explain why they are made of that material.

5.23. Professional cooks generally prefer carbon-steel to stainless-steel knives, even though the latter are more popular with consumers. Explain the reasons for those preferences.

5.24. Why is the control of the structure of an ingot important?

5.25. Explain why continuous casting has been such an important technological advancement.

5.26. Describe applications in which you would not want to use carbon steels.

5.27. Explain what would happen if the speed of the continuous-casting process shown in Fig. 5.4a is (a) higher or (b) lower than that indicated, typically 25 mm/s.

5.28. The cost of mill products of metals increases with decreasing thickness and section size. Explain why.

5.29. Describe your observations regarding the information given in Table 5.9.

5.30. How do trace elements affect the ductility of steels?

5.31. Comment on your observations regarding Table 5.1.

5.32. In Table 5.9, D2 steel is listed as a more common tool and die material for most applications. Why is this so?

5.33. List the common impurities in steel. Which of these are the ones most likely to be minimized if the steel is melted in a vacuum furnace?

5.34. Explain the purpose of the oil shown at the top left of Fig. 5.4a given that the molten-steel temperatures are far above the ignition temperatures of the oil.

5.35. Recent research has identified mold-surface textures that will either (a) inhibit a solidified steel from separating from the mold or (b) force it to stay in contact in continuous casting. What is the advantage of a mold that maintains intimate contact with the steel?

5.36. Identify products that cannot be made of steel, and explain why this is so. (For example, electrical contacts commonly are made of gold or copper, because their softness results in low contact resistance, whereas for steel, the contact resistance would be very high.)

5.37. List and explain the advantages and disadvantages of using advanced high-strength steels.

Quantitative Problems

5.38. Conduct an internet search and determine the chemical composition of (a) TRIP 450/800; (b) 304 stainless steel; (c) 4140 steel. If a foundry ladle will pour 60,000 kg, calculate the weight of each element in the ladle.

5.39. Refer to the available literature, and estimate the cost of the raw materials for (a) an aluminum beverage can, (b) a stainless-steel two-quart cooking pot, and (c) the steel hood of a car.

5.40. In Table 5.1, more than one type of steel is listed for some applications. Refer to data available in the technical literature listed in the bibliography, and determine the range of properties for these steels in various conditions, such as cold worked, hot worked, and annealed.

5.41. Some soft drinks are now available in steel cans (with aluminum tops) that look similar to aluminum cans. Obtain one of each type, weigh them when empty, and determine their respective wall thicknesses.

5.42. Using strength and density data, determine the minimum weight of a 1 m-long tension member that must support a load of 4 kN, manufactured from (a) annealed 303 stainless steel, (b) normalized 8620 steel, (c) as-rolled 1080 steel, (d) any two aluminum alloys, (e) any brass alloy, and (f) pure copper.

5.43. The endurance limit (fatigue life) of steel is approximately one-half the ultimate tensile strength (see Fig. 2.16), but never higher than 700 MPa. For iron, the endurance limit is 40% of the ultimate strength, but never higher than 170 MPa. Plot the endurance limit vs. the ultimate strength for the steels described in this chapter and for the cast irons shown in Table 12.3. On the same plot, show the effect of surface finish by plotting the endurance limit, assuming that the material is in the as-cast state (see Fig. 2.29).

5.44. Using the data given in Table 5.4, obtain the power-law curves for the advanced high-strength steels shown and plot the curves. Compare these materials with those given in Table 2.3.

Synthesis, Design, and Projects

5.45. Based on the information given in Section 5.5.1, make a table with columns for each improved property, such as hardenability, strength, toughness, and machinability. In each column, list the elements that improve that particular property and identify the element that has the most influence.

5.46. Assume that you are in charge of public relations for a large steel-producing company. Outline all of the attractive characteristics of steels that you would like your customers to be informed about.

5.47. Assume that you are in competition with the steel industry and are asked to list all of the characteristics of steels that are not attractive. Make a list of those characteristics and explain their relevance to engineering applications.

5.48. Section 5.5.1 noted the effects of various individual elements, such as lead alone or sulfur alone, on the properties and characteristics of steels. What was not discussed, however, was the role of combinations of these elements (such as lead and sulfur together). Review the technical literature, and prepare a table indicating the combined effects of several elements on steels.

5.49. In the past, waterfowl hunters used lead shot in their shotguns, but this practice resulted in lead poisoning of unshot birds that ingested lead pellets (along with gravel) to help them digest food. Steel and tungsten are being used as replacement materials. If all pellets have the same velocity upon exiting the shotgun barrel, what concerns would you have regarding this substitution of materials? Consider both performance and environmental effects.

5.50. Aluminum is being used as a substitute material for steel in automobiles. Describe your concerns, if any, in purchasing an aluminum automobile.

5.51. In the 1940s (The Second World War), the *Yamato* and its sister ship, the *Musashi*, were the largest battleships ever built. Find out the weight of these ships, and estimate the number of automobiles that could have been built from the steel used in just one such ship. Estimate the time it would take to cast that much steel by continuous casting.

5.52. Search the technical literature, and add more parts and materials to those shown in Table 5.1.

5.53. Referring to Fig. 5.4, note that the mold has cooling channels incorporated to remove heat. Can continuous casting be done without such cooling channels? Can it be done with a heated mold? Explain your answer.

Chapter 6

Nonferrous Metals and Alloys: Production, General Properties, and Applications

Case Studies:

- Nonferrous metals include a wide variety of materials, with special properties that are indispensable in most products.
- This chapter introduces each class of nonferrous metal and its alloys, and briefly describes their methods of production.
- Their physical and mechanical properties are then summarized, along with general guidelines for their selection and applications.
- Shape-memory alloys, amorphous alloys, metal foams, and rare earth metals are also described, with examples of their unique applications.

6.1 Introduction

Nonferrous metals and alloys cover a very wide range, from the more common metals (such as aluminum, copper, and magnesium) to high-strength, high-temperature alloys (such as those of tungsten, tantalum, and molybdenum). Although generally more expensive than ferrous metals (Table 6.1), **nonferrous** metals have numerous important applications because of such properties as corrosion resistance, high thermal and electrical conductivity, low density, and ease of fabrication (Table 6.2).

Typical examples of nonferrous metal and alloy applications include aluminum for aircraft bodies and cooking utensils, copper wire for electrical power cords, zinc for galvanized sheet metal for car bodies, titanium for jet-engine turbine blades and for orthopedic implants, and tantalum for rocket engine components. As an example, the turbofan jet engine (Fig. 6.1) for the Boeing 757 aircraft typically contains the following nonferrous metals and alloys: 38% Ti, 37% Ni, 12% Cr, 6% Co, 5% Al, 1% Nb, and 0.02% Ta.

Table 6.1: Approximate Cost-per-unit-volume for Wrought Metals and Plastics Relative to the Cost of Carbon Steel. Table data are representative of 25-mm diameter bar stock.

Material	Relative cost
Gold	70,000
Silver	680
Molybdenum alloys	200–250
Nickel	40
Titanium alloys	25–40
Copper alloys	8–10
Zinc alloys	1.5–3.5
Stainless steels	2–9
Magnesium alloys	2–4
Aluminum alloys	1.5–3
High-strength low-alloy steels	1.4
Gray cast iron	1.2
Carbon steel	1
Nylons, acetals, and silicon rubber*	1.1–2
Rubber*	0.2–1
Other plastics and elastomers*	0.2–2

* As molding compounds.

Note: Costs vary significantly with quantity of purchase, supply and demand, size and shape, and various other factors.

Table 6.2: General Characteristics of Nonferrous Metals and Alloys.

Material	Characteristics
Nonferrous alloys	More expensive than steels and plastics; wide range of mechanical, physical, and electrical properties; good corrosion resistance; high-temperature applications
Aluminum	Alloys have high strength-to-weight ratio; high thermal and electrical conductivity; good corrosion resistance; good manufacturing properties
Magnesium	Lightest metal; good strength-to-weight ratio
Copper	High electrical and thermal conductivity; good corrosion resistance; good manufacturing properties
Superalloys	Good strength and resistance to corrosion at elevated temperatures; can be iron-, cobalt-, and nickel-based alloys
Tin	Good corrosion resistance and bright appearance; used also in solders and as bearing materials.
Titanium	Highest strength-to-weight ratio of all metals; good strength and corrosion resistance at high temperatures
Refractory metals	Molybdenum, niobium, tungsten, and tantalum; high strength at elevated temperatures
Precious metals	Gold, silver, and platinum; generally good corrosion resistance and aesthetic characteristics.
Zinc	Very good corrosion resistance; commonly used in castings and galvanizing steel sheet for corrosion protection.
Rare earths	Unique combinations of magnetic and electrical properties; commonly used in magnetic devices, high-capacity storage batteries and microelectronics.

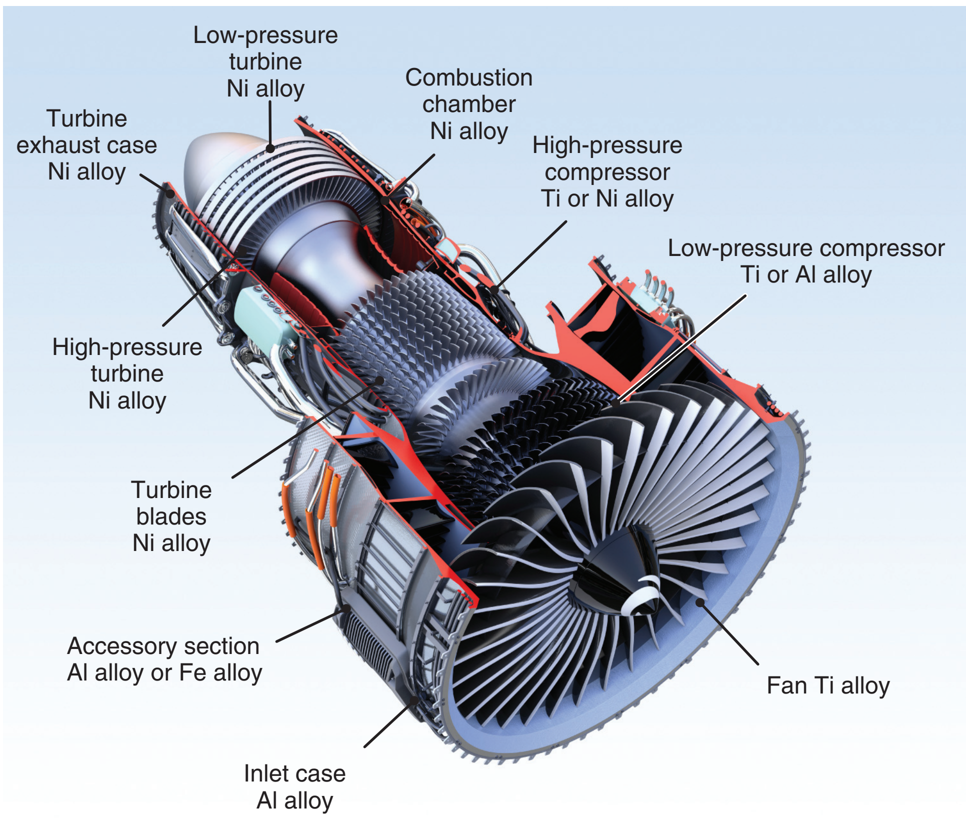


Figure 6.1: Cross section of a jet engine (PW2037) showing various components and the alloys used in manufacturing them. *Source:* Shutterstock/Chesky.

This chapter introduces the general properties, production methods, and engineering applications of nonferrous metals and their alloys. The manufacturing properties of these materials, such as formability, machinability, and weldability, are described in various chapters throughout this text.

6.2 Aluminum and Aluminum Alloys

The important characteristics of **aluminum** (Al) and its alloys are their high strength-to-weight ratios, resistance to corrosion, high thermal and electrical conductivities, nontoxicity, reflectivity, appearance, and ease of formability and machinability; also, they are nonmagnetic. The principal uses of aluminum and its alloys, in decreasing order of consumption, are in containers and packaging (aluminum beverage cans and foil); architectural and structural applications; transportation (aircraft and aerospace applications, buses, automobiles, railroad cars, and marine craft); electrical applications (as economical and nonmagnetic electrical conductors); computer, tablets, and smart phone casings; consumer durables (appliances, cooking utensils, and furniture), and portable tools (Tables 6.3 and 6.4). Nearly all high-voltage transmission wiring is made of aluminum.

In its structural (load-bearing) components, 82% of a Boeing 747 aircraft and 70% of a Boeing 777 aircraft is aluminum. Although the Boeing 787 Dreamliner (first placed into service in late 2011) is well recognized for its carbon fiber-reinforced composite fuselage, it still uses 20% aluminum, by weight, as compared to 15% titanium. The frame and the body panels of the Rolls Royce Phantom coupe are made of aluminum, improving the car's strength-to-weight and torsional rigidity-to-weight ratios.

Aluminum alloys are available as mill products, that is, as wrought products made into various shapes by rolling, extrusion, drawing, and forging (Chapters 13 through 15). Aluminum ingots are available for casting, as is aluminum in powder form for powder-metallurgy applications (Chapter 17). Most aluminum alloys can be machined, formed, and welded with relative ease. There are two types of wrought alloys of aluminum: (a) Alloys that can be hardened by mechanical processing and are not heat treatable, and (b) alloys that can be hardened by heat treatment.

Unified Numbering System. Aluminum and other nonferrous metals and alloys are identified internationally by the Unified Numbering System (UNS), consisting of a letter, indicating the general class of the alloy, followed by five digits, indicating its chemical composition. For example, A for aluminum, C for copper, N for nickel alloys, P for precious metals, and Z for zinc. Also, in the UNS designation, 2024 wrought aluminum alloy is A92024.

Table 6.3: Properties of Selected Aluminum Alloys at Room Temperature.

Alloy (UNS)	Temper	Ultimate tensile strength (MPa)	Yield strength (MPa)	Elongation in 50 mm (%)
1100 (A91100)	O	90	35	35–45
1100	H14	125	120	9–20
2024 (A92024)	O	190	75	20–22
2024	T4	470	325	19–20
2099	T83	560	525	9
2099	T8E67	530	485	10
3003 (A93003)	O	110	40	30–40
3003	H14	150	145	8–16
5052 (A95052)	O	190	90	25–30
5052	H34	260	215	10–14
6061 (A96061)	O	125	55	25–30
6061	T6	310	275	12–17
7075 (A97075)	O	230	105	16–17
7075	T6	570	500	11

Table 6.4: Manufacturing Characteristics and Typical Applications of Selected Wrought Aluminum Alloys.

	Characteristics*			
Alloy	Corrosion resistance	Machinability	Weldability	Typical Applications
1100	A	C–D	A	Sheet-metal work, spun hollowware, tin stock, power transmission lines
2024	C	B–C	B–C	Truck wheels, screw machine products, most widely used aluminum alloy for aircraft structures
2099	A	B–C	A	Lightweight lithium alloy with limited fracture toughness, used in aerospace and lightweight applications.
3003	A	C–D	A	Cooking utensils, chemical equipment, pressure vessels, sheet-metal work, builders' hardware, storage tanks
5052	A	C–D	A	Sheet-metal work, hydraulic tubes, and appliances; pressure vessels; bus, truck, and marine uses
6061	B	C–D	A	Heavy-duty structures where corrosion resistance is needed; truck and marine structures, railroad cars, furniture, pipelines, bridge railings, hydraulic tubing. Used increasingly for hydroformed tubing and extrusions for lightweight vehicles
7075	C	B–D	D	Aircraft and other structures, keys, hydraulic fittings

*A, excellent; D, poor.

Aluminum is widely seen as an essential metal for lightweight applications, such as vehicles where fuel economy goals have to be met. While the 2000 and 6000 series aluminum alloys are widely used for vehicles, new alloys, especially those containing lithium as the main alloying element, are now receiving significant interest. Aluminum 2099, for example, contains up to 3.0% copper and 2.0% lithium, along with other elements, and provides exceptional corrosion resistance and high strength, with a 4% reduction in density and 3% increase in stiffness as compared to other aluminum alloys. The main drawbacks to aluminum–lithium alloys are cost and limited fatigue and fracture strengths, although material innovations continue.

Porous Aluminum. Blocks of aluminum are produced that are 37% lighter than solid aluminum and have uniform permeability and *microporosity*. This characteristic allows their use in applications where a vacuum or differential pressure has to be maintained. Examples are vacuum holding of fixtures for assembly and automation (Section 37.8), and vacuum forming or thermoforming of plastics (Section 19.6). The blocks are 70–90% aluminum powder; the rest is epoxy resin. They can be machined with relative ease and can be joined using adhesives or special welding processes.

Case Study 6.1 Aluminum Production from Ore: The Hall-Héroult Process

First produced in 1825, aluminum is the most abundant metallic element, making up about 8% of the earth's crust, and is produced in a quantity second only to iron. The principal ore for aluminum is *bauxite*, which is hydrous (water-containing) aluminum oxide and includes other oxides. Bauxite generally contains 30–60% aluminum oxide (alumina, Al_2O_3) combined with other elements such as iron. The first step in producing aluminum is to extract the aluminum oxide from the bauxite. This is done in a multi-step process:

1. The bauxite is first crushed in a comminution mill to an aggregate form.
2. The ore is then heated to a temperature of 150–200°C in a sodium hydroxide (NaOH) solution, this results in the production of sodium aluminate ($NaAlO_2$).
3. The sodium aluminate produces aluminum hydroxide, $Al(OH)_3$, which is then converted to alumina in rotary kilns at temperatures up to 1000°C.

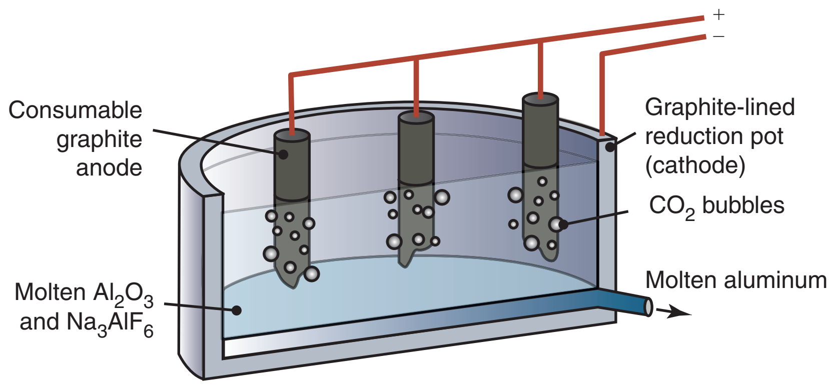


Figure 6.2: Schematic illustration of a Hall-Héroult reduction pot, used to produce molten aluminum from aluminum oxide.

The Hall-Héroult process (Fig. 6.2) produces pure aluminum from the alumina in a carbon- or graphite-lined container, referred to as a reduction pot. In the reduction pot, a small voltage, as low as 6 volts, is used, but the current can be extremely high-over 150,000 amperes. The Hall-Héroult process uses a carbon anode that combines with the oxygen in the alumina, forming a combination of carbon dioxide and carbon monoxide, leaving molten aluminum to collect in the bottom of the pot, where it is siphoned periodically.

The Hall-Héroult process enabled aluminum to be an inexpensive commodity metal instead of a precious metal. The demand for aluminum is so high that it has been estimated that 5% of all electricity generated in the United States is used in Hall-Héroult cells. This has serious implications, since it has been estimated that 11.4 metric tons of CO_2 emissions are generated per ton of aluminum produced (see Section 40.5).

Case Study 6.2 The Tesla Model S 60

Aluminum use in automobiles and in light trucks has been increasing steadily. As recently as 1990, there were no aluminum-structured passenger cars in production in the United States. It is expected that the automotive industry will need to reduce the mass of vehicles by around 9%, by 2028, in order to achieve mandated fuel efficiency (CAFÉ) standards, requiring material advances with most metals. However, its light weight, high strength, formability, weldability and corrosion resistance of aluminum makes it very attractive for the production of lightweight vehicles. This would require the average aluminum content in vehicles to increase from the current weight of around 187 to 212 kg.

The Tesla Model S is an electric vehicle, requiring around 455 kg of batteries to supply power to its electric motors. In order to achieve a curb weight of around 1900 kg, while seating up to seven passengers, a number of design innovations were required, including a lightweight aluminum space frame and an aluminum body (though with some selected steel reinforcement in some pillars for the operator zone). The aluminum sections are held together using a mix of cold-metal transfer welding, conventional spot welding, self-piercing rivets, and structural adhesives. The design also allows the batteries to be mounted low in the car, thus lowering the center of gravity and maximizing interior space.

Many of the design innovations involve non-structural features such as several cameras to allow near self-driving, numerous airbags and a unique chassis design that places the battery at the base of the vehicle to maximize stability. However, since the electric vehicle does not have a large motor, it can use advanced extrusions and castings to produce an energy-absorbing frame.

(a) (b)

Figure 6.3: The Tesla S 60. (a) Image of the Tesla S 60 electric automobile, with an aluminum alloy body; (b) Chasis showing the battery supported and protected by a titanium underbody. *Source:* (a) Shutterstock/Dimitris Leonidas (b) Shutterstock/Alexander Kondratenko.

6.3 Magnesium and Magnesium Alloys

Magnesium (Mg) is the lightest engineering metal and has good vibration-damping characteristics. Its alloys are used in structural as well as nonstructural applications wherever weight is of primary importance. Magnesium is also an alloying element in various nonferrous metals. The main drawback to magnesium is its high cost (Table 6.1).

Typical uses of magnesium alloys are in aircraft and missile components, material-handling equipment, portable power tools, ladders, luggage, bicycles, sporting goods, and general lightweight components. Magnesium is finding increased use in the automotive sector, mainly for weight savings. Magnesium alloys are available either as castings (such as die-cast camera frames) or as wrought products (such as extruded bars and shapes, forgings, and rolled plates and sheet). Its alloys are also used in printing and textile machinery to minimize inertial forces in high-speed components of machinery

Because it is not sufficiently strong in its pure form, magnesium is alloyed with various elements (Table 6.5) in order to impart certain specific properties, particularly high strength-to-weight ratio. A variety of magnesium alloys have good casting, forming, and machining characteristics. Because magnesium as powder or in chip forms (as from machining) oxidizes rapidly (*i.e.*, they are *pyrophoric*), a fire hazard exists; thus precautions must be taken when machining, grinding, or sand-casting magnesium alloys. Products made of magnesium and its alloys are, however, not a fire hazard during their normal use.

Table 6.5: Properties and Typical Forms of Selected Wrought Magnesium Alloys.

Alloy	Nominal composition	Condition	Ultimate tensile strength (MPa)	Yield strength (MPa)	Elongation in 50 mm (%)	Typical forms
AZ31B	3.0 Al, 1.0 Zn, 0.2 Mn	F	260	200	15	Extrusions
		H24	290	220	15	Sheet and plate
AZ80A	8.5 Al, 0.5 Zn, 0.2 Mn	T5	380	275	7	Extrusions and forgings
AZ91D	9.0 Al, 0.03 Cu	F	230	160	3	Most common die cast alloy
HK31A	0.7 Zr, 3 Th	H24	255	200	8	Sheet and plates
ZE10	1.0 Zn, 1.0 Ce	F	263	163	16	Sheet and plates
ZEK199	1.0 Zn, 0.3 Zr, 1.0 Ce	F	311	308	19	Extrusions and sheet
ZK60A	5.7 Zn, 0.55 Zr	T5	365	300	11	Extrusions and forgings

Magnesium is easy to cast but normally difficult to form. Efforts have been made to promote the increased use of magnesium in automobiles through improved welding and sheet formability. Alloys ZEK100, AZ31 and ZE10 are of current high interest.

Production. Magnesium is the third-most-abundant metallic element (2%) in the earth's crust, after iron and aluminum. Most magnesium comes from seawater, which contains 0.13% magnesium as magnesium chloride. First produced in 1808, magnesium metal can be obtained either electrolytically or by thermal reduction. In the *electrolytic method*, seawater is mixed with lime (calcium hydroxide) in settling tanks. The magnesium hydroxide developed precipitates to the bottom, and is then filtered and mixed with hydrochloric acid. The resulting solution is subjected to electrolysis (as is done with aluminum), producing magnesium metal; it is then cast into ingots for further processing into various shapes.

In the *thermal-reduction method*, magnesium ores (dolomite, magnesite, and others) are broken down with reducing agents (such as powdered ferrosilicon, an alloy of iron and silicon) by heating the mixture in a vacuum chamber. As a result of this reaction, vapors of magnesium form, and they condense into magnesium crystals; they are then melted, refined, and poured into ingots to be processed further into various shapes.

6.4 Copper and Copper Alloys

First produced in about 4000 B.C., **copper** (Cu, from the Latin *cuprum*) and its alloys have properties somewhat similar to those of aluminum and its alloys. In addition, they have good corrosion resistance and are among the best conductors of electricity and heat (Tables 3.1 and 3.2). Copper and its alloys can be processed easily by forming, machining, casting, and joining techniques.

Copper alloys often are attractive for applications in which a combination of qualities, such as electrical, mechanical, nonmagnetic, corrosion-resistance, thermally conductivity, and wear-resistance are required. Applications include electrical and electronic components, springs, coins, plumbing components, heat exchangers, marine hardware, and consumer goods (such as cooking utensils, jewelry, and decorative objects). Although aluminum is the most common material for dies in polymer injection molding (Section 19.3), copper is often used because of its better thermal properties. Pure copper can be used as a solid lubricant in hot metal-forming operations (Section 33.7.6).

Copper alloys now have improved manufacturing characteristics, and can be heat treated to improve their mechanical properties. The most common copper alloys are brasses and bronzes. **Brass** (an alloy of copper and zinc) is one of the earliest alloys developed and has numerous applications, including decorative objects (Table 6.6). **Bronze** is an alloy of copper and tin (Table 6.7); there are also other bronzes, such as aluminum bronze (an alloy of copper and aluminum) and bismuth bronze. Beryllium copper (or beryllium bronze) and phosphor bronze have good strength and hardness, with applications such as springs and bearings. Other major copper alloys are copper nickels and nickel silvers.

Production. Copper is found in several types of ores, the most common being sulfide ores. The ores are generally of low grade (containing typically less than 5% copper) and usually are obtained from open-pit mines. The ore is ground into fine particles in ball mills (rotating cylinders with metal balls inside to crush the ore, as illustrated in Fig. 17.6b); the resulting particles are then suspended in water to form a slurry. Reducing chemicals and oil are added, and the mixture is agitated. The mineral particles form a *froth*, which is scraped and dried. The dry copper concentrate (as much as one-third of which is copper) is traditionally **smelted** (melted and fused) and refined, a process known as **pyrometallurgy**, because heat is used to refine the metal. For such applications as electrical conductors, the copper is further refined electrolytically to a purity of at least 99.95% (*oxygen-free electrolytic copper*). Copper is also processed by **hydrometallurgy**, involving both chemical and electrolytic reactions.

Table 6.6: Properties and Typical Applications of Selected Wrought Copper and Brasses.

Type and UNS number	Nominal composition (%)	Ultimate tensile strength (MPa)	Yield strength (MPa)	Elongation in 50 mm (%)	Typical applications
Electrolytic tough-pitch copper (C11000)	99.90 Cu, 0.04 O	220–450	70–365	55–4	Downspouts, gutters, roofing, gaskets, auto radiators, bus bars, nails, printing rolls, rivets
Red brass, 85% (C23000)	85.0 Cu, 15.0 Zn	270–725	70–435	55–3	Weather stripping, conduits, sockets, fasteners, fire extinguishers, condenser and heat-exchanger tubing
Cartridge brass, 70% (C26000)	70.0 Cu, 30.0 Zn	300–900	75–450	66–3	Radiator cores and tanks, flashlight shells, lamp fixtures, fasteners, locks, hinges, ammunition components, plumbing accessories
Free-cutting brass (C36000)	61.5 Cu, 3.0 Pb, 35.5 Zn	340–470	125–310	53–18	Gears, pinions, automatic high-speed screw machine parts
Naval brass (C46400 to C46700)	60.0 Cu, 39.25 Zn, 0.75 Sn	380–610	170–455	50–17	Aircraft: turnbuckle barrels, balls, bolts; marine hardware: propeller shafts, rivets, valve stems, condenser plates

Table 6.7: Properties and Typical Applications of Selected Wrought Bronzes.

Type and UNS number	Nominal composition (%)	Ultimate tensile strength (MPa)	Yield strength (MPa)	Elongation in 50 mm (%)	Typical applications
Architectural bronze (C38500)	57.0 Cu, 3.0 Pb, 40.0 Zn	415 (as extruded)	140	30	Architectural extrusions, storefronts, thresholds, trim, butts, hinges
Phosphor bronze, 5% A (C51000)	95.0 Cu, 5.0 Sn, trace P	325–960	130–550	64–2	Bellows, clutch disks, cotter pins, diaphragms, fasteners, wire brushes, chemical hardware, textile machinery
Free-cutting phosphor bronze (C54400)	88.0 Cu, 4.0 Pb, 4.0 Zn, 4.0 Sn	300–520	130–435	50–15	Bearings, bushings, gears, pinions, shafts, thrust washers, valve parts
Low-silicon bronze, (C65100)	98.5 Cu, 1.5 Si	275–655	100–475	55–11	Hydraulic pressure lines, bolts, marine hardware, electrical conduits, heat-exchanger tubing
Nickel silver, 65–10 (C74500)	65.0 Cu, 25.0 Zn, 10.0 Ni	340–900	125–525	50–1	Rivets, screws, slide fasteners, hollowware, nameplates

6.5 Nickel and Nickel Alloys

Nickel (Ni) is a silver-white metal and a major alloying element in metals imparting strength, toughness, and corrosion resistance. It is used extensively in stainless steels and in nickel-based alloys (also called **superalloys**). Nickel alloys are used in high-temperature applications (such as jet engine components, rockets, and nuclear power plants), food-handling and in chemical-processing equipment, coins, and marine applications. Because nickel is magnetic, its alloys are used in electromagnetic applications, such as solenoids.

The principal use of nickel as a metal is in the electroplating of parts for their appearance and for improvement of their corrosion and wear resistance. Nickel alloys have high strength and corrosion resistance at elevated temperatures. Common alloying elements in nickel are chromium, cobalt, and molybdenum. The behavior of nickel alloys in machining, forming, casting, and welding can be modified by various other alloying elements.

Table 6.8: Properties and Typical Applications of Selected Nickel Alloys (All Are Trade Names).

Type and UNS number	Nominal composition (%)	Ultimate tensile strength (MPa)	Yield strength (MPa)	Elongation in 50 mm (%)	Typical applications
Nickel 200 (annealed)	–	380–550	100–275	60–40	Chemical and food processing industry, aerospace equipment, electronic parts
Duranickel 301 (age hardened)	4.4 Al, 0.6 Ti	1300	900	28	Springs, plastics extrusion equipment, molds for glass, diaphragms
Monel R-405 (hot rolled)	30 Cu	525	230	35	Screw-machine products, water meter parts
Monel K-500 (age hardened)	29 Cu, 3 Al	1050	750	30	Pump shafts, valve stems, springs
Inconel 600 (annealed)	15 Cr, 8 Fe	640	210	48	Gas turbine parts, heat-treating equipment, electronic parts, nuclear reactors
Hastelloy C-4 (solution treated and quenched)	16 Cr, 15 Mo	785	400	54	Parts requiring high-temperature stability and resistance to stress-corrosion cracking

A variety of nickel alloys, with a wide range of strengths at different temperatures, have been developed (Table 6.8). Although trade names are still in wide use, nickel alloys are identified in the UNS system with the letter N; thus, for example, Hastelloy G is N06007. Other common trade names are as follow:

- **Monel** is a nickel–copper alloy.
- **Hastelloy** (also a nickel–chromium alloy) has good corrosion resistance and high strength at elevated temperatures.
- **Nichrome** (an alloy of nickel, chromium, and iron) has high electrical resistance and high resistance to oxidation and is used for electrical heating elements.
- **Invar** and **Kovar** (alloys of iron and nickel) have relatively low sensitivity to temperature changes (Section 3.6).

Production. The main sources of nickel are sulfide and oxide ores, all of which have low concentrations of nickel. The metal is produced by sedimentary and thermal processes, followed by electrolysis; this sequence yields 99.95% pure nickel.

6.6 Superalloys

Superalloys are important in high-temperature applications, hence they are also known as **heat-resistant** or **high-temperature alloys**. Superalloys generally have good resistance to corrosion, mechanical and thermal fatigue, mechanical and thermal shock, and creep and erosion at elevated temperatures.

Major applications of superalloys are in jet engines and gas turbines; other applications are in reciprocating engines, rocket engines, tools and dies for hot working operations, and in the nuclear, chemical, and petrochemical industries. Generally, superalloys are identified by trade names or by special numbering systems, and are available in a variety of shapes. Most superalloys have a maximum service temperature of about 1000°C in structural applications. For non-load bearing components, temperatures can be as high as 1200°C.

Superalloys are generally referred to as *iron-based*, *cobalt-based*, or *nickel-based*.

- **Iron-based superalloys** generally contain from 32 to 67% Fe, 15 to 22% Cr, and 9 to 38% Ni. Common alloys in this group are the *Incoloy* series.

Table 6.9: Properties and Typical Applications of Selected Superalloys at 870°C (All Are Trade Names).

Alloy	Condition	Ultimate tensile strength (MPa)	Yield strength (MPa)	Elongation in 50 mm (%)	Typical applications
Astroloy	Wrought	770	690	25	Forgings for high-temperature use
Co-28Cr-6Mo	Cast	655	450	8	Surgical implants
Hastelloy X	Wrought	255	180	50	Jet engine sheet parts
IN-100	Cast	885	695	6	Jet engine blades and wheels
IN-102	Wrought	215	200	110	Superheater and jet engine parts
Inconel 625	Wrought	285	275	125	Aircraft engines and structures, chemical processing equipment
Inconel 718	Wrought	340	330	88	Jet engine and rocket parts
MAR-M 200	Cast	840	760	4	Jet engine blades
MAR-M 432	Cast	730	605	8	Integrally cast turbine wheels
René 41	Wrought	620	550	19	Jet engine parts
Udimet 700	Wrought	690	635	27	Jet engine parts
Waspaloy	Wrought	525	515	35	Jet engine parts

- **Cobalt-based superalloys** generally contain from 35 to 65% Co, 19 to 30% Cr, and up to 35% Ni. These superalloys are not as strong as nickel-based superalloys, but they retain their strength at higher temperatures.

- **Nickel-based superalloys** are the most common of the superalloys and are available in a wide variety of compositions (Table 6.9). The proportion of nickel is from 38 to 76% and also contain up to 27% Cr and 20% Co. Common alloys in this group are the *Hastelloy, Inconel, Nimonic, René, Udimet, Astroloy,* and *Waspaloy* series.

6.7 Titanium and Titanium Alloys

Titanium (Ti, named after the Greek god Titan) is a silvery white metal discovered in 1791, but not produced commercially until the 1950s. Although titanium is expensive, its high strength-to-weight ratio and corrosion resistance at room and elevated temperatures make it attractive for many applications, including aircraft; jet engines (see Fig. 6.1); racing cars; golf clubs; chemical, petrochemical, and marine components; submarine hulls; armor plate; and medical applications, such as orthopedic implants (Table 6.10). Titanium alloys are available for service at 550°C for long periods of time, and service at up to 750°C for shorter periods.

Unalloyed titanium, known as *commercially pure titanium*, has excellent corrosion resistance for applications where strength considerations are secondary. Aluminum, vanadium, molybdenum, manganese, and other alloying elements impart special properties, such as improved workability, strength, and hardenability.

The properties and manufacturing characteristics of titanium alloys are extremely sensitive to small variations in both alloying and residual elements. The control of composition and processing are therefore important, especially for the prevention of surface contamination by hydrogen, oxygen, or nitrogen during processing. These elements cause embrittlement (Section 1.5.2) of titanium and, consequently, reduce toughness and ductility.

The body-centered cubic structure of titanium (*beta-titanium*) is above 880°C and is ductile, whereas its hexagonal close-packed structure (*alpha-titanium*) is somewhat brittle and is very sensitive to stress corrosion. A variety of other structures (alpha, near-alpha, alpha–beta, and beta) can be obtained by alloying

Table 6.10: Properties and Typical Applications of Selected Wrought Titanium Alloys at Various Temperatures.

UNS number	Nominal composition (%)	Condition	Temperature (°C)	Ultimate tensile strength (MPa)	Yield strength (MPa)	Elongation (%)	Reduction in area (%)
R50250	99.5 Ti	Annealed	25	330	240	30	55
			300	150	95		
R54520	5 Al, 2.5 Sn	Annealed	25	860	810	16	40
			300	565	450		
R56400	6 Al, 4 V	Annealed	25	1000	925	14	30
			300	725	650		
		Solution + age	25	1175	1100	10	20
			300	980	900		
R58010	13 V, 11 Cr, 3 Al	Solution + age	25	1275	1210	8	–
			425	1100	830		

and heat treating, so that the properties can be optimized for specific applications. **Titanium aluminide intermetallics** (TiAl and Ti_3Al; see Section 4.2.2) have higher stiffness and lower density than conventional titanium alloys, and can withstand higher temperatures.

By far, the most widely used alloy of titanium is titanium—6% aluminum and 4% vanadium, or Ti-6-4; it is for aerospace structural applications and medical implants, because of its high strength, good fatigue and corrosion resistance, and (within the body) for its high biocompatibility.

Production. Ores containing titanium are first reduced to titanium tetrachloride in an arc furnace, then converted to titanium chloride in a chlorine atmosphere. The compound is reduced further to titanium metal by distillation and leaching (dissolving). This sequence forms *sponge titanium*, which is then pressed into billets, melted, and poured into ingots to be later processed into various shapes. The complexity of these multistep thermochemical operations (the *Kroll process*) adds considerably to the cost of titanium.

6.8 Refractory Metals and Alloys

There are four **refractory metals**: molybdenum, niobium, tungsten, and tantalum; they are called *refractory* because of their high melting points; they are also important alloying elements in steels and superalloys. More than most other metals and alloys, refractory metals maintain their strength at elevated temperatures, and thus are of great importance in rocket engines, gas turbines, and various other aerospace applications; in the electronic, nuclear-power, and chemical industries; and as tool and die materials. The temperature range for some of these applications is on the order of 1100 to 2200°C, where strength and oxidation are of major concern.

6.8.1 Molybdenum

Molybdenum (Mo) is a silvery white metal, with a high melting point, high modulus of elasticity, good resistance to thermal shock, and good electrical and thermal conductivity. It is used in higher amounts than any other refractory metal, in applications such as solid-propellant rockets, jet engines, honeycomb structures, electronic components, heating elements, and dies for die casting. The principal alloying elements for molybdenum are titanium and zirconium. Molybdenum is itself also an important alloying element in cast and wrought alloy steels and in heat-resistant alloys, imparting strength, toughness, and corrosion resistance. A major limitation of molybdenum alloys is their low resistance to oxidation at temperatures above 500°C, necessitating the need for protective coatings.

Production. The main source of molybdenum is the mineral *molybdenite* (molybdenum disulfide). The ore is first processed whereby the molybdenum is concentrated; it is then chemically reduced, first with oxygen and then with hydrogen. Powder-metallurgy techniques (Chapter 17) also are used to produce ingots, for further processing into various shapes.

6.8.2 Niobium (Columbium)

Niobium (Nb, for niobium, after Niobe, the daughter of the mythical Greek king Tantalus) was first identified in 1801; it is also called **columbium** (after its source mineral, *columbite*). Niobium possesses good ductility and formability, and has higher oxidation resistance than other refractory metals. With various alloying elements, niobium alloys can be produced with moderate strength and good fabrication characteristics. These alloys are generally used in rockets and missiles and in nuclear, chemical, and superconductor applications. Niobium is also an alloying element in various alloys and superalloys. The metal is processed from ores by reduction and refinement, and from powder by first melting and shaping into ingots.

6.8.3 Tungsten

Tungsten (W, for *wolfram*, its European name, and from its source mineral, *wolframite*; in Swedish, *tung* means heavy and *sten* means stone) is the most abundant of all the refractory metals. Tungsten has the highest melting point of any metal (3410°C) and is notable for its high strength at elevated temperatures. However, it has high density, and hence used for balancing weights and counterbalances in mechanical systems, including self-winding watches. It is brittle at low temperatures and has poor resistance to oxidation. As an alloying element, tungsten imparts elevated-temperature strength and hardness to steels.

Tungsten alloys are used for applications involving temperatures above 1650°C, such as nozzle throat liners in missiles and in the hottest parts of jet and rocket engines, circuit breakers, welding electrodes, tooling for electrical-discharge machining, and spark-plug electrodes. Tungsten carbide, with cobalt as a binder for the carbide particles, is one of the most important tool and die materials (Chapter 22). Tungsten is processed from ore concentrates by chemical decomposition, then reduced, and further processed by powder-metallurgy techniques in a hydrogen atmosphere.

6.8.4 Tantalum

Tantalum (Ta, after the mythical Greek king, Tantalus) is characterized by its high melting point (3000°C), high density, good ductility, and resistance to corrosion. However, it has poor chemical resistance at temperatures above 150°C. Tantalum is used extensively in electrolytic capacitors and in various components in the electrical, electronic, medical product (see Section 6.16), and chemical industries. It also is used for thermal applications, such as in furnaces and acid-resistant heat exchangers, and as an alloying element. A variety of tantalum-based alloys are available for use in missiles and aircraft. It is processed by techniques similar to those used for processing niobium.

6.9 Beryllium

Steel gray in color, **beryllium** (Be, from the ore *beryl*) has a high strength-to-weight ratio. Unalloyed beryllium is used in rocket nozzles, space and missile structures, aircraft disc brakes, and precision instruments and mirrors. It is also used in nuclear and X-ray applications because of its low neutron absorption. Beryllium is also an alloying element, and its alloys of copper and nickel are used in various applications, including springs (*beryllium copper*), electrical contacts, and in non-sparking tools for use in explosive environments, such as mines and metal-powder production. Beryllium and its oxide are toxic.

6.10 Zirconium

Zirconium (Zr) is silvery in appearance; it has good strength and ductility at elevated temperatures and has good corrosion resistance because of an adherent oxide film. Zirconium is used in electronic components and in nuclear-power reactor applications because of its low neutron absorption.

6.11 Lithium

Lithium, from the Greek *lithos* or stone, is the lightest metal, with a silvery-white appearance. It is widely used in rechargeable batteries, with applications ranging from computers and cell phones to automobiles. It is also used to form alloys with aluminum and magnesium for lightweighting applications. *Lithium carbonate* is added to glass to improve its mechanical properties. *Lithium stearate* is an effective lubricating grease.

Lithium is produced from ocean water brines or by crushing ore and converting the lithium compounds to lithium chloride, which is then melted and a current passed through the material to produce lithium and chlorine gas.

6.12 Low-melting Alloys

Low-melting alloys are so named because of their relatively low melting temperatures. The major metals in this category are lead, zinc, tin, and their alloys.

6.12.1 Lead

Lead (Pb, after *plumbum*, the root of the word plumber) has the properties of high density, resistance to corrosion (by virtue of the stable lead-oxide layer that forms to protect the surface), softness, low strength, ductility, and good workability. Alloying it with various elements (such as antimony and tin) enhances its desirable properties, making it suitable for piping, collapsible tubing, bearing alloys (*Babbitt*), cable sheathing, foil (as thin as 0.01 mm), roofing, and lead–acid storage batteries. Lead also is used for damping vibrations, radiation shielding against X-rays, ammunition, and in the chemical industry.

The oldest known lead artifacts were made in about 3000 B.C. Lead pipes made by the Romans and installed in the Roman baths in Bath, England, two millennia ago, are still in use. An additional use of lead is as a solid lubricant for hot metal-forming operations. Because of its toxicity, however, major efforts continue to be made to replace lead with other elements, such as *lead-free solders* (Section 32.3.1). The most important mineral source of lead is *galena* (PbS); it is mined, smelted, and refined by chemical treatments.

6.12.2 Zinc

Zinc (Zn) is bluish white in color and is the metal that is fourth most utilized industrially, after iron, aluminum, and copper. It has three major uses: (1) for galvanizing iron, steel sheet, and wire, (2) as an alloying element in other metals, and (3) as a metal for castings. In **galvanizing**, zinc serves as an anode and protects steel (cathode) from corrosive attack should the coating be scratched or punctured. Zinc is also used as an alloying element; brass, for example, is an alloy of copper and zinc. In zinc-based alloys, the major alloying elements are aluminum, copper, and magnesium; they impart strength and provide dimensional control during casting of the metal.

Zinc-based alloys are used extensively in die casting (Section 11.4.5), for making such products as fuel pumps and grills for automobiles, components for household appliances such as vacuum cleaners and washing machines, kitchen equipment, machinery parts, and photoengraving equipment. Another use for zinc is in superplastic alloys (Section 2.2.7). A very-fine grained 78% Zn–22% Al sheet is a common example of a superplastic zinc alloy that can be formed by methods used for forming plastics or metals (Part III).

Production. The principal mineral source for zinc is zinc sulfide, also called *zincblende*. The ore is first roasted in air and converted to zinc oxide. It is then reduced to zinc either electrolytically (using sulfuric acid) or by heating it in a furnace with coal, which causes the molten zinc to separate.

6.12.3 Tin

Although used in small amounts, as compared to iron, aluminum, or copper, **tin** (Sn, from the Latin *stannum*) is an important metal. A silver-white, lustrous metal, its most extensive use is as a protective coating on steel sheets (*tin plates*) used in making containers (*tin cans*), for food and for various other products. The low shear strength of the tin coatings on steel sheet improves its deep drawability (Section 16.7.1). Unlike galvanized steels, if this coating is punctured or destroyed, the steel corrodes because the tin is cathodic.

Unalloyed tin is used in such applications as lining for water distillation plants and as a molten layer of metal in the production of float glass plate (Section 18.3.1). Tin-based alloys (also called **white metals**) generally contain copper, antimony, and lead. These alloying elements impart hardness, strength, and corrosion resistance. Tin itself is an alloying element for dental alloys and for bronze (copper–tin alloy), titanium, and zirconium alloys. Tin–lead alloys are common soldering materials (Section 32.3), with a wide range of compositions and melting points.

Because of their low friction coefficients (which result from low shear strength and low adhesion), some tin alloys are used as journal-bearing materials. Known as **babbitts** (after I. Babbitt, 1799–1862), these alloys contain tin, copper, and antimony. **Pewter**, an alloy of tin, copper, and antimony, is used for tableware, hollowware, and decorative artifacts. Tin alloys are also used in making organ pipes. The most important tin mineral is *cassiterite* (a low grade tin oxide). The ore is first mined, then concentrated using various techniques, smelted, refined, and cast into ingots for further processing.

6.13 Precious Metals

The most important precious (costly) metals, also called **noble metals**, are the following:

- **Gold** (Au, from the Latin *aurum*) is soft and ductile, and has good corrosion resistance at any temperature. Typical applications include jewelry, coinage, reflectors, gold leaf for decorative purposes, dental work, electroplating, and electrical contacts and terminals.

- **Silver** (Ag, from the Latin *argentum*) is ductile and has the highest electrical and thermal conductivity of any metal (see Table 3.2). However, it develops an oxide film that adversely affects its surface characteristics and appearance. Typical applications for silver include tableware, jewelry, coinage, electroplating, solders, bearing linings, and food and chemical equipment. *Sterling silver* is an alloy of silver and 7.5% copper.

- **Platinum** (Pt) is a soft, ductile, grayish-white metal that has good corrosion resistance, even at elevated temperatures. Platinum alloys are used as electrical contacts; for spark-plug electrodes; as catalysts for automobile pollution-control devices; in filaments and nozzles; in dies for extruding glass fibers (Section 18.3.4); in thermocouples; and in jewelry and dental work.

6.14 Shape-memory Alloys (Smart Materials)

Shape-memory alloys are unique in that, after being plastically deformed at room temperature into various shapes, they return to their original shape upon heating. For example, a piece of straight wire made of such a material, can be wound into the shape of a helical spring; when heated, the spring uncoils and returns to its original straight shape. Shape-memory alloys can be used to generate motion and/or force in temperature-sensitive actuators. The behavior of these alloys, also called **smart materials**, can be reversible; that is, the shape can switch back and forth repeatedly upon application and removal of heat.

A typical shape-memory alloy is 55% Ni–45% Ti (*Nitinol*); other alloys are copper–aluminum–nickel, copper–zinc–aluminum, iron–manganese–silicon, and titanium–nickel–hafnium. Shape-memory alloys generally also have such properties as good ductility, corrosion resistance, and high electrical conductivity.

Applications of shape-memory alloys include sensors, relays, pumps, switches, connectors, clamps, fasteners, seals, and stents for blocked arteries. As an example, a nickel–titanium valve has been made to protect people from being scalded in sinks, tubs, and showers. It is installed directly into the piping system and brings the water flow down to a trickle within three seconds after the water temperature reaches 47°C. More recent developments include thin-film shape-memory alloys deposited on polished silicon substrates for use in microelectromechanical (MEMS) devices (Chapter 29).

6.15 Amorphous Alloys (Metallic Glasses)

A class of metal alloys that, unlike metals, do not have a long-range crystalline structure is called *amorphous alloys*. They have no grain boundaries, and their atoms are packed randomly and tightly. The amorphous structure was first obtained in the late 1960s by **rapid solidification** of a molten alloy (Section 11.6). Because their structure resembles that of glasses, these alloys are also called *metallic glasses*.

Amorphous alloys typically contain iron, nickel, and chromium, alloyed with carbon, phosphorus, boron, aluminum, and silicon. They are available as wire, ribbon, strip, and powder: One application is for faceplate inserts on golf-club heads; the alloy has a composition of zirconium, beryllium, copper, titanium, and nickel and is made by die casting. Another application is in hollow aluminum baseball bats, coated with a composite of amorphous metal by thermal spraying, and is said to improve the performance of the bat.

Amorphous alloys exhibit excellent corrosion resistance, good ductility, high strength, and very low magnetic hysteresis (utilized in the production of magnetic steel cores for transformers, generators, motors, lamp ballasts, magnetic amplifiers, and linear accelerators). They have low magnetic hysteresis loss, providing greatly improved efficiency; however, fabrication costs are significant. Amorphous steels have been demonstrated to have strengths twice those of high-strength steels, and have potential applications in large structures; however, they are presently cost prohibitive. A major application for the superalloys of rapidly solidified powders is the consolidation into near-net shapes for parts used in aerospace engines.

6.16 Metal Foams and Metamaterials

Metal foams are structures where the metal consists of only 5–20% of the structure's volume, as shown in Fig. 6.4. Usually made of aluminum alloys (but also of titanium, tantalum, and others), metal foams can be produced by blowing air into molten metal and tapping the froth that forms at the surface. The froth then solidifies into a foam.

Other approaches to producing metal foam include (a) chemical vapor deposition (Section 34.6.2) onto a carbon foam lattice, (b) depositing metal powders from a slurry onto a polymer foam lattice, followed by sintering (Section 17.4) to fuse the metals and burn off the polymer, (c) doping molten or powder metals (Chapter 17) with titanium hydride (TiH_2), which then releases hydrogen gas at the elevated casting or sintering temperatures, and (d) pouring molten metal into a porous salt and, upon cooling, leaching out the salt with acid.

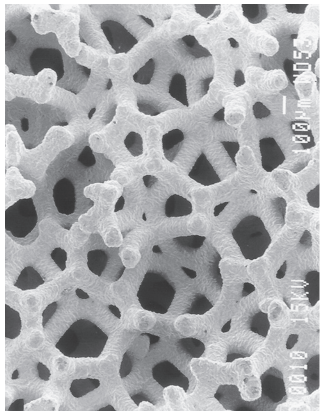


Figure 6.4: Structure of a metal foam used in orthopedic implants to encourage bone ingrowth. *Source:* Courtesy of Zimmer, Inc.

Metal foams have unique combinations of strength-to-density and stiffness-to-density ratios, although these ratios are not as high as the base metals themselves. However, metal foams are very lightweight and thus are attractive materials, especially for aerospace applications. Because of their porosity, other applications of metal foams are filters and orthopedic implants. More recent developments include nickel–manganese–gallium metal foams with shape-memory characteristics (Section 6.14).

Metamaterials are similar to metal foams, in that they have a very high porosity; the difference is that metamaterials are *designed* to achieve certain mechanical, thermal, or electrical properties. The metamaterial shown in Fig. 6.5 consists of a number of struts, just as with a metal foam. However, the struts do not have a random orientation or location; as is common, this metamaterial has a unit cell that is repeated to fill a volume. By changing the geometry, it is possible to obtain desired characteristics such as high strength-to-weight or stiffness-to-weight ratios, or zero (or negative) Poisson's ratio. An application for metamaterials is described in Case Study 20.2.

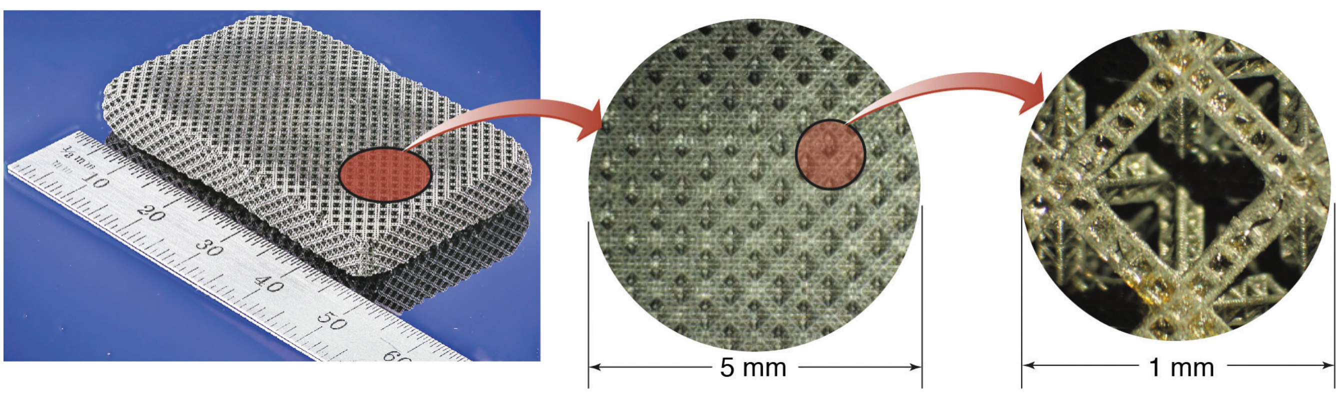


Figure 6.5: Nickel alloy hierarchical metamaterial and critical features across two orders of magnitude in length scale. (a) Large-area, high-resolution additive manufacturing of hierarchical metamaterials. (b)–(c) Optical microscope images of bulk hierarchical lattice material with a network of hierarchical stretch-dominated octet unit cells. *Source:* Xiaoyu (Rayne) Zheng, Virginia Tech.

Metamaterials have mainly been manufactured by additive manufacturing (Chapter 20) because of the ability to design the structure; an alternative approach uses a additive manufactured structure that is then investment cast to produce a metamaterial. The slender nature of the struts in Fig. 6.5 would not be suitable for investment casting, however, and the struts cell dimensions of such metamaterials are much larger.

6.17 Rare Earth Metals

Rare earth metals. Rare earth metals are so named because they are generally difficult to mine and are available only in small quantities. However, they are often more common in the Earth's crust than **precious metals**, such as gold and platinum. Because of their unique magnetic, luminescent, and electrical properties, rare earth metals are essential in modern technologies (see below). Among the more important rare earth metals are:

1. **Yttrium**, which is used in compact fluorescent lamps, light-emitting diodes, flat-panel monitors, laser technology, and superconductor applications.
2. **Lanthanum**, which is used in catalytic converters and anode materials in high-performance batteries. Most hybrid automobiles depend on lanthanum anodes in their batteries; for example, it is estimated that each Toyota Prius uses 10–15 kg of lanthanum.
3. **Dysprosium**, which is used in the production of lasers and commercial lighting, as well as in dosimeters to measure radiation exposure. When exposed to radiation, dysprosium emits light, which can be measured and correlated to radiation strength; it is also highly *magnetostrictive*, deforming under a magnetic field, thus making it useful for transducers and resonators.
4. **Neodymium** More common than cobalt, nickel, and copper, neodymium is the second most common rare earth metal (after cerium). Its alloy, $Nd_2Fe_{14}B$, produces the strongest permanent magnets known. As such, it is used where small but powerful magnets are required, such as in-ear headphones and microphones.
5. **Cerium**, which is mainly used in catalytic converters in automobiles, for the oxidation of carbon monoxide and nitrous oxide; it is also used in glass manufacture, permanent magnets, fuel cells, and in polishing optical components.
6. **Samarium**, which is used in compounds of cobalt (usually $SmCo_5$ or $SmCo_{17}$) and is the second strongest permanent magnet known, next to neodymium magnets. However, samarium-cobalt magnets have better stability and can be used to temperatures as high as 700°C, whereas neodymium magnets are limited to 300°C or so.
7. **Terbium**, which is used as a dopant in solid-state electronic devices and in various sensors. It has the highest magnetostriction (Section 3.7) of any alloy.

Several other rare-earth metals have been used in industrial applications, usually for magnets (praseodymium, holmium), computer and portable electronic device displays (scandium, europium), and in radiation shielding (gadolinium, erbium) or in generation of radiation (thulium).

Summary

- Nonferrous metals and alloys include a very broad range of materials. The most common are aluminum, magnesium, and copper and their alloys, with a wide range of applications. For high temperature service, nonferrous metals include nickel, titanium, refractory alloys (molybdenum, niobium, tungsten, tantalum), and superalloys. Other nonferrous metal categories include low-melting alloys (lead, zinc, tin) and precious metals (gold, silver, platinum).

- Nonferrous alloys have a wide variety of desirable properties, such as strength, toughness, hardness, and ductility; resistance to high temperature, creep, and oxidation; a wide range of physical, thermal, and chemical properties; and high strength-to-weight and stiffness-to-weight ratios (particularly for aluminum and titanium). Nonferrous alloys can be heat treated to impart certain specific properties.
- Shape-memory alloys (smart materials) have unique properties, with numerous applications in a variety of products as well as in manufacturing operations.
- Amorphous alloys (metallic glasses) have properties that are superior to other materials; available in various forms, they have numerous applications.
- Metal foams are very lightweight and thus are attractive for aerospace as well as various other applications.
- Rare earth metals are not actually rare, but are difficult to mine in large amounts. These metals have widespread applications, including permanent magnets and devices that exploit them, computer and portable electronic displays, batteries and fuel cells, and radiation shielding and generation.
- As with all materials, the selection of a nonferrous material for a particular application requires a careful consideration of several factors, including design and service requirements, long-term effects, chemical affinity to other materials, environmental attack, and cost.

Key Terms

Amorphous alloys
Babbitts
Brass
Bronze
Galvanizing
Low-melting alloys
Metal foam
Metallic glasses
Nonferrous
Pewter
Precious metals
Pyrometallurgy
Refractory metals
Shape-memory alloys
Smart materials
Smelting
Superalloys
Temper designation

Bibliography

Ashby, M.F., **Materials Selection in Mechanical Design**, 4th ed., Pergamon, 2010.
ASM Handbook, Vol. 2: **Properties and Selection: Nonferrous Alloys and Special-Purpose Materials**, ASM International, 1990.
ASM Specialty Handbook: **Aluminum and Aluminum Alloys**, ASM International, 1993.
ASM Specialty Handbook: **Copper and Copper Alloys**, ASM International, 2001.
ASM Specialty Handbook: **Heat-Resistant Materials**, ASM International, 1997.
ASM Specialty Handbook: **Magnesium and Magnesium Alloys**, ASM International, 1999.
ASM Specialty Handbook: **Nickel, Cobalt, and Their Alloys**, ASM International, 2000.
Donachie, M.J. (ed.), **Titanium: A Technical Guide**, 2nd ed., ASM International, 2000.
Donachie, M.J., and Donachie, S.J., **Superalloys: A Technical Guide**, 2nd ed., ASM International, 2002.
Geddes, B., Leon, H., and Huang, X., **Superalloys: Alloying and Performance**, ASM International, 2010.
Kaufman, J.G., **Introduction to Aluminum Alloys and Tempers**, ASM International, 2000.
Lexcellent, C., **Shape-memory Alloys Handbook**, Wiley, 2013.

Lutjering, G., and Williams, J.C., **Titanium**, 2nd ed., Springer, 2007.
Reed, R.C,. **The Superalloys: Fundamentals and Applications**, Cambridge, 2006.
Russew, K., and Stojanova, L. **Glassy Metals**, Springer, 2016.
Schwartz, M., **Smart Materials**, CRC Press, 2008.
Solanski, K.N., and Orlov, D. (eds.), **Magnesium Technology**, Springer, 2017.
Totten, G. and MacKenzie, D.S. (eds.), **Handbook of Aluminum, Vol. 1: Physical Metallurgy and Processes**, CRC Press, 2003.

Review Questions

6.1. Given the abundance of aluminum in the earth's crust, explain why it is more expensive than steel.

6.2. Why is magnesium often used as a structural material in power hand tools? Why are its alloys used instead of pure magnesium?

6.3. What are the major uses of copper? What are the alloying elements in brass and bronze, respectively?

6.4. What are superalloys? Why are they so named?

6.5. What properties of titanium make it attractive for use in race-car and jet-engine components? Why is titanium not used widely for engine components in passenger cars?

6.6. Which properties of each of the major refractory metals define their most useful applications?

6.7. What are the main industrial uses of lithium?

6.8. What are metallic glasses? Why is the word "glass" used for these materials?

6.9. What is the composition of (a) babbitts, (b) pewter, and (c) sterling silver?

6.10. Name the materials described in this chapter that have the highest (a) density, (b) electrical conductivity, (c) thermal conductivity, (d) strength, and (e) cost.

6.11. What are the major uses of gold and silver, other than in jewelry?

6.12. Describe the advantages to using zinc as a coating for steel.

6.13. What are nanomaterials? Why are they being developed?

6.14. Why are aircraft fuselages made of aluminum alloys, even though magnesium is a lighter metal? Why isn't lithium used for such applications?

6.15. How is metal foam produced?

6.16. What metals have the lowest melting points? What applications for these metals take advantage of their low melting points?

6.17. What are the main applications of rare earth metals?

Qualitative Problems

6.18. Explain why cooking utensils generally are made of stainless steels, aluminum, or copper.

6.19. Would it be advantageous to plot the data in Table 6.1 in terms of cost per unit weight rather than cost per unit volume? Explain and give some examples.

6.20. Compare the contents of Table 6.3 with those in various other tables and data on materials in this book, and then comment on which of the two hardening processes (heat treating and work hardening) is more effective in improving the strength of aluminum alloys.

6.21. What factors other than mechanical strength should be considered in selecting metals and alloys for high-temperature applications? Explain.

6.22. Assume that, for geopolitical reasons, the price of copper increases rapidly. Name two metals with similar mechanical and physical properties that can be substituted for copper. Comment on your selection and any observations you make.

6.23. If aircraft, such as a Boeing 757, are made of 79% aluminum, why are automobiles made predominantly of steel?

6.24. Portable (notebook) computers and digital cameras can have their housing made of magnesium. Why?

6.25. Most household wiring is made of copper wire. By contrast, grounding wire leading to satellite dishes and the like is made of aluminum. Explain the reason.

6.26. The example in this chapter showed the benefits of making cars from aluminum alloys. However, the average amount of steel in cars has increased in the past decade. List reasons to explain these two observations.

6.27. If tungsten is the highest melting-point metal, why are no high temperature parts in Fig. 6.1 made from tungsten?

Quantitative Problems

6.28. A simply supported rectangular beam is 50 mm wide and 2 m long, and it is subjected to a vertical load of 50 kg at its center. Assume that this beam could be made of any of the materials listed in Table 6.1. Select three different materials, and for each, calculate the beam height that would cause each beam to have the same maximum deflection. Calculate the ratio of the cost for each of the three beams.

6.29. Obtain a few aluminum beverage cans, cut them, and measure their wall thicknesses. Using data in this chapter and simple formulas for thin-walled, closed-end pressure vessels, calculate the maximum internal pressure these cans can withstand before yielding. (Assume that the can is a thin-walled, closed-end, internally pressurized vessel.)

6.30. Beverage cans usually are stacked on top of each other in stores. Use the information from Problem 6.24, and, referring to textbooks on the mechanics of solids, estimate the crushing load each of these cans can withstand.

6.31. Using strength and density data, determine the minimum weight of a 1 m-long tension member that must support 3000 N if it is manufactured from (a) 3003-O aluminum, (b) 5052-H34 aluminum, (c) AZ31B-F magnesium, (d) any brass alloy, and (e) any bronze alloy.

6.32. Plot the following for the materials described in this chapter: (a) yield strength vs. density, (b) modulus of elasticity vs. strength, (c) modulus of elasticity vs. relative cost, and (d) electrical conductivity vs. density.

Synthesis, Design, and Projects

6.33. Because of the number of processes involved in making metals, the cost of raw materials depends on the condition (hot or cold rolled), shape (plate, sheet, bar, tubing), and size of the metals. Make a survey of the technical literature, obtain price lists or get in touch with suppliers, and prepare a list indicating the cost per 100 kg of the nonferrous materials described in this chapter, available in different conditions, shapes, and sizes.

6.34. The materials described in this chapter have numerous applications. Make a survey of the available literature in the bibliography, and prepare a list of several specific parts or components and applications, indicating the types of materials used.

6.35. Name products that would not have been developed to their advanced stages (as we find them today) if alloys having high strength, high corrosion resistance, and high creep resistance (all at elevated temperatures) had not been developed.

6.36. Assume that you are the technical sales manager of a company that produces nonferrous metals. Choose any one of the metals and alloys described in this chapter, and prepare a brochure, including some illustrations, for use as sales literature by your staff in their contact with potential customers.

6.37. Give some applications for (a) amorphous metals, (b) precious metals, (c) low-melting alloys, and (d) nanomaterials.

6.38. Describe the advantages of making products with multilayer materials. (For example, aluminum bonded to the bottom of stainless-steel pots.)

6.39. In the text, magnesium was described as the lightest engineering metal. Is it also the lightest metal? Explain.

6.40. Review the technical literature and the Internet and summarize the rare earth metals, their sources, and their main applications.

6.41. Review the technical literature, and write a detailed description of how magnesium is produced from sea water.

6.42. If you were to design an implant for use in the human body, what materials would you exclude? Which metals are possible for such applications? Of these, list three that you feel are best.

6.43. Perform an Internet search and obtain a typical design for an "ear bud," intended to play sound in a device inserted into an ear. What role do rare earth elements play in ear buds? Estimate the required size of ear buds if rare earth elements were unavailable.

Chapter 7

Polymers: Structure, General Properties, and Applications

- Polymers display a very wide range of properties and have several advantages over metallic materials, including low cost and ease of manufacturing; they continue to be among the most commonly used materials.
- This chapter first describes the structure of polymers, the polymerization process, crystallinity, and the glass-transition temperature.
- Mechanical properties and how they are affected by temperature and deformation rate are then described.
- Two basic types of polymers are thermoplastics and thermosets. Thermoplastics follow a basic manufacturing procedure of heating them until they soften or melt, and then shaping them into the desired product. Thermosets involve precursors that are formed to a desired shape and set through polymerization or cross-linking between polymer chains.

- The chapter then describes the properties and uses of elastomers.
- The general properties, typical applications, advantages, and limitations of polymers are all described throughout the chapter, with several specific examples.

7.1 Introduction

The word **plastics** was first used as a noun in 1909, and is commonly interchanged as a synonym for **polymers**, a term first used in 1866. Plastics are unique in that they have extremely large molecules (*macromolecules* or *giant molecules*). Consumer and industrial products made of plastics include food and beverage containers, packaging, signs, housewares, housings for computers and monitors, textiles (clothing), medical devices, foams, paints, safety shields, toys, appliances, lenses, gears, electronic and electrical products, and automobile and aircraft bodies and numerous components.

Because of their unique and diverse properties, polymers increasingly have replaced metallic components, reflecting the advantages of polymers in terms of the following characteristics:

- Relatively low cost (Table 6.1) and ease of manufacture
- Resistance to chemicals
- Low electrical and thermal conductivity
- Low density
- High strength-to-weight ratio, particularly when reinforced
- Noise reduction
- Wide choice of colors and transparencies
- Complex design possibilities and ease of manufacturing

Characteristics of polymers that may or may not be desirable are low strength and stiffness (Table 7.1), high coefficient of thermal expansion, low useful-temperature range, and lower dimensional stability in service over a period of time.

The word *plastic* is from the Greek word *plastikos*, meaning capable of being molded and shaped. Plastics can be formed, cast, machined, and joined into various shapes with relative ease. Little or no additional surface finishing is required; this characteristic provides an important advantage over metals. Plastics are available as film, sheet, plate, rod, and tubing of various cross sections.

An outline of the basic process for making synthetic polymers is given in Fig. 7.1. In polyethylene, only carbon and hydrogen atoms are present, but other polymer compounds can be made by including chlorine, fluorine, sulfur, silicon, nitrogen, and oxygen. As a result, an extremely wide range of polymers, with an equally wide range of properties, has been developed.

7.2 The Structure of Polymers

The properties of polymers depend largely on the structures of individual polymer molecules, molecule shape and size, and the arrangement of molecules to form a polymer structure. Polymer molecules are characterized by their *very large size*, a feature that distinguishes them from most other organic chemical compositions. Polymers are **long-chain molecules** formed by *polymerization*, that is, by the linking and cross-linking of different monomers. A **monomer** is the basic building block of a polymer. The word mer

Table 7.1: Range of Mechanical Properties for Various Engineering Plastics at Room Temperature.

Material	Ultimate tensile strength (MPa)	Elastic modulus (GPa)	Elongation (%)	Poisson's ratio, ν
Thermoplastics:				
Acrylonitrile-butadiene-styrene (ABS)	28–55	1.4–2.8	75–5	0.35
ABS, reinforced	100	7.5	—	0.35
Acetal	55–70	1.4–3.5	75–25	0.35
Acetal, reinforced	135	10	—	0.35–0.40
Acrylic	40–75	1.4–3.5	50–5	0.37
Cellulosic	10–48	0.4–1.4	100–5	0.39
Fluorocarbon	7–48	0.7–2	300–100	0.46–0.48
Nylon	55–83	1.4–2.8	200–60	0.32–0.40
Nylon, reinforced	70–210	2–10	10–1	—
Polycarbonate	55–70	2.5–3	125–10	0.38
Polycarbonate, reinforced	110	6	6–4	—
Polyester	55	2	300–5	0.38
Polyester, reinforced	110–160	8.3–12	3–1	—
Polyethylene	7–40	0.1–1.4	1000–15	0.46
Polypropylene	20–35	0.7–1.2	500–10	0.43
Polypropylene, reinforced	40–100	3.5–6	4–2	—
Polystyrene	14–83	1.4–4	60–1	0.35
Polyvinyl chloride	7–55	0.014–4	450–40	0.40
Thermosets:				
Epoxy	35–140	3.5–17	10–1	0.30–0.35
Epoxy, reinforced	70–1400	21–52	4–2	—
Phenolic	28–70	2.8–21	2–0	0.41
Polyester, unsaturated	30	5–9	1–0	0.40
Elastomers:				
Chloroprene (neoprene)	15–25	1–2	100–500	0.5
Natural rubber	17–25	1.3	75–650	0.5
Silicone	5–8	1–5	100–1100	0.5
Styrene-butadiene	10–25	2–10	250–700	0.5
Urethane	20–30	2–10	300–450	0.5

(from the Greek word *meros*, meaning part) indicates the smallest repetitive unit, thus the term is similar to that of *unit cell* in crystal structures of metals (Section 1.3).

The word **polymer** means many mers, repeated hundreds or thousands of times in a chainlike structure. Most monomers are *organic materials*, in which carbon atoms are joined through *covalent* (electron sharing) bonds with other atoms (such as hydrogen, oxygen, nitrogen, fluorine, chlorine, silicon, and sulfur). An ethylene molecule (Fig. 7.2) is an example of a simple monomer, consisting of carbon and hydrogen atoms.

7.2.1 Polymerization

Monomers can be linked in repeating units to make longer and larger molecules by a chemical process called a **polymerization reaction**. Although there are several variations, two polymerization processes are important: condensation and addition polymerization.

In **condensation polymerization** (Fig. 7.3a), polymers are produced by the formation of bonds between two types of reacting mers. A characteristic of this reaction is that the reaction by-products (such

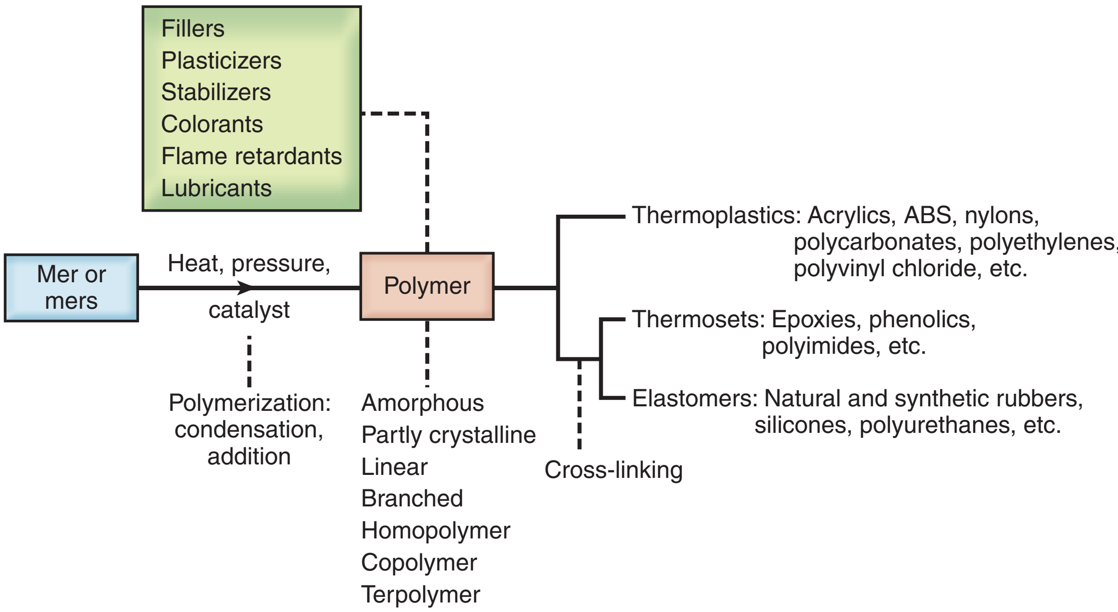


Figure 7.1: Outline of the topics described in Chapter 7.

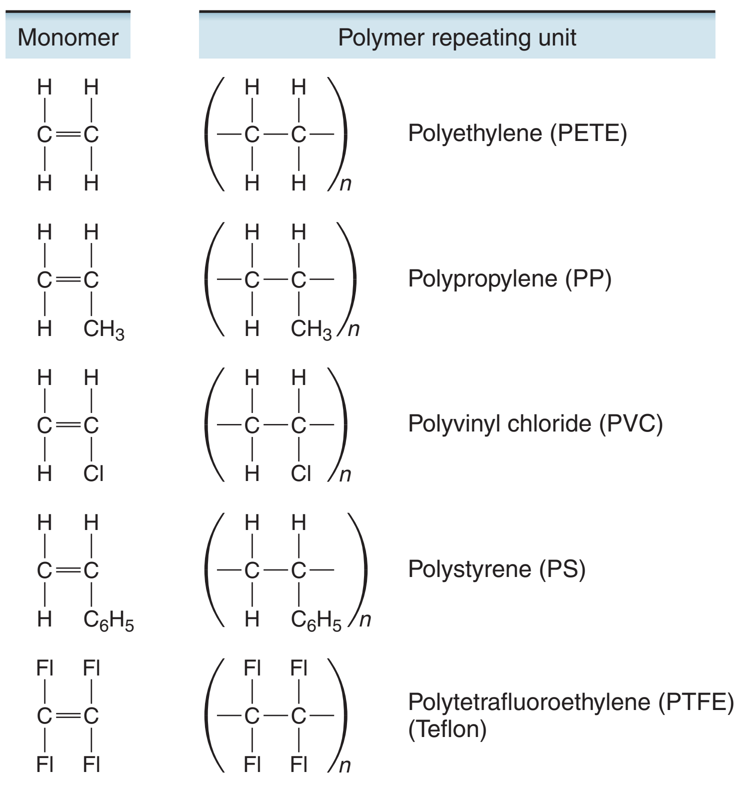


Figure 7.2: Molecular structure of various polymers. These are examples of the basic building blocks for plastics.

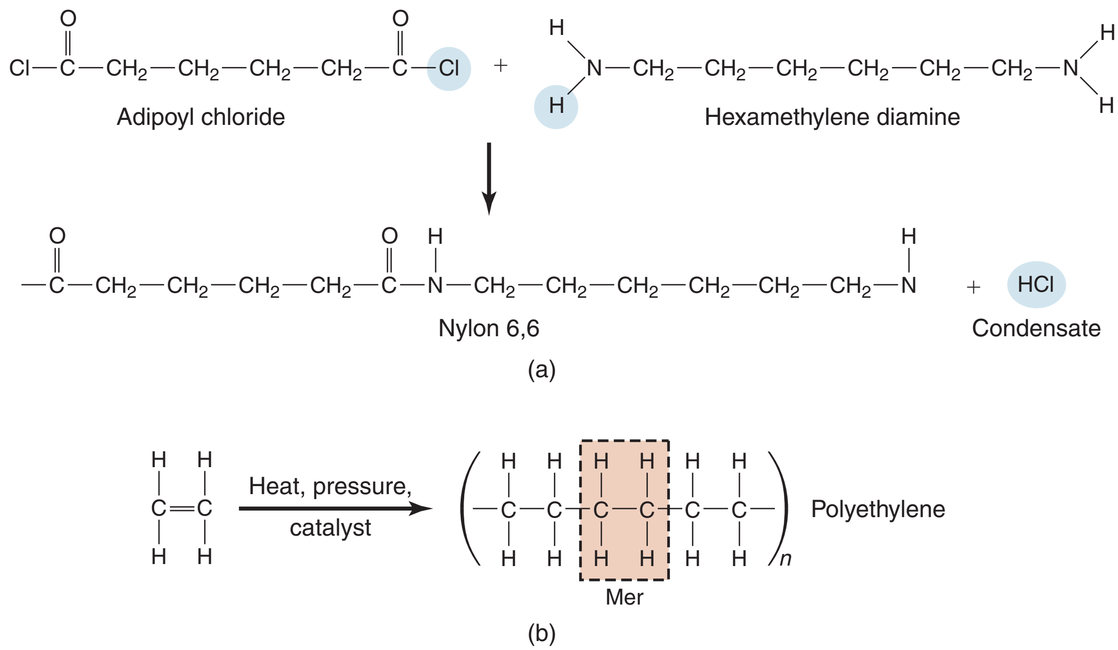


Figure 7.3: Examples of polymerization. (a) Condensation polymerization of nylon 6,6 and (b) addition polymerization of polyethylene molecules from ethylene mers.

as water) are condensed out (hence the word *condensation*). This process is also known as **step-growth** or **step-reaction polymerization**, because the polymer molecule grows step-by-step until all of one reactant is consumed.

In **addition polymerization**, also called **chain-growth** or **chain-reaction polymerization**, bonding takes place without reaction by-products, as shown in Fig. 7.3b. It is called *chain reaction* because of the high rate at which long molecules form simultaneously, usually within a few seconds, a rate much higher than that in condensation polymerization. In addition polymerization, an *initiator* is added to open the double bond between two carbon atoms, which then begins the linking process by adding several more monomers to a growing chain. For example, ethylene monomers (Fig. 7.3b) link to produce *polyethylene*; other examples of addition-formed polymers are given in Fig. 7.2.

Molecular Weight. The sum of the molecular weights of the mers in a representative chain is known as the *molecular weight* of the polymer. The higher the molecular weight of a given polymer, the greater is the average chain length. Most commercial polymers have a molecular weight between 10,000 and 10,000,000. Because polymerization is a random event, the polymer chains produced are not all of equal length, although the chain lengths fall into a traditional distribution curve (described in Section 36.7). The molecular weight of a polymer is determined on a statistical basis by averaging.

The spread of the molecular weights in a chain is called the **molecular weight distribution** (MWD). A polymer's molecular weight and its distribution have a major influence on its properties. For example, the tensile and the impact strength, the resistance to cracking, and the viscosity (in the molten state) of the polymer all increase with increasing molecular weight (Fig. 7.4).

Degree of Polymerization. It is convenient to express the size of a polymer chain in terms of the *degree of polymerization* (DP), defined as the ratio of the molecular weight of the polymer to the molecular weight of the repeating unit. For example, polyvinyl chloride (PVC) has a mer weight of 62.5; thus, with a molecular weight of 50,000 its DP is $50{,}000/62.5 = 800$. In terms of polymer processing (Chapter 19), the higher the DP,

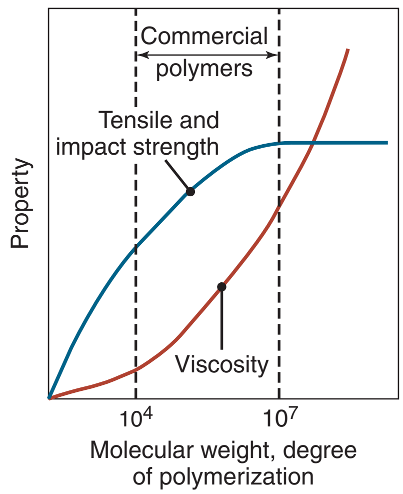


Figure 7.4: Effect of molecular weight and degree of polymerization on the strength and viscosity of polymers.

the higher is the polymer's *viscosity*, or its resistance to flow (Fig. 7.4). High viscosity adversely affects the ease of shaping the polymer, thus raising the cost of processing; however, high DP can result in stronger polymers.

Bonding. During polymerization, the monomers are linked together by **covalent bonds** (Section 1.2), forming a polymer chain; because of their strength, covalent bonds are also called **primary bonds**. The polymer chains are held together by **secondary bonds**, such as van der Waals bonds, hydrogen bonds, and ionic bonds (Section 1.2). Secondary bonds are weaker than primary bonds by one to two orders of magnitude.

In a given polymer, the increase in strength and viscosity with molecular weight is due, in part, to the fact that the longer the polymer chain, the greater is the energy needed to overcome the combined strength of the secondary bonds. For example, ethylene polymers having DPs of 1, 6, 35, 140, and 1350, at room temperature, are, respectively, in the form of gas, liquid, grease, wax, and hard plastic.

Linear Polymers. The chainlike polymers shown in Fig. 7.2 are called *linear polymers* because of their sequential structure (Fig. 7.5a); however, a linear molecule is not necessarily straight in structure. In addition to those shown in this figure, other linear polymers include polyamides (nylon 6,6) and polyvinyl fluoride. Generally, a polymer consists of more than one type of structure; thus, a linear polymer may contain some branched and some cross-linked chains. As a result , the polymer's properties are changed significantly.

Branched Polymers. The properties of a polymer depend not only on the type of monomers but also on their arrangement in the molecular structure. In *branched polymers* (Fig. 7.5b), side-branch chains are attached to the main chain during the synthesis of the polymer. Branching interferes with the relative movement of the molecular chains; as a result, their resistance to deformation and stress cracking is increased. The density of branched polymers is lower than that of linear-chain polymers, because the branches interfere with the packing efficiency of the polymer chains.

The behavior of branched polymers can be compared to that of linear-chain polymers, by making an analogy with a pile of tree branches (*branched polymers*) and a bundle of straight logs (*linear polymers*). Note that it is more difficult to move a branch within the pile of branches than to move a log within its bundle. The three-dimensional entanglements of branches make movements more difficult, a phenomenon akin to increased strength of the polymer.

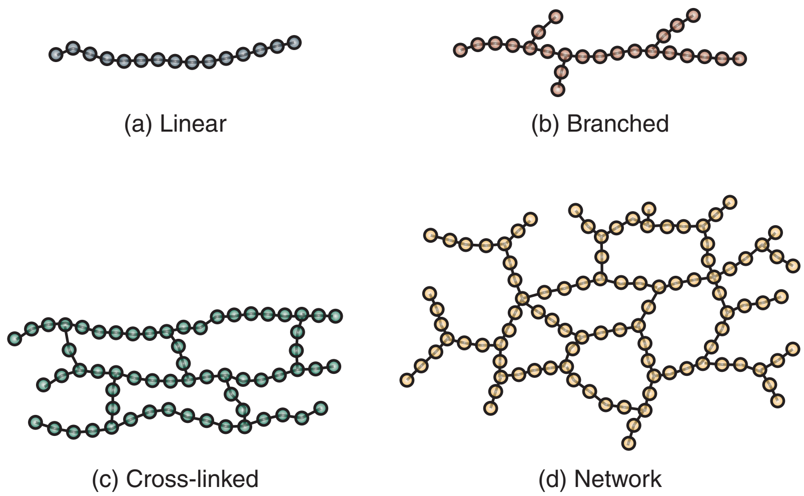


Figure 7.5: Schematic illustration of polymer chains. (a) Linear structure-thermoplastics such as acrylics, nylons, polyethylene, and polyvinyl chloride have linear structures. (b) Branched structure, such as in polyethylene. (c) Cross-linked structure—many rubbers, or elastomers, have this structure, and the vulcanization of rubber produces this structure. (d) Network structure, which is basically highly cross-linked. Examples are thermosetting plastics, such as epoxies and phenolics.

Cross-linked Polymers. Generally three-dimensional in structure, *cross-linked polymers* have adjacent chains linked by covalent bonds (Fig. 7.5c). Polymers with a cross-linked structure are called **thermosets** or **thermosetting plastics**, such as epoxies, phenolics, and silicones. Cross-linking has a major influence on the properties of polymers, typically imparting hardness, strength, stiffness, brittleness, and better dimensional stability (see Fig. 7.6) and the **vulcanization** of rubber (Section 7.9).

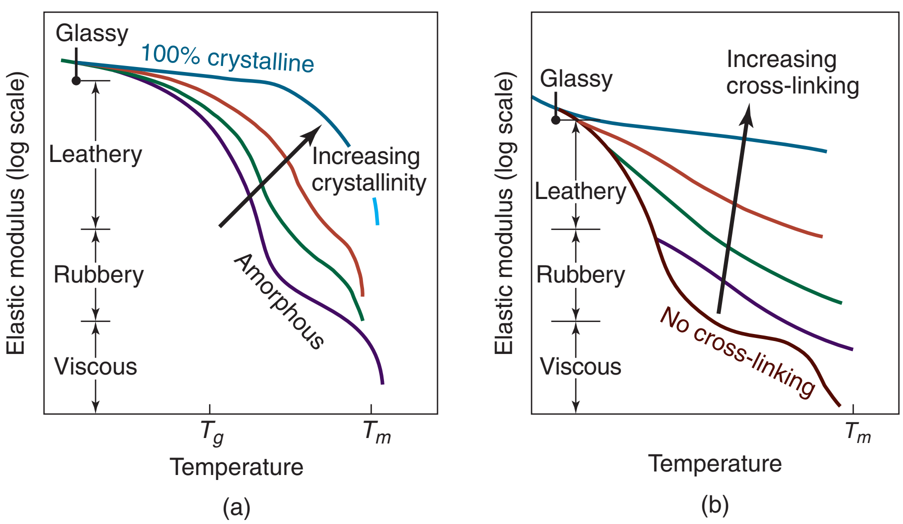


Figure 7.6: Behavior of polymers as a function of temperature and (a) degree of crystallinity and (b) cross-linking. The combined elastic and viscous behavior of polymers is known as viscoelasticity.

Network Polymers. These polymers consist of *spatial* (three-dimensional) networks consisting of three or more active covalent bonds (Fig. 7.5d); a highly cross-linked polymer also is considered a network polymer. Thermoplastic polymers that already have been formed or shaped can be cross-linked to induce higher strength, by subjecting them to high-energy radiation, such as ultraviolet light, X-rays, or electron beams. Excessive radiation can, however, cause degradation of the polymer.

Copolymers and **Terpolymers**. If the repeating units in a polymer chain are all of the same type, the molecule is called a *homopolymer*. As with solid-solution metal alloys (Section 4.2), two or three different types of monomers can be combined to develop specific properties and characteristics, such as improved strength, toughness, and formability of the polymer. *Copolymers* contain two types of polymers, such as styrene-butadiene, which is used widely for automobile tires. *Terpolymers* contain three types of polymers, such as acrylonitrile-butadiene-styrene (ABS), which is used for helmets, telephones, and refrigerator liners.

Case Study 7.1 Dental and Medical Bone Cement

Polymethylmethacrylate (PMMA) is an acrylic polymer, commonly used in dental and medical applications as an adhesive, often referred to as *bone cement*. There are several forms of PMMA, but the adhesive is one common form undergoing an addition-polymerization reaction. PMMA is delivered to the manufacturer in two parts: a powder and a liquid, which are hand-mixed. The liquid wets and partially dissolves the powder, resulting in a liquid with a viscosity similar to that of vegetable oil. The viscosity increases significantly until a doughy state is reached in about five minutes; it fully hardens in an additional five minutes.

The powder consists of high-molecular-weight poly[(methylmethacrylate)-costyrene] particles, about 50 μm in diameter, containing a small volume fraction of benzoyl peroxide. The liquid consists of methyl methacrylate (MMA) monomer, with a small amount of dissolved n, n dimethyl-p-toluidine (DMPT). When the liquid and the powder are mixed, the DMPT cleaves the benzoyl peroxide molecule into two parts, forming a catalyst with a free electron (also referred to as a free radical). This catalyst causes rapid growth of PMMA from the MMA mers, so that the final material is a composite of high-molecular-weight PMMA particles interconnected by PMMA chains. An illustration of a fully set bone cement is given in Fig. 7.7.

7.2.2 Crystallinity

Polymers such as PMMA, polycarbonate, and polystyrene are generally **amorphous**; that is, the polymer chains exist without long-range order (see also *amorphous alloys*, Section 6.15). The amorphous arrangement of polymer chains is often described as being like a bowl of spaghetti, or like worms in a bucket, all intertwined with each other. In some polymers, however, it is possible to impart some crystallinity and, thereby, modify their characteristics. This arrangement may be fostered either during the synthesis of the polymer or by deformation during its subsequent processing.

The crystalline regions in polymers are called **crystallites** (Fig. 7.8). They are formed when the long molecules arrange themselves in an orderly manner, similar to the folding of a fire hose in a cabinet or of facial tissues in a box. A partially crystalline (**semicrystalline**) polymer can be regarded as a two-phase material; one phase being crystalline and the other amorphous.

By controlling the chain structure and the rate of solidification during cooling, it is possible to impart different **degrees of crystallinity** to polymers, although never 100%. Crystallinity ranges from an almost complete crystal (up to about 95% by volume in the case of polyethylene) to slightly crystallized (and mostly amorphous) polymers. The degree of crystallinity is also affected by branching. A linear polymer can become highly crystalline; a highly branched polymer cannot, although it may develop some low level of crystallinity. It will never achieve high crystallite content, because the branches interfere with the alignment of the chains into a regular crystal array.

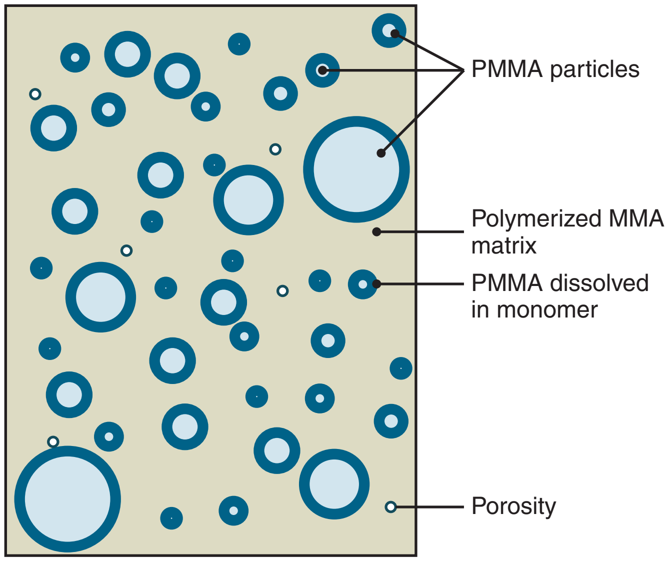


Figure 7.7: Schematic illustration of the microstructure of polymethylmethacrylate cement used in dental and medical applications.

Effects of Crystallinity. The mechanical and physical properties of polymers are greatly influenced by the degree of crystallinity. As crystallinity increases, polymers become stiffer, harder, less ductile, denser, less rubbery, and more resistant to solvents and heat (Fig. 7.6). The increase in density with increasing crystallinity is called *crystallization shrinkage*, and is caused by more efficient packing of the molecules in the crystal lattice. For example, the highly crystalline form of polyethylene, known as *high-density polyethylene* (HDPE), has a specific gravity in the range of 0.941–0.970 (80–95% crystalline). It is stronger, stiffer, tougher, and less ductile than low-density polyethylene (LDPE), which is about 60–70% crystalline and has a specific gravity in the range of 0.910–0.925.

Optical properties of polymers also are affected by the degree of crystallinity. The reflection of light from the boundaries between the crystalline and the amorphous regions in the polymer (Fig. 7.8) causes

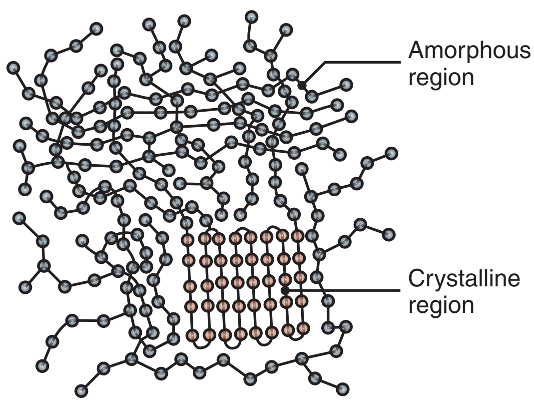


Figure 7.8: Amorphous and crystalline regions in a polymer. The crystalline region (crystallite) has an orderly arrangement of molecules. The higher the crystallinity, the harder, stiffer, and less ductile the polymer.

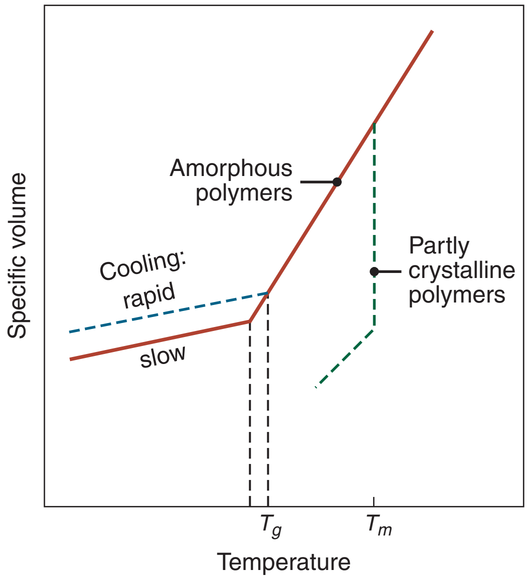


Figure 7.9: Specific volume of polymers as a function of temperature. Amorphous polymers, such as acrylic and polycarbonate, have a glass-transition temperature, but do not have a specific melting point. Partly crystalline polymers, such as polyethylene and nylons, contract sharply while passing through their melting temperatures during cooling.

opaqueness. Furthermore, because the index of refraction is proportional to density, the greater the density difference between the amorphous and the crystalline phases, the greater is the opaqueness of the polymer. Polymers that are completely amorphous can be transparent, such as polycarbonate and acrylics.

7.2.3 Glass-transition Temperature

Although amorphous polymers do not have a specific melting point, they undergo a distinct change in their mechanical behavior across a narrow range of temperatures. At low temperatures, they are hard, rigid, brittle, and glassy; at high temperatures, they are rubbery or leathery. The temperature at which a transition occurs is called the **glass-transition temperature** (T_g), also called the *glass point* or *glass temperature*. The term glass is used in this description because glasses, which are amorphous solids, behave in the same manner (see *metallic glasses*, Section 6.15 and *glass*, Section 8.4). Although most amorphous polymers exhibit this behavior, an exception is polycarbonate, which is neither rigid nor brittle below its glass-transition temperature. Polycarbonate is tough at ambient temperatures and is used for safety helmets and shields.

To determine T_g, a plot of the specific volume of the polymer as a function of temperature is produced; T_g occurs where there is a sharp change in the slope of the curve (Fig. 7.9). For highly cross-linked polymers, the slope of the curve changes gradually near T_g, making it difficult to determine their T_g. Glass-transition temperature varies with the type of polymer (Table 7.2), and it can be above or below room temperature. Unlike amorphous polymers, partly crystalline polymers have a distinct melting point, T_m (Fig. 7.9; see also Table 7.2). Because of the structural changes (called *first-order changes*) that occur, the specific volume of the polymer drops rapidly as its temperature is reduced.

7.2.4 Polymer Blends

The brittle behavior of amorphous polymers below their glass-transition temperature can be reduced by *blending* them, usually with small quantities of an **elastomer** (Section 7.9). The tiny particles that make

Table 7.2: Glass-transition and Melting Temperatures of Some Polymers.

Material	T_g (°C)	T_m (°C)
Nylon 6,6	57	265
Polycarbonate	150	265
Polyester	73	265
Polyethylene		
High density	–90	137
Low density	–110	115
Polymethylmethacrylate	105	—
Polypropylene	–14	176
Polystyrene	100	239
Polytetrafluoroethylene	–90	327
Polyvinyl chloride	87	212
Rubber	–73	—

up the elastomer are dispersed throughout the amorphous polymer, enhancing its toughness and impact strength by improving its resistance to crack propagation; these polymer blends are known as **rubber-modified polymers**.

Blending involves combining several components, developing **polyblends** that utilize the favorable properties of different polymers. **Miscible blends** (meaning mixing without separation of two phases) are produced by a process similar to the alloying of metals that enables polymer blends to become more ductile. Polymer blends account for about 20% of all polymer production.

7.3 Thermoplastics

It was noted above that within each molecule, the bonds between adjacent long-chain molecules (secondary bonds) are much weaker than the covalent bonds between mers (primary bonds). It is the strength of the secondary bonds that determines the overall strength of the polymer; linear and branched polymers have weak secondary bonds.

As the temperature is raised above the glass-transition temperature, T_g, or melting point, T_m, some polymers become easier to shape or to mold into desired shapes. When the polymer is cooled, it returns to its original hardness and strength; in other words, the process is reversible. Polymers that exhibit this behavior are known as **thermoplastics**, common examples of which are acrylics, cellulosics, nylons, polyethylenes, and polyvinyl chloride.

The behavior of thermoplastics also depends on other variables, including their structure and composition; among the most important are temperature and deformation rate. Below the glass-transition temperature, most polymers are *glassy* (brittle) and they behave like an elastic solid. The relationship between stress and strain is linear, as shown in Fig. 2.2. The behavior also depends on the particular polymer; for example, PMMA is glassy below its T_g, whereas polycarbonate is not. When the applied stress is increased further, polycarbonate eventually fractures, just as a piece of glass does at room temperature.

Typical stress–strain curves for some thermoplastics and thermosets at room temperature are shown in Fig. 7.10; their behavior may be described as rigid, soft, brittle, flexible, and so on. As can be noted from the mechanical properties of the polymers listed in Table 7.1, thermoplastics are about two orders of magnitude less stiff than metals, and their ultimate tensile strength is about one order of magnitude lower than that of metals (see Table 2.2).

Effects of Temperature. If the temperature of a thermoplastic polymer is raised above its T_g, it first becomes *leathery* and then, with increasing temperature, *rubbery* (Fig. 7.6). Finally, at higher temperatures (above T_m for crystalline thermoplastics), the polymer becomes a *viscous fluid*, and its viscosity decreases with

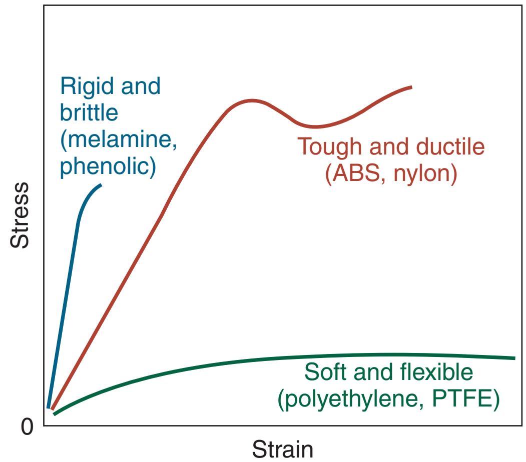


Figure 7.10: General terminology describing the behavior of three types of plastics. PTFE (polytetrafluoroethylene) has *Teflon* as its trade name. *Source:* After R.L.E. Brown.

increasing temperature. As a viscous fluid, it can be softened, molded into shapes, resolidified, remelted, and remolded several times. In practice, however, repeated heating and cooling causes **degradation** or **thermal aging** of thermoplastics.

As with metals, the strength and the modulus of elasticity of thermoplastics both decrease with increasing temperature and the ductility increases (Fig. 7.11). The effect of temperature on impact strength is shown in Fig. 7.12; note the large difference in the impact behavior among various polymers.

Effect of Rate of Deformation. When deformed rapidly, the behavior of thermoplastics is somewhat similar to metals, as shown by the strain-rate sensitivity exponent, m, in Eq. (2.9). Thermoplastics, in general, have high m values, indicating that they can undergo large *uniform deformation* in tension before they fracture. Note in Fig. 7.13 how, unlike in common metals, the necked region of the specimen elongates considerably.

This phenomenon can easily be demonstrated by stretching a piece of the plastic holder for a six-pack of beverage cans, and observing the sequence of necking and stretching behavior shown in Fig. 7.13a.

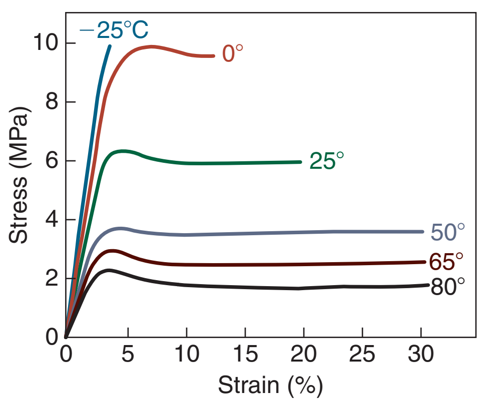


Figure 7.11: Effect of temperature on the stress–strain curve for cellulose acetate, a thermoplastic. Note the large drop in strength and the large increase in ductility with a relatively small increase in temperature. *Source:* After T.S. Carswell and H.K. Nason.

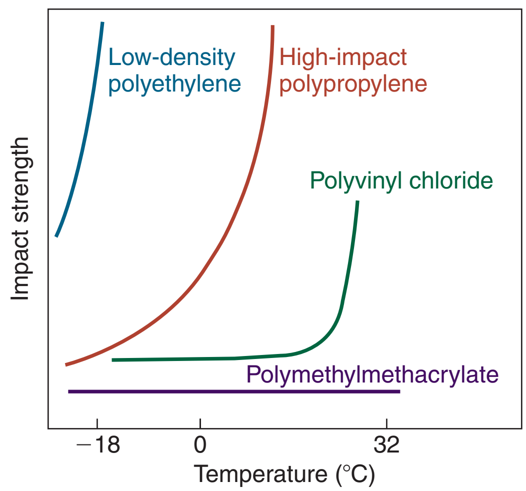


Figure 7.12: Effect of temperature on the impact strength of various plastics. Small changes in temperature can have a significant effect on impact strength. *Source:* After P.C. Powell.

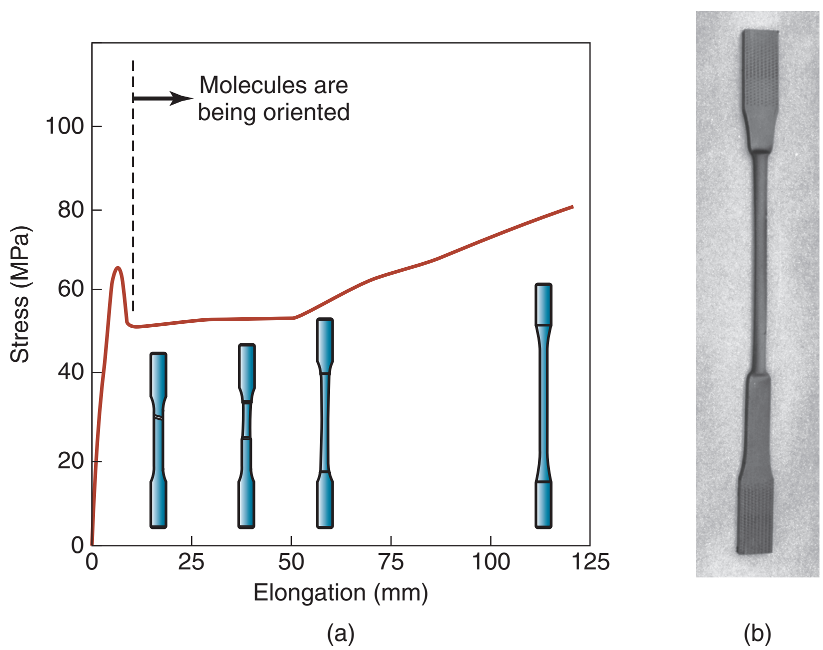


Figure 7.13: (a) Stress–elongation curve for polycarbonate, a thermoplastic. *Source:* Courtesy of R.P. Kambour and R.E. Robertson. (b) High-density polyethylene tensile-test specimen, showing uniform elongation (the long, narrow region in the specimen).

This characteristic, which is the same in the *superplastic metals* (Section 2.2.7), enables the thermoforming of thermoplastics (Section 19.6) into such shapes as candy trays, lighted signs, and packaging.

Orientation. When thermoplastics are deformed (say by stretching), the long-chain molecules tend to align themselves in the general direction of the elongation, a behavior called *orientation*. As in metals, the polymer becomes *anisotropic* (see also Section 1.6), and the specimen becomes stronger and stiffer in the elongated (stretched) direction than in its transverse direction. Stretching is an important technique for enhancing the strength and toughness of polymers, and is especially exploited in producing high-strength fibers for use in reinforced plastics (Chapter 9).

Creep and Stress Relaxation. Because of their viscoelastic behavior, thermoplastics are particularly susceptible to creep and stress relaxation (Section 2.8), and to a larger extent than metals. The extent of these phenomena depends on the particular polymer, stress level, temperature, and time. Thermoplastics exhibit creep and stress relaxation at room temperature, whereas most metals do so only at elevated temperatures.

Crazing. When subjected to tensile or bending stresses, some thermoplastics (such as polystyrene and PMMA) develop localized, wedge-shaped narrow regions of highly deformed material, called *crazing*. Although they may appear to be cracks, crazes are spongy material, typically containing about 50% voids. With increasing tensile stress, these voids coalesce and form a crack, which eventually can lead to fracture of the polymer. Crazing has been observed both in transparent, glassy polymers and in other types. The environment, particularly the presence of solvents, **lubricants**, or water vapor, can enhance the formation of crazes, called **environmental-stress cracking** and **solvent crazing**. Residual stresses in the material also contribute to crazing and cracking of the polymer.

A phenomenon related to crazing is **stress whitening**. When subjected to tensile stresses, such as those caused by folding or bending, the polymer becomes lighter in color, usually attributed to the formation of microvoids in the material. As a result, the polymer becomes less translucent (transmits less light), or more opaque. This behavior can easily be demonstrated by bending plastic components commonly found in colored binder strips for report covers, household products, and toys.

Water Absorption. An important characteristic of some polymers, such as nylons, is their ability to absorb water (*hygroscopy*). Water acts as a plasticizing agent, making the polymer more plastic (see Section 7.5); in a sense, it lubricates the chains in the amorphous regions. With increasing moisture absorption, the glass-transition temperature, the yield stress, and the elastic modulus of the polymer typically become rapidly lower. Dimensional changes also occur, especially in a humid environment.

Thermal and Electrical Properties. Compared to metals, plastics generally are characterized by low thermal and electrical conductivity, low specific gravity (ranging from 0.90 to 2.2), and high coefficient of thermal expansion (about an order of magnitude higher, as shown in Tables 3.1 and 3.2. Because most polymers have low electrical conductivity, they can be used for insulators and as packaging material for electronic components.

The electrical conductivity of some polymers can be increased by **doping** (introducing impurities, such as metal powders, salts, and iodides, into the polymer). Discovered in the late 1970s, **electrically conducting polymers** include polyethylene oxide, polyacetylene, polyaniline, polypyrrole, polythiophene, and poly(3,4-ethylene dioxitiophene), or PEDOT, commonly used for flexible electronics (see Section 28.14). The electrical conductivity of polymers increases with moisture absorption. Their electronic properties also can be changed by irradiation. Applications for conducting polymers include adhesives, microelectronic devices, rechargeable batteries, capacitors, catalysts, fuel cells, fuel-level sensors, deicer panels, radar dishes, antistatic coatings, and thermoactuating motors (used in linear-motion applications such as for power antennae, sun roofs, and power windows). One of the earliest applications of conducting polymers was in rechargeable batteries. Modern lithium rechargeable batteries use lithium or an oxide of lithium as the cathode, and lithium carbide (Li_yC_6) as the anode, separated by a conducting polymer layer.

Thermally conducting polymers are for applications requiring dimensional stability and heat transfer, such as heat sinks, and for reducing cycle times in molding and processing of thermoplastics. These polymers are typically thermoplastics, such as polypropylene, polycarbonate, and nylon, embedded with thermally conducting particles. Their conductivity can be as much as 100 times that of conventional plastics (see also *sprayed-metal tooling*, Section 20.10).

Shape-memory Polymers. Polymers also can behave in a manner similar to shape-memory alloys (Section 6.14). They can be stretched or compressed to very large strains, and then, when subjected to heat, light, or a chemical environment, they recover to their original shape. The potential applications for these polymers are similar to those for shape-memory metals, such as in opening blocked arteries, probing neurons in the brain, and improving the toughness of spines.

7.4 Thermosetting Plastics

When the long-chain molecules of a polymer are cross-linked in a three-dimensional arrangement, the structure in effect becomes one *giant molecule*, with strong covalent bonds. These polymers are called **thermosetting polymers** or **thermosets**, because, during polymerization, the network is completed and the shape of the part is permanently set. The **curing** (**cross-linking**) reaction, unlike that of thermoplastics, is irreversible. The response of a thermosetting plastic to a sufficiently elevated temperature can be likened to what happens when baking a cake or boiling an egg: Once the cake is baked and cooled, or the egg boiled and cooled, reheating it will not change its shape.

Some thermosets, such as epoxy, polyester, and urethane, cure at room temperature, because the heat produced by the *exothermic* reaction is sufficient to cure the plastic. A common thermoset is **phenolic**, which is a product of the reaction between phenol and formaldehyde. Typical products made of this polymer are the handles and knobs on cooking pots and pans and components of light switches and outlets.

The polymerization process for thermosets generally takes place in two stages. The first occurs at the chemical plant, where the molecules are partially polymerized into linear chains. The second stage occurs during the final step of part production, where cross-linking is completed under *heat* and *pressure* during molding and shaping of the part (Chapter 19).

Thermosetting polymers do not have a sharply defined glass-transition temperature. Because of the nature of the bonds, the strength and hardness of a thermoset are not affected by temperature or by rate of deformation, unlike those for thermoplastics. If the temperature is increased sufficiently, the thermosetting polymer instead will begin to burn, degrade, and char. Thermosets generally have better mechanical, thermal, and chemical properties; electrical resistance; and dimensional stability than do thermoplastics.

7.5 Additives in Plastics

Polymers usually are compounded with *additives*, which modify and improve specific characteristics of the polymer, such as stiffness, strength, color, weatherability, flammability, electric arc resistance, and ease of subsequent processing. Additives may consist of:

- **Plasticizers.** These are added to polymers to impart *flexibility* and *softness*, by lowering their glass-transition temperature. They are low-molecular-weight solvents, with high boiling points (nonvolatile); they reduce the strength of the secondary bonds between the long-chain molecules, and thus make the polymer flexible and soft. The most common use of plasticizers is in polyvinyl chloride (PVC), which remains flexible during its numerous uses, such as thin sheets, films, tubing, shower curtains, and clothing materials.

- **Colorants.** The wide variety of colors available in plastics is obtained by adding colorants, which are either *dyes* (organic) or *pigments* (inorganic).

- **Fillers**. Because of their low cost, fillers are important in reducing the cost of polymers. Depending on their type, fillers may also improve the strength, hardness, toughness, abrasion resistance, dimensional stability, and stiffness of plastics. These properties are greatest at specific percentages of different types of polymer-filler combinations. Fillers are generally wood flour (fine sawdust), silica flour (fine silica powder), clay, powdered mica, talc, calcium carbonate, and short fibers of cellulose, glass, and asbestos.

- **Flame retardants.** If the temperature is sufficiently high, most polymers will ignite and burn. The **flammability** (ability to support combustion) of polymers varies considerably, depending on their composition and especially on their chlorine and fluorine content. The flammability of polymers can be reduced either by making them from less flammable raw materials or by adding flame retardants, such as compounds of chlorine, bromine, and phosphorus. Cross-linking also reduces polymer flammability.

- **Other additives.** Most polymers are affected adversely by *ultraviolet radiation*, such as from sunlight, and by *oxygen*; they weaken and break the primary bonds and cause the scission (splitting) of the long-chain molecules. The polymer then degrades and becomes stiff and brittle. An example of protection against ultraviolet radiation is the compounding of certain polymers and rubber with **carbon black** (soot). Protection against degradation caused by oxidation, particularly at elevated temperatures, is achieved by adding **antioxidants** to the polymer.

7.6 General Properties and Applications of Thermoplastics

The general characteristics and typical applications of major classes of thermoplastics, particularly as they relate to the manufacturing and service life of plastic products and components, are outlined in this section. *General recommendations* for various plastics applications are given in Table 7.3, and Table 7.4 lists some of the more common trade names for thermoplastics.

Acetals (from *acetic* and *alcohol*) have good strength, good stiffness, and good resistance to creep, abrasion, moisture, heat, and chemicals. Typical applications include mechanical parts and components

Table 7.3: General Recommendations for Plastic Products.

Design requirement	Typical applications	Plastics
Mechanical strength	Gears, cams, rolls, valves, fan blades, impellers, pistons	Acetals, nylon, phenolics, polycarbonates, polyesters, polypropylenes, epoxies, polyimides
Wear resistance	Gears, wear strips and liners, bearings, bushings, roller blades	Acetals, nylon, phenolics, polyimides, polyurethane, ultrahigh-molecular-weight polyethylene
Frictional properties		
High	Tires, nonskid surfaces, footware, flooring	Elastomers, rubbers
Low	Sliding surfaces, artificial joints	Fluorocarbons, polyesters, polyimides
Electrical resistance	All types of electrical components and equipment, appliances, electrical fixtures	Polymethylmethacrylate, ABS, fluorocarbons, nylon, polycarbonate, polyester, polypropylenes, ureas, phenolics, silicones, rubbers
Chemical resistance	Containers for chemicals, laboratory equipment, components for chemical industry, food and beverage containers	Acetals, ABS, epoxies, polymethylmethacrylate, fluorocarbons, nylon, polycarbonate, polyester, polypropylene, ureas, silicones
Heat resistance	Appliances, cookware, electrical components	Fluorocarbons, polyimides, silicones, acetals, polysulfones, phenolics, epoxies
Functional and decorative	Handles, knobs, camera and battery cases, trim moldings, pipe fittings	ABS, acrylics, cellulosics, phenolics, polyethylenes, polypropylenes, polystyrenes, polyvinyl chloride
Functional and transparent	Lenses, goggles, safety glazing, signs, food-processing equipment, laboratory hardware	Acrylics, polycarbonates, polystyrenes, polysulfones
Housings and hollow shapes	Power tools, housings, sport helmets, telephone cases	ABS, cellulosics, phenolics, polycarbonates, polyethylenes, polypropylene, polystyrenes

Table 7.4: Trade Names for Thermoplastic Polymers.

Acetal: Delrin, Duracon, Lupital, Ultraform
Acrylic: Lucite, Acrylite, Acrysteel, Cyrolite, Diakon, Implex, Kamax, Korad, Plexiglass, XT, Zylar
Acrylic-polyvinyl chloride: Kydex
Acrylonitrile-butadiene-styrene: Cycolac, Delta, Denka, Magnum, Novodur, Royalite, Terluran
Aramid: Kevlar
Fluorocarbon: Teflon (polytetrafluoroethylene)
Polyamide: Capron, Celanese, Durethan, Grilamid, Maranyl, Nylon, Rilsan, Ultramid, Vespel, Vydyne, Zytel
Polycarbonate: APEC, Calibre, Hyzod, Lexan, Makrolon, Merlon
Polyester: Dacron, Eastpac, Ektar, Kodel, Mylar, Rynite
Polyetherimide: Ultem
Polyethylene: Alathon, Dowlex, Forar, Fortiflex, Hostalen, Marlex, Petrothene
Polyimide: Aurum, Avimid, Estamid, Envex, Kapton, Lenzing, VTEC
Polyphenylene: Forton, Fortron, Noryl
Polypropylene: Fortilene, Oleplate, Olevac, Pro-Fax
Polystyrene: Dylene, Fosta Tuf-Flex, Fostalite, Fostarene, Lustrex, Polystrol, Styron, Syrofoam
Polysulfone: Mindel, Udel
Polyurethane: Estane, Isoplast, Pellethane
Polyvinyl chloride: Fiberloc, Geon, Saran, Sintra, Tygon
Polyvinylidene fluoride: Foraflon, Kynar
Styrene-methylmethacrylate: Zerlon

requiring high performance over a long period (e.g., bearings, cams, gears, bushings, and rolls), impellers, wear surfaces, pipes, valves, shower heads, and housings.

Acrylics, such as PMMA, possess moderate strength, good optical properties, and weather resistance. They are transparent (but can be made opaque), are generally resistant to chemicals, and have good electrical resistance. Typical applications include lenses, lighted signs, displays, window glazing, skylights, bubble tops, automotive lenses, windshields, lighting fixtures, and furniture.

Acrylonitrile-butadiene-styrene (ABS) is rigid and dimensionally stable. It has good impact, abrasion, and chemical resistance; good strength and toughness; good low-temperature properties; and high electrical resistance. Typical applications include pipes, fittings, chrome-plated plumbing supplies, helmets, tool handles, automotive components, boat hulls, telephones, luggage, housing, appliances, refrigerator liners, and decorative panels.

Cellulosics have a wide range of mechanical properties, depending on their composition. They can be made rigid, strong, and tough; however, they weather poorly and are affected by heat and chemicals. Typical applications include tool handles, pens, knobs, frames for eyeglasses, safety goggles, machine guards, helmets, tubing and pipes, lighting fixtures, rigid containers, steering wheels, packaging film, signs, billiard balls, toys, and decorative parts.

Fluorocarbons possess good resistance to high temperature (*Teflon*, for example, has a melting point of 327°C), chemicals, weather, and electricity. They also have unique non-adhesive properties and low friction. Typical applications include linings for chemical-processing equipment, nonstick coatings for cookware, electrical insulation for high-temperature wire and cable, gaskets, low-friction surfaces, bearings, and seals.

Polyamides (from the words *poly*, *amine*, and *carboxyl acid*) are available in two main types: *nylons* and *aramids*:

- **Nylons**, a coined word, have good mechanical properties and abrasion resistance; they are also self-lubricating and resistant to most chemicals. All nylons are *hygroscopic* (absorb water); the moisture absorption reduces desirable mechanical properties and increases part dimensions. Typical applications include gears, bearings, bushings, rolls, fasteners, zippers, electrical parts, combs, tubing, wear-resistant surfaces, guides, and surgical equipment.

- **Aramids** (aromatic polyamides) have very high tensile strength and stiffness. Typical applications include fibers for reinforced plastics, bulletproof vests, cables, and radial tires.

Polycarbonates have good mechanical and electrical properties, high impact resistance, and they can be made resistant to chemicals. Typical applications include safety helmets, optical lenses, bullet-resistant window glazing, signs, bottles, food-processing equipment, windshields, load-bearing electrical components, electrical insulators, medical apparatus, business machine components, guards for machinery, and parts requiring dimensional stability.

Polyesters (thermoplastic polyesters; see also Section 7.7) have good mechanical, electrical, and chemical properties, good abrasion resistance, and low friction. Typical applications include gears, cams, rolls, load-bearing members, pumps, and electromechanical components.

Polyethylenes possess good electrical and chemical properties; their mechanical properties depend on composition and structure. Three major polyethylene classes are: (1) *low density* (LDPE), (2) *high density* (HDPE), and (3) *ultrahigh molecular weight* (UHMWPE). Typical applications for LDPE and HDPE are housewares, bottles, garbage cans, ducts, bumpers, luggage, toys, tubing, bottles, and packaging materials. UHMWPE is used in parts requiring high-impact toughness and resistance to abrasive wear; examples include artificial knee and hip joints.

Polyimides have the structure of a thermoplastic but the nonmelting characteristic of a thermoset (see also Section 7.7).

Polypropylenes have good mechanical, electrical, and chemical properties and good resistance to tearing. Typical applications include automotive trim and components, medical devices, appliance parts, wire insulation, TV cabinets, pipes, fittings, drinking cups, dairy-product and juice containers, luggage, ropes, and weather stripping.

Polystyrenes generally have average properties and are somewhat brittle, but inexpensive. Typical applications include disposable containers; packaging; trays for meats, cookies, and candy; foam insulation; appliances; automotive and radio/TV components; housewares; and toys and furniture parts (as a substitute for wood).

Polysulfones have excellent resistance to heat, water, and steam; they have dielectric properties that remain virtually unaffected by humidity, are highly resistant to some chemicals, but are attacked by organic solvents. Typical applications include steam irons, coffeemakers, hot-water containers, medical equipment that requires sterilization, power-tool and appliance housings, aircraft cabin interiors, and electrical insulators.

Polyvinyl chloride has a wide range of properties, is water resistant, inexpensive, and can be made rigid or flexible. It is not suitable for applications requiring strength and heat resistance. *Rigid* PVC is tough and hard; it is used for signs and in the construction industry. *Flexible* PVC is used in wire and cable coatings, in low-pressure flexible tubing and hose, and in footwear, imitation leather, upholstery, records, gaskets, seals, trim, film, sheet, and coatings.

7.7 General Properties and Applications of Thermosetting Plastics

This section outlines the general characteristics and typical applications of the major thermosetting plastics.

Alkyds (from *alkyl*, meaning alcohol, and *acid*) possess good electrical insulating properties, impact resistance, dimensional stability, and low water absorption. Typical applications are in electrical and electronic components.

Aminos have properties that depend on composition; generally, they are hard, rigid, and resistant to abrasion, creep, and electric arcing. Typical applications include small-appliance housings, countertops, toilet seats, handles, and distributor caps. **Urea** typically is used for electrical and electronic components; and **melamine** for dinnerware.

Epoxies have excellent mechanical and electrical properties, good dimensional stability, strong adhesive properties, and good resistance to heat and chemicals. Typical applications include electrical components requiring mechanical strength and high insulation, tools and dies, and adhesives.

Fiber-reinforced epoxies have excellent mechanical properties and are used in pressure vessels, rocket-motor casings, tanks, and similar structural components.

Phenolics are rigid, though brittle, and dimensionally stable, and they have high resistance to heat, water, electricity, and chemicals. Typical applications include knobs, handles, laminated panels, and telephones; bonding material to hold abrasive grains together in grinding wheels; and electrical components (such as wiring devices, connectors, and insulators).

Polyesters (thermosetting polyesters; see also Section 7.7) have good mechanical, chemical, and electrical properties. They generally are reinforced with glass (or other) fibers and also are available as casting resins. Typical applications include boats, luggage, chairs, automotive bodies, swimming pools, and materials for impregnating cloth and paper.

Polyimides possess good mechanical, physical, and electrical properties at elevated temperatures; they also have good creep resistance, low friction, and low wear characteristics. Polyimides have the nonmelting characteristic of a thermoset, but the structure of a thermoplastic. Typical applications include pump components (bearings, seals, valve seats, retainer rings, and piston rings), electrical connectors for high-temperature use, aerospace parts, high-strength impact-resistant structures, sports equipment, and safety vests.

Silicones have properties that depend on their composition; generally, they weather well, possess excellent electrical properties over a wide range of humidity and temperature, and resist chemicals and heat (Section 7.9). Typical applications include electrical components requiring strength at elevated temperatures, oven gaskets, heat seals, and waterproof materials.

Health Hazards. Some of the chemicals used in polymers may present health hazards, especially in products such as polycarbonate water containers and baby bottles, and also medical devices, sports safety equipment, and eating utensils. The chemical that is of particular concern is bisphenol A (BPA), which is widely used.

Case Study 7.2 Materials for a Refrigerator Door Liner

In selecting candidate materials for a refrigerator door liner (where eggs, butter, salad dressings, and small bottles are stored), the following factors should be considered:

1. *Mechanical requirements*: strength, toughness (to withstand impacts, door slamming, and racking), stiffness, resilience, and resistance to scratching and wear at operating temperatures.

2. *Physical requirements*: dimensional stability and electrical insulation.

3. *Chemical requirements*: resistance to staining, odor, chemical reactions with food and beverages, and cleaning fluids.

4. *Appearance*: color, stability of color over time, surface finish, texture, and feel.

5. *Manufacturing properties*: methods of manufacturing and assembly, effects of processing on material properties and behavior over a period of time, compatibility with other components in the door, and cost of materials and manufacturing.

Considering all of the factors involved, a study identified two candidate materials for door liners: ABS (acrylonitrile-butadiene-styrene) and HIPS (high-impact polystyrene). One aspect of the study concerned the effect of vegetable oils, such as from salad dressing stored in the door shelf, on the strength of these two plastics. Experiments showed that the presence of vegetable oils significantly reduced the load-bearing capacity of HIPS. It was also found that it becomes brittle in the presence of oils (a case of solvent-stress cracking), whereas ABS is not affected to any significant extent.

7.8 Biodegradable Plastics

Plastic wastes contribute about 16% of municipal solid waste by weight, and make up 50–80% of waste littering beaches, oceans, and sea beds. On a volume basis, they contribute between two and three times their weight. Only about one-third of plastic production goes into disposable products, such as bottles, packaging, and garbage bags. With the growing use of plastics and continuing concern over environmental issues regarding the disposal of plastic products and the shortage of landfills, major efforts continue to develop completely biodegradable plastics.

Traditionally, most plastic products have been made from synthetic polymers that are (a) derived from nonrenewable natural resources, (b) not biodegradable, and (c) difficult to recycle. **Biodegradability** means that microbial species in the environment (e.g., microorganisms in soil and water) will degrade all or part of the polymeric material under the proper environmental conditions, without producing toxic by-products. The end products of the degradation of the biodegradable portion of the material are carbon dioxide and water. Because of the variety of constituents in biodegradable plastics, these plastics can be regarded as composite materials (Chapter 9); consequently, only a portion of these products may be truly biodegradable.

Three main *biodegradable plastics* have thus far been developed. They have different degradability characteristics, and they degrade over different periods of time, from a few months to a few years.

1. The **starch-based system** is the farthest along in terms of production capacity. Starch may be extracted from potatoes, wheat, rice, or corn. The starch granules are processed into a powder, which is heated and becomes a sticky liquid. The liquid is then cooled, shaped into pellets, and processed in conventional plastic-processing equipment (Chapter 19). Various additives and binders are blended with the starch to impart specific characteristics to the bioplastic materials. For example, a composite of polyethylene and starch is produced commercially as degradable garbage bags.

2. In the **lactic-based system**, fermenting feedstocks produce lactic acid, which is then polymerized to form a polyester resin. Typical uses include medical and pharmaceutical applications.

3. In **fermentation of sugar**, organic acids are added to a sugar feedstock. The resulting reaction produces a highly crystalline and very stiff polymer, which, after further processing, behaves in a manner similar to polymers developed from petroleum.

Studies continue to be conducted on producing fully biodegradable plastics by using various agricultural waste (*agrowastes*), plant carbohydrates, plant proteins, and vegetable oils. Typical applications of this approach include the following:

- Disposable tableware made from a cereal substitute, such as rice grains or wheat flour.
- Plastics made almost entirely from starch extracted from potatoes, wheat, rice, or corn.
- Plastic articles made from coffee beans and rice hulls that are dehydrated and molded under high pressure and temperature.
- Water-soluble and compostable polymers for medical and surgical use.
- Food and beverage containers made from potato starch, limestone, cellulose, and water, which can dissolve in storm sewers and oceans without affecting wildlife or marine life.

Recycling of Plastics. Much effort continues to be expended globally on collecting and recycling of used plastic products. Thermoplastics are recycled by melting, blending, and reforming them into other products. *Recycling symbols*, in the shape of a triangle outlined by three clockwise arrows and with a number in the middle, are now commonly used. These numbers identify the following plastics:

1. PETE (polyethylene)
2. HDPE (high-density polyethylene)
3. V (vinyl)
4. LDPE (low-density polyethylene)
5. PP (polypropylene)
6. PS (polystyrene)
7. Other

7.9 Elastomers (Rubbers)

Elastomers, derived from the words *elastic* and *mer*, consist of a large family of amorphous polymers (Section 7.2.1) with a low glass-transition temperature. They have the characteristic ability to undergo large elastic deformations without rupture; they are soft and have low elastic modulus.

The structure of elastomer molecules is highly kinked (tightly twisted or curled): they stretch but then return to their original shape after the load is removed (Fig. 7.14). Elastomer can also be cross-linked, the best example of which is the elevated-temperature **vulcanization** of rubber with sulfur (discovered by C. Goodyear, in 1839, and named for Vulcan, the Roman god of fire). Once the elastomer is cross-linked, it cannot be reshaped. An automobile tire (a giant molecule) cannot be softened and reshaped.

The terms *elastomer* and *rubber* often are used interchangeably. Generally, however, an **elastomer** is defined as being capable of recovering substantially in shape and size after the load has been removed. **Rubber** is defined as being capable of recovering quickly from large deformations.

The hardness of elastomers, which is measured with a *durometer* (Section 2.6.1), increases with the cross-linking of the molecular chains. As with plastics, a variety of additives can be blended into elastomers to impart specific properties. Elastomers have a wide range of applications, such as high-friction and nonskid surfaces, protection against corrosion and abrasion, electrical insulation, and shock and vibration insulation. Examples include tires, hoses, weather stripping, footwear, linings, gaskets, seals, printing rolls, and flooring.

An important property of elastomers is their hysteresis loss in stretching or compression (Fig. 7.14). The clockwise loop indicates energy loss, whereby mechanical energy is converted into heat. This property is important for absorbing vibrational energy (damping) and sound insulation.

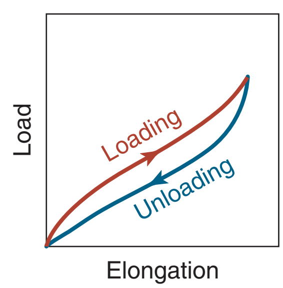


Figure 7.14: Typical load–elongation curve for rubbers. The clockwise loop, indicating the loading and the unloading paths, displays the *hysteresis loss*. Hysteresis gives rubbers their capacity to dissipate energy, damp vibration, and absorb shock loading, as is necessary in automobile tires and in vibration dampers placed under machinery.

Natural Rubber. The source for natural rubber is **latex**, a milk-like sap obtained from the inner bark of a tropical tree. Natural rubber has good resistance to abrasion and fatigue, and high friction but low resistance to oil, heat, ozone, and sunlight. Typical applications are tires, seals, couplings, and engine mounts.

Synthetic Rubbers. Examples of synthetic rubbers are butyl, styrene butadiene, polybutadiene, and ethylene propylene. Compared to natural rubber, they have better resistance to heat, gasoline, and chemicals, and they have a higher temperature range. Synthetic rubbers resistant to oils are neoprene, nitrile, urethane, and silicone. Typical applications are tires, shock absorbers, seals, and belts.

Silicones. Silicones (Section 7.9) have the highest useful temperature range of elastomers (up to 315°C), but properties such as strength and resistance to wear and oils generally are inferior to those in other elastomers. Typical applications of silicones are seals, gaskets, thermal insulation, high-temperature electrical switches, and electronic apparatus.

Polyurethane. This elastomer has very good overall properties of high strength, stiffness, and hardness, and it also has exceptional resistance to abrasion, cutting, and tearing. Typical applications include seals, gaskets, cushioning, diaphragms for the rubber forming of sheet metals (Section 16.8), and auto body parts.

7.10 Gels

A **gel** is mostly liquid by weight, but contains a cross-linked network within its structure. Edible jelly is a common example of a gel. The liquid in a gel is held by the network through surface tension, and physical and/or chemical bonds.

Hydrogels have water as the contained liquid, and are highly absorbant. Common examples of hydrogels are polyvinyl alcohol, sodium polyacrylate, hydroxyethylmethacrylate (HEMA), polyethylene glycol (PEG), and several others. Because they can be made to have mechanical properties similar to that of soft tissue, there are numerous biomedical applications of hydrogels. The most common application is that of soft contact lenses, which are abrasive when dry, but absorb liquid from the user, providing a smooth interface between the lens and the eye.

Aerogels are derived from gels, where the liquid has been replaced by air. The result is a soft material with extremely low density (160 g/m^3). Some aerogels are over 98% air and have a density as low as 1900 g/m^3. Aerogels are very effective at sound isolation and thermal insulation, and are also very effective desiccants (water absorbing agents). Common aerogels are produced from silica, graphite, and graphene (Section 8.6.3), or from organic polymers. The lightest material produced to date is *graphene aerogel*. Its density is around 1160 g/m^3, including the air in the aerogel. *Metal oxide aerogels* are used as catalysts in chemical processing industries.

Summary

- Polymers are a major class of materials, and possess a very wide range of mechanical, physical, chemical, and optical properties. Compared to metals, polymers are generally characterized by (a) lower density, strength, elastic modulus, thermal and electrical conductivity, cost; (b) higher strength-to-weight ratio, higher resistance to corrosion, higher thermal expansion, (c) wide choice of colors and transparencies; and (d) greater ease of manufacture into complex shapes.

- Plastics are composed of polymer molecules and various additives. The smallest repetitive unit in a polymer chain is called a mer. Monomers are linked by polymerization processes (condensation or addition) to form larger molecules. The glass-transition temperature separates the region of brittle behavior in polymers from that of ductile behavior.

- The properties of polymers depend on their molecular weight, structure (linear, branched, cross-linked, or network), degrees of polymerization and crystallinity, and on additives present in their formulation. Additives have such functions as improving strength, flame retardation, lubrication, imparting flexibility and color, and providing stability against ultraviolet radiation and oxygen. Polymer structures can be modified by several means to impart a wide range of desirable properties.
- Two major classes of polymers are thermoplastics and thermosets. Thermoplastics become soft and easy to form at elevated temperatures. Their behavior includes such phenomena as creep and stress relaxation, crazing, and water absorption. Thermosets are produced by cross-linking polymer chains; they do not become soft to any significant extent with increasing temperature, and are much more rigid and harder than thermoplastics.
- Elastomers have a characteristic ability to undergo large elastic deformations and then return to their original shapes when unloaded. Consequently, they have important applications in tires, seals, footwear, hoses, belts, and shock absorbers.
- Among important considerations in polymers are their recyclability and biodegradability. Several formulations of biodegradable plastics are available, and others are under continued development.
- Gels are cross-linked polymers combined with water (hydrogels) or air (aerogels), yielding unique properties. Hydrogels are useful for soft biological applications (notably contact lenses) while aerogels are widely used in chemical process industries as catalysts.

Key Terms

Additives
Biodegradable
Blends
Bonding
Branched polymers
Colorants
Crazing
Cross-linked polymers
Crystallinity
Curing
Degradation
Degree of crystallinity
Degree of polymerization
Doping
Elastomer
Fillers
Flame retardants
Gels
Glass-transition temperature
Latex
Linear polymers
Lubricants
Mer
Molecular weight
Monomer
Network polymers
Orientation
Plasticizers
Plastics
Polyblends
Polymer
Polymerization
Primary bonds
Recycling
Rubber
Secondary bonds
Shape-memory polymers
Silicones
Stress whitening
Thermal aging
Thermoplastics
Thermosets
Vulcanization

Bibliography

Bhowmick, A.K., and Stephens, H.L., **Handbook of Elastomers**, 2nd ed., CRC Press, 2000.
Campo, E.A., **Selection of Polymeric Materials**, William Andrew, 2008.
Chanda, M., **Plastics Technology Handbook**, 5th ed., CRC Press, 2017.
Concise Encyclopedia of Medical Polymers and Polymeric Materials, CRC Press, 2017.
Drobny, J.G., **Handbook of Thermoplastic Elastomers**, William Andrew, 2007.
Ellis, B., and Smith, R., **Plastics: A Property Database**, 2nd ed., CRC Press, 2008.
Fink, J.K., **High Performance Polymers**, 2nd ed., William Andrew, 2014.
Flick, C.W., **Plastics Additives**, William Andrew, 2002.
Gilbert, M., **Brydson's Plastics Materials**, 8th ed., Butterworth-Heinemann, 2016.
Goodship, V., **Introduction to Plastics Recycling**, 2nd ed., Smithers Rapra Press, 2008.
Harper, C., **Handbook of Plastics, Elastomers, and Composites**, 4th ed., McGraw-Hill, 2003.
Harper, C.A., **Modern Plastics Handbook**, McGraw-Hill, 2000.
Khemani, K., and Scholz, C., **Degradable Polymers and Materials: Principles and Practice**, 2nd ed., American Chemical Society, 2013.
Klingender, R.C. (ed.), **Handbook of Specialty Elastomers**, CRC Press, 2018.
Kutz, M., **Applied Plastics Engineering Handbook: Processing and Materials**, 2nd ed., William Andrew, 2016.
Margolis, J., **Engineering Plastics Handbook**, McGraw-Hill, 2006.
Mark, J.E. (ed.), **Physical Properties of Polymers Handbook**, 2nd ed., Springer, 2006.
Mills, N., **Plastics: Microstructure and Engineering Applications**, 3rd ed., Butterworth-Heinemann, 2006.
Mittal, V. (ed.), **High Performance Polymers and Engineering Plastics**, 2nd ed., Wiley Scrinever, 2014.
Osswald, T.A., and Menges, G., **Materials Science of Polymers for Engineers**, 3rd ed., Hanser, 2012,
Strong, A.B., **Plastics: Materials and Processing**, 3rd ed., Prentice Hall, 2005.
Ward, I.M., and Sweeny, J., **An Introduction to the Mechanical Properties of Solid Polymers**, 2nd ed., Wiley, 2004.
Wypych, G., **Handbook of Polymers**, 2nd ed., ChemTech Publishing, 2016.
Xanthos, M., **Functional Fillers for Plastics**, 2nd ed., Wiley-VCH, 2010.
Young, R.J., and Lovell, P., **Introduction to Polymers**, 3rd ed., CRC Press, 2008.
Zweifel, H., Maier, H., and Schiller, M., **Plastics Additives Handbook**, 6th ed., Hanser, 2009.

Review Questions

7.1. Summarize the important mechanical and physical properties of plastics.

7.2. What are the major differences between (a) the mechanical and (b) the physical properties of plastics and metals?

7.3. List properties that are influenced by the degree of polymerization.

7.4. What is the difference between condensation polymerization and addition polymerization?

7.5. Explain the differences between linear, branched, and cross-linked polymers.

7.6. What is the glass-transition temperature?

7.7. List and explain the additives commonly used in plastics.

7.8. What is crazing?

7.9. What are polyblends?

7.10. List the major differences between thermoplastics and thermosets.

7.11. What is an elastomer?

7.12. What effects does a plasticizing agent have on a polymer?

7.13. Define the following abbreviations: PMMA, PVC, ABS, HDPE, and LDPE.

7.14. Explain why it would be advantageous to produce a polymer with a high degree of crystallinity.

7.15. What are the differences and similarities of addition and condensation polymerization?

7.16. Are molecular weight and degree of polymerization related? Explain.

7.17. Why do polymers need to be dried before processing?

Qualitative Problems

7.18. What characteristics of polymers make them attractive for clothing?

7.19. Do polymers strain harden more than metals or vice versa? Explain.

7.20. Inspect various plastic components in an automobile, and state whether they are made of thermoplastic materials or of thermosetting plastics.

7.21. Give applications for which flammability of plastics would be of major importance.

7.22. What characteristics make polymers advantageous for applications such as gears? What characteristics are drawbacks in such applications?

7.23. What properties do elastomers have that thermoplastics in general do not have?

7.24. Do you think that the substitution of plastics for metals in products traditionally made of metal may be viewed negatively by the public at large? If so, why?

7.25. Is it possible for a material to have a hysteresis behavior that is the opposite of that shown in Fig. 7.14, so that the two arrows run counterclockwise? Explain.

7.26. Observe the behavior of the specimen shown in Fig. 7.13, and state whether the material has a high or a low strain-rate sensitivity exponent, m (see Section 2.2.7).

7.27. Add more to the applications column in Table 7.3.

7.28. Discuss the significance of the glass-transition temperature, T_g, in engineering applications.

7.29. Describe how a rechargeable lithium battery works.

7.30. Explain how cross-linking improves the strength of polymers.

7.31. Describe the methods by which the optical properties of polymers can be altered.

7.32. How can polymers be made to conduct electricity? Explain.

7.33. Explain the reasons for which elastomers were developed.

7.34. Give several examples of plastic products or components in which creep and stress relaxation would be important considerations.

7.35. Describe your opinions regarding the recycling of plastics versus the development of plastics that are biodegradable.

7.36. Explain how you would go about determining the hardness of plastics.

7.37. Compare the values of the elastic modulus given in Table 7.1 to the values for metals given in Chapters 2, 5, and 6.

7.38. Why is there so much variation in the stiffness of products made of polymers? Explain.

7.39. Explain why thermoplastics are easier to recycle than thermosets.

7.40. Give an example of a process where crazing is desirable.

7.41. Describe the principle behind shrink wrapping.

7.42. List and explain some environmental pros and cons of using plastic shopping bags vs. paper bags.

7.43. List the characteristics required of a polymer for (a) a bucket, (b) a golf ball, (c) an automobile dashboard, (d) clothing, (e) flooring, and (f) fishing nets.

7.44. How can you tell whether a part is made of a thermoplastic or a thermoset?

7.45. As you know, there are plastic paper clips available in various colors. Why are there no plastic staples?

7.46. By incorporating small amounts of a blowing agent, it is possible to manufacture hollow polymer fibers with gas cores. List possible applications for such fibers.

7.47. In injection-molding operations, it is common practice to remove the part from its runner, to place the runner into a shredder, and to recycle the resultant pellets. List the concerns you would have in using such recycled pellets as opposed to so-called virgin pellets.

7.48. From an environmental standpoint, do you feel it is best to incorporate polymers or metals into designs? Explain your answer.

Quantitative Problems

7.49. Calculate the areas under the stress–strain curve (toughness) for the materials shown in Fig. 7.11, plot them as a function of temperature, and describe your observations.

7.50. Note in Fig. 7.11 that, as expected, the elastic modulus of the polymer decreases as temperature increases. Using the stress–strain curves in the figure, make a plot of the modulus of elasticity versus the temperature. Comment on the shape of the curve.

7.51. A rectangular cantilever beam 75 mm high, 20 mm wide, and 1 m long is subjected to a concentrated load of 50 kg at its end. From Table 7.1, select three unreinforced and three reinforced materials and calculate the maximum deflection of the beam in each case. Then select aluminum and steel for the same beam dimensions, calculate the maximum deflection, and compare the results.

7.52. Estimate the number of molecules in a typical automobile tire. Estimate the number of atoms in the tire.

7.53. Using strength and density data, determine the minimum weight of a 2-m-long tension member that must support a load of 10,000 N if it is manufactured from (a) high-molecular-weight polyethylene, (b) polyester, (c) rigid PVC, (d) ABS, (e) polystyrene, and (e) reinforced nylon.

7.54. Plot the following for any five polymers described in this chapter: (a) ultimate tensile strength vs. density and (b) elastic modulus vs. ultimate tensile strength. Where appropriate, plot a range of values.

Synthesis, Design, and Projects

7.55. Conduct an Internet search, and describe differential scanning calorimetry. What does this measure?

7.56. Describe the design considerations involved in replacing a metal beverage container with one made of plastic.

7.57. Assume that you are manufacturing a product in which all of the gears are made of metal. A salesperson visits you and asks you to consider replacing some of these metal gears with plastic ones. Make a list of the questions that you would raise before making a decision.

7.58. Assume you work for a company that produces polymer gears. You have arranged to meet with a potential new customer, who currently uses gears made of metal. Make a list of the benefits that plastic gears present, and prepare a presentation for the meeting.

7.59. Sections 7.6 and 7.7 list several plastics and their applications. Rearrange this information by making a table of products (gears, helmets, luggage, electrical parts, etc.) that shows the types of plastic that can be used to make these products.

7.60. Make a list of products or parts that currently are not made of plastics and offer possible reasons why they are not.

7.61. Review the three curves shown in Fig. 7.10 and give some applications for each type of behavior. Explain your choices.

7.62. Repeat Problem 7.61 for the curves shown in Fig. 7.12.

7.63. In order to use a steel or aluminum container for an acidic liquid, such as tomato sauce, a polymeric barrier is usually placed between the container and its contents. Describe some methods of producing such a barrier.

7.64. Perform a study of plastics used for some products. Measure the hardness and stiffness of these plastics. (For example, dog chew toys use plastics with a range of properties.)

7.65. Add a column to Table 7.1 that describes the appearance of these plastics, including available colors and opaqueness.

7.66. With Table 7.3 as a guide, inspect various products both in a typical kitchen and in an automobile, and describe the types of plastics that could be used in making their individual components.

Chapter 8

Ceramics, Glass, Graphite, Diamond, and Nanomaterials: Structure, General Properties, and Applications

Case Studies:

- Ceramics, glass, and various forms of carbon present unique combinations of mechanical and physical properties that cannot be obtained with other metallic or nonmetallic materials.
- Ceramic materials are first described in terms of their chemistry, microstructure, mechanical and physical properties, and applications.
- The basic types of ceramics include oxide ceramics, such as aluminum or zirconium oxide, and carbides and nitrides.
- Glasses have numerous formulations, all containing at least 50% silica. Their general properties and typical uses are described.
- Several forms of carbon are commercially important; graphite is the most common, with numerous uses, including as reinforcement in composite materials, electrodes for electrical discharge machining, and solid lubricant.

- Diamond is the hardest material known and, as such, is used for precision and abrasive machining and for polishing operations.
- Nanomaterials, such as carbon nanotubes and graphene, are becoming increasingly important, with numerous applications for nanoscale electrical and microelectronic devices.

8.1 Introduction

The various types of materials described in the preceding chapters are not suitable for certain engineering applications, including the following:

1. An electrical or thermal insulator for use at high temperatures
2. Floor tiles to resist scuffing and abrasion
3. A transparent baking dish
4. Small ball bearings that are light, rigid, hard, and resist high temperatures
5. Automobile windshields that are hard, abrasion resistant, and transparent
6. Cutting tools that remain hard and wear resistant at high pressure and temperature
7. Optics applications such as lenses and mirrors that require high levels of transparency.

It is apparent from these examples that the specific properties required include high-temperature strength; hardness; desired optical properties of transparency; inertness to chemicals, foods, and the environment; resistance to wear and corrosion; and low electrical and thermal conductivity.

The general characteristics and applications of those ceramics, glasses, and glass ceramics that are of importance in engineering applications and in manufacturing are first described. Because of their unique properties and uses, the various forms of carbon (graphite, diamond, carbon nanotubes, and graphene) are described next. The manufacturing of ceramic and of glass components and various shaping and finishing operations are detailed in Chapter 18. Composites, which contain combinations of the materials described, are described in Chapter 9.

8.2 The Structure of Ceramics

Ceramics are compounds of metallic and nonmetallic elements. The term *ceramics* (from the Greek *keramos*, meaning potter's clay, and keramikos, meaning clay products) refers both to the material and to the ceramic product itself. Because of the large number of possible combinations of various elements, a wide variety of ceramics is now available for a broad range of consumer and industrial applications. The earliest use of ceramics was in pottery and bricks, dating back to before 4000 B.C. They have become increasingly important in tool and die materials, medical products and automotive components (such as exhaust-port liners, automotive spark plugs, coated pistons, food processing equipment, and cylinder liners).

Ceramics may be divided into two general categories:

1. **Traditional ceramics**, such as whiteware, tiles, brick, pottery, and abrasive wheels
2. **Industrial ceramics** (also called **engineering, high-tech**, or **fine ceramics**), such as automotive, turbine, structural, and aerospace components heat exchangers, semiconductors, and cutting tools.

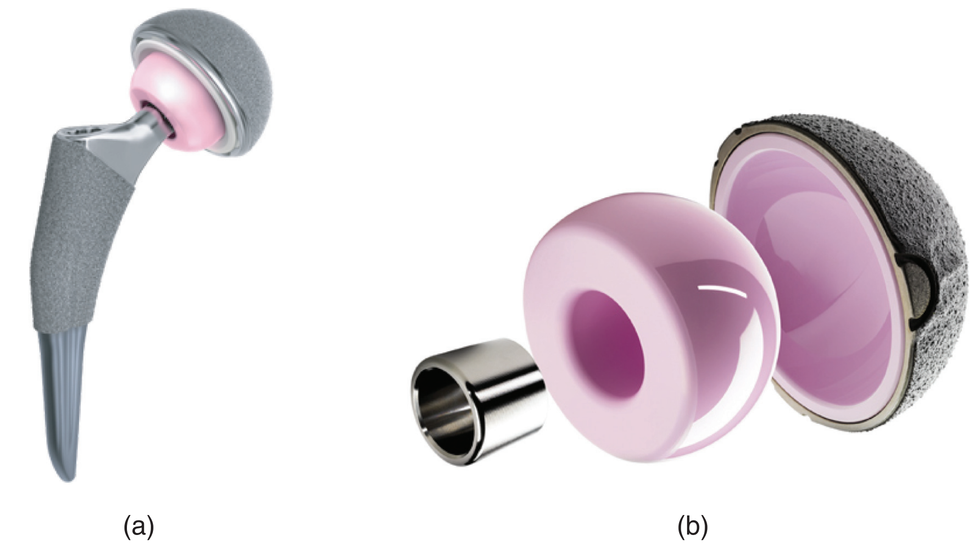


Figure 8.1: (a) A ceramic total hip replacement. (b) Detail of the ceramic ball and cup produced from a zirconia-alumina blend. Ceramics perform well in such applications because of their high hardness and wear resistance. *Source:* Courtesy of DePuy, a Johnson & Johnson Company.

The structure of ceramic crystals, containing various atoms of different sizes, is among the most complex of all material structures. The bonding between these atoms is generally **covalent** or **ionic** (Section 1.2), and as such are much stronger than metallic bonds. Consequently, properties such as hardness and thermal and electrical resistance are significantly higher in ceramics than in metals (Tables 3.1 and 3.2). Ceramics are available in *single-crystal* or *polycrystalline* form. Grain size has a major influence on the strength and properties of ceramics; the finer the grain size (hence the term **fine ceramics**), the higher the strength and toughness.

8.2.1 Raw Materials

Among the oldest of the raw materials used for making ceramics is **clay**, which has a fine-grained sheet-like structure. The most common example is *kaolinite* (from Kaoling, a hill in China) which is a white clay consisting of silicate of aluminum, with alternating weakly–bonded layers of silicon and aluminum ions (Fig. 8.2). When blended with kaolinite, water attaches itself to these layers (*adsorption*); this makes the layers slippery and gives wet clay both its well-known softness and the plastic properties (*hydroplasticity*) that make it easily formable.

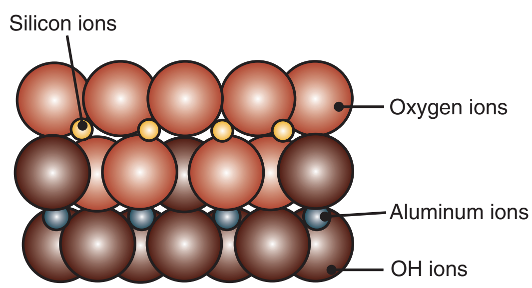


Figure 8.2: The crystal structure of kaolinite, commonly known as clay; compare with Figs. 1.3–1.5 for metals.

Other major raw materials for ceramics that are found in nature include **flint** (a rock composed of very fine grained silica, SiO_2) and **feldspar** (a group of crystalline minerals consisting of aluminum silicates and potassium, calcium, or sodium). **Porcelain** is a white ceramic composed of kaolin, quartz, and feldspar; its largest use is in appliances and kitchen and bath ware. In their natural state, these raw materials generally contain impurities of various types, which have to be removed prior to their further processing into specific products.

8.2.2 Oxide Ceramics

There are two major types of oxide ceramics: alumina and zirconia (Table 8.1).

Alumina. Also called **corundum** or **emery**, *alumina* (aluminum oxide, Al_2O_3) is the most widely used *oxide ceramic*, either in pure form or as a raw material to be blended with other oxides. It has high hardness and moderate strength. Although alumina exists in nature, it contains varying levels of impurities and

Table 8.1: Types, General Characteristics and Principal Uses of Ceramics.

Type	General characteristics and uses
Oxide ceramics	
Alumina	High hardness and moderate strength; most widely used ceramic; cutting tools; abrasives; electrical and thermal insulation.
Zirconia	High strength and toughness; thermal expansion close to cast iron; suitable for high-temperature applications such as metallurgical furnace linings, jet engine components and nuclear fuel cladding.
Carbides	
Tungsten carbide	Hardness, strength, and wear resistance depend on cobalt binder content; commonly used for dies and cutting tools.
Titanium carbide	Not as tough as tungsten carbide; has nickel and molybdenum as the binder; used as cutting tools.
Silicon carbide	High-temperature strength and wear resistance; used for heat engines and as abrasives in grinding wheels.
Nitrides	
Cubic boron nitride	Second-hardest substance known, after diamond; used as abrasives and cutting tools.
Titanium nitride	Gold in color; used as coatings because of low frictional characteristics.
Silicon nitride	High resistance to creep and thermal shock; used in high-temperature applications such as turbocharger components, rolling element bearings and cutting tools.
Sialon	Consists of silicon nitrides and other oxides and carbides; used as cutting tools and feed tubes and linings for non-ferrous metal casting.
Cermets	Consist of oxides, carbides, and nitrides; used in high-temperature applications such as cutting tools and composite armor for military applications.
Silica	High-temperature resistance; quartz exhibits piezoelectric effect; silicates containing various oxides are used in nonstructural applications such as fiber glass, plate glass, and optical glass.
Glasses	Contain at least 50% silica; amorphous structures; several types available with a wide range of mechanical and physical properties.
Glass ceramics	Have a high crystalline component to their structure; good thermal-shock resistance and strong. Typical applications include glass-ceramic cooking tops for stoves and cookware.
Graphite	Crystalline form of carbon; high electrical and thermal conductivity; good thermal-shock resistance, used for structural reinforcement in composite materials, electrical discharge machining electrodes, piston rings.
Diamond	Hardest substance known; available as single crystal or in polycrystalline form; used as cutting tools and abrasives and as dies for drawing fine wire.
Carbon nanotubes	Unique crystalline form of graphite, with high strength and electrical and thermal conductivity; in use in some structural composites and under investigation for MEMS and microelectronics applications.
Graphene	Single layer form of graphite with good electrical, magnetic and mechanical properties; used as cathodes in some batteries and fuel cells, and also under investigation for applications in display screens, microelectronics, and solar cells.
Nanophase ceramics	Stronger and easier to fabricate and machine than conventional ceramics; used in automotive and jet-engine applications.

possesses nonuniform properties; as a result, its performance also varies. Aluminum oxide, silicon carbide, and most other ceramics are now manufactured almost totally synthetically, so that their quality can be controlled at a consistently high level. First made in 1893, *synthetic aluminum oxide* is obtained from the fusion of molten bauxite (an aluminum-oxide ore, which is the principal source of aluminum), iron filings, and coke in electric furnaces. The cooled product is crushed and graded, by size, by passing it through standard screens. Aluminum oxide can also be blended with small amounts of other ceramics, such as titanium oxide and titanium carbide.

Structures containing alumina and various other oxides are known as **mullite** and **spinel**, typically used as refractory materials for high-temperature applications. The mechanical and physical properties of alumina are suitable particularly in applications such as electrical and thermal insulation and in cutting tools and abrasives.

Zirconia. Zirconia (zirconium oxide, ZrO_2, white in color) has toughness, resistance to thermal shock, wear, and corrosion, low thermal conductivity, and a low friction coefficient. **Partially stabilized zirconia** (PSZ) has higher strength and toughness and better reliability in performance than zirconia. It is obtained by blending zirconia with oxides of calcium, yttrium, or magnesium. This process forms a material with fine particles of tetragonal zirconia in a cubic lattice. Typical applications include dies for the hot extrusion of metals, and zirconia beads used as grinding and polishing media for aerospace coatings, for automotive primers for paint and for fine glossy print on flexible food packaging.

Two important characteristics of PSZ are its high coefficient of thermal expansion (only about 20% lower than that of cast iron), and its low thermal conductivity (about one-third that of other ceramics). Consequently, PSZ is highly suitable for heat-engine components, such as cylinder liners and valve bushings, to help keep the cast-iron engine assembly intact. **Transformation-toughened zirconia** (TTZ) has higher toughness because of dispersed tough phases in the ceramic matrix.

Case Study 8.1 Ceramic Knives

Generally made of zirconium oxide, ceramic knives are produced by a process described in Section 18.2. It starts with a blend of ceramic powder mixed with various binders, and compacted (molded) into blanks under high pressure. The blanks are then fired or *sintered* at temperatures above 1000°C for several days. An optional hot isostatic pressing operation (Section 17.3.2) can be applied to densify and toughen the ceramic. Next, they are ground and polished on a diamond wheel to form a sharp edge, and the handle is attached. The Mohs hardness (Section 2.6) of the zirconium oxide ceramic is 8.2, as compared to 6 for hardened steel and a maximum of 10 for diamond.

Among the advantages of ceramic knives over steel knives are: (a) Because of their very high hardness and wear resistance, ceramic knives can last months and even years before sharpening, depending on their frequency of use. (b) The knives are chemically inert; consequently, they do not stain and food does not stick to them, hence they are easy to clean, and leave no metallic taste or smell. (c) Because they are lightweight, they are easier to use.

The knives should be stored in wooden knife blocks and handled carefully. Sharp impact against other objects (such as dishes or dropping it on its edge on a hard surface) should be avoided, as their sharp edges can chip. Also, they should be used only for cutting (not for prying), and in cutting meat, for example, contact with bones is not advisable. The knives have to be sharpened professionally to a precise edge, using diamond grinding wheels.

Source: Courtesy of Kyocera Corporation.

8.2.3 Other Ceramics

Carbides. *Carbides* are typically used as cutting tools and die materials, and as an abrasive, especially in grinding wheels. Common examples of carbides are:

- **Tungsten carbide** (WC) consists of tungsten-carbide particles with cobalt as a binder. The amount of binder has a major influence on the material's properties; toughness increases with cobalt content, whereas hardness, strength, and wear resistance decrease.
- **Titanium carbide** (TiC) has nickel and molybdenum as its binder, and is not as tough as tungsten carbide.
- **Silicon carbide** (SiC) has high resistance to wear (thus suitable for use as an abrasive), thermal shock, and corrosion. It has a low friction coefficient and retains strength at elevated temperatures, thus suitable for high-temperature components in combustion and jet engines. First produced in 1891, synthetic silicon carbide is made from silica sand, coke, and small amounts of sodium chloride and sawdust. The process is similar to that for making synthetic aluminum oxide (Section 8.2.2).

Nitrides. Common examples of *nitrides* are:

- **Cubic boron nitride** (cBN) is the second-hardest known substance (after diamond), and has special applications, such as in cutting tools and as abrasives in grinding wheels. It does not exist in nature, and was first made synthetically in the 1970s, using techniques similar to those used in making synthetic diamond (Section 8.7).
- **Titanium nitride** (TiN) is used widely as a coating on cutting tools; it improves tool life by virtue of its low friction characteristics.
- **Silicon nitride** (Si_3N_4) has high resistance to creep at elevated temperatures, low thermal expansion, and high thermal conductivity, thus it resists thermal shock (Section 3.6). It is suitable for high-temperature structural applications, such as components in automotive engines and gas turbines, cam-follower rollers, bearings, sandblast nozzles, and components for the paper industry.

Sialon. Derived from the words *si*licon, *al*uminum, *o*xygen, and *n*itrogen, sialon consists of silicon nitride, with various additions of aluminum oxide, yttrium oxide, and titanium carbide. It has higher strength and thermal-shock resistance than silicon nitride, and is used primarily as a cutting-tool material.

Cermets. *Cermets* are combinations of a *cer*amic phase bonded with a *met*allic phase. Introduced in the 1960s and also called **black ceramics** or **hot-pressed ceramics**, they combine the high-temperature oxidation resistance of ceramics with the toughness, thermal-shock resistance, and ductility of metals. A common application of cermets is in cutting tools, with a typical composition being 70% Al_2O_3 and 30% TiC. Other cermets contain various oxides, carbides, and nitrides.

Cermets have been developed for high-temperature applications, such as nozzles for jet engines and brakes for aircraft, as well as electrical components like resistors and capacitors subjected to high temperatures. Cermets can be regarded as composite materials (Chapter 9) and can be used in various combinations of ceramics and metals bonded by powder-metallurgy techniques (Chapter 17).

8.2.4 Silica

Abundant in nature, **silica** is a polymorphic material; that is, it can have different crystal structures. The cubic structure is found in refractory bricks, used for high-temperature furnace applications. Most glasses contain more than 50% silica. The most common form of silica is **quartz**, a hard, abrasive hexagonal crystal, used extensively in communications applications, as an oscillating crystal of fixed frequency, because it exhibits the piezoelectric effect (Section 3.7).

Silicates are products of the reaction of silica with oxides of aluminum, magnesium, calcium, potassium, sodium, and iron; examples are clay, asbestos, mica, and silicate glasses. **Lithium aluminum silicate** has very low thermal expansion and thermal conductivity, and high thermal-shock resistance. Because it has very low strength and fatigue life, it is suitable only for nonstructural applications, such as catalytic converters, regenerators, and heat-exchanger components.

8.2.5 Nanoceramics and Composites

In order to improve the ductility and manufacturing characteristics of ceramics, the particle size in ceramics can be reduced by means of various techniques, most commonly gas condensation, use of sol-gels, or by combustion synthesis. Called *nanoceramics* or **nanophase ceramics**, the structure of these materials consists of atomic clusters, each containing a few thousand atoms. Control of particle size, distribution, and contamination are important.

Nanoceramics exhibit ductility at significantly lower temperatures than do conventional ceramics, and are stronger and easier to fabricate and machine, with fewer flaws. Applications of nanoceramics are in automotive components, such as valves, rocker arms, turbocharger rotors, and cylinder liners, and in jet-engine components. Nanocrystalline second-phase particles (on the order of 100 nm or less) and fibers also are used as reinforcements in composites. These composites have enhanced properties, such as improved tensile strength and creep resistance.

8.2.6 Porous Ceramics

Porous ceramics can be produced through a number of methods, including slip casting (Section 18.2.1) with a sacrificial insert or by freeze casting (Section 11.4.5). They have a combination of nano-scale and microscale structures, and are used in biomedical applications, heating elements, thermocouples, and diaphragms. Porous ceramics can be infiltrated with glass or polymers, such as in dental applications.

Bioceramics. Because of their strength and inertness, ceramics are also used as biomaterials (*bioceramics*) to replace joints in the human body, as prosthetic devices, and in dental work. Commonly used bioceramics are aluminum oxide, hydroxyapatite, tricalcium phosphate, silicon nitride, and various compounds of silica. Ceramic implants can be made porous, so that bone can grow into the porous structure (as is the case with porous titanium implants for dental work), developing a strong bond with structural integrity.

8.3 General Properties and Applications of Ceramics

Compared with metals, ceramics typically have the following relative characteristics: brittleness, high strength and hardness at elevated temperatures, high elastic modulus, low toughness, density, and thermal expansion, and low thermal and electrical conductivity. Because of the wide variety of compositions and grain size, the mechanical and physical properties of ceramics can vary considerably. Properties can also vary widely because of their sensitivity to flaws, defects, and surface or internal cracks. The presence of different types and levels of impurities and different methods of manufacturing also affect properties.

8.3.1 Mechanical Properties

The mechanical properties of selected engineering ceramics are given in Fig. 8.3 and Table 8.2. Note that their strength in tension is approximately one order of magnitude lower than their compressive strength, because of their sensitivity to cracks, impurities, and porosity. Such defects lead to the initiation and propagation of cracks under tensile stresses, thus significantly reducing the tensile strength of the ceramic. Reproducibility and reliability are therefore important aspects in the service life of ceramic components.

The tensile strength of polycrystalline ceramic increases with decreasing grain size and porosity. This relationship is represented approximately by the expression

$$S_{\rm ut} = S_{\rm ut,o} e^{-nP}, \tag{8.1}$$

where P is the volume fraction of pores in the solid (thus, if the **porosity** is 15%, $P = 0.15$), $S_{\rm ut,o}$ is the tensile strength at zero porosity; and the exponent n ranges between 4 and 7. The modulus of elasticity of ceramics is related to porosity by the expression

$$E \simeq E_o \left(1 - 1.9P + 0.9P^2\right), \tag{8.2}$$

where E_o is the elastic modulus at zero porosity.

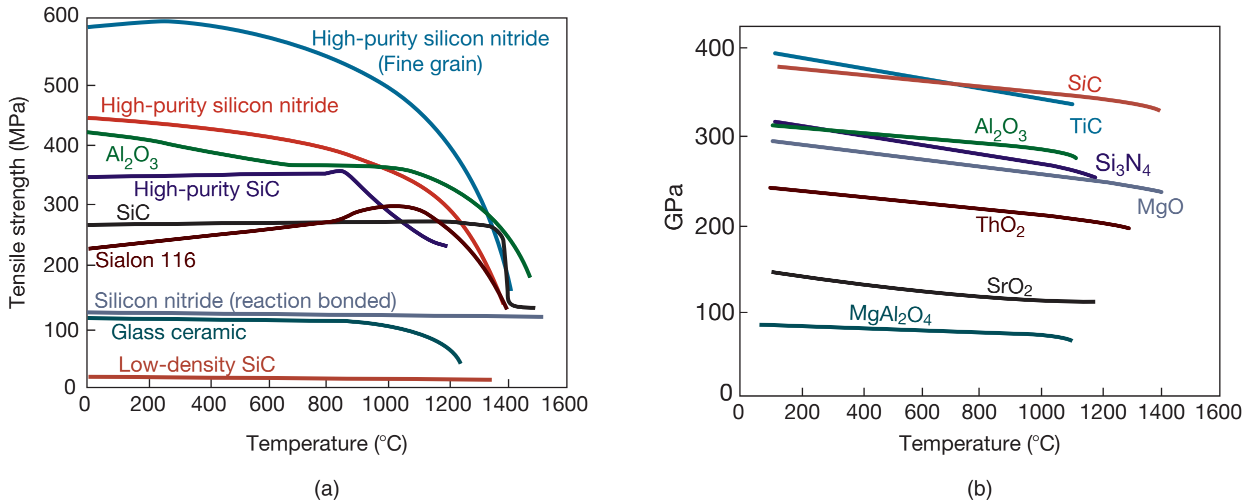


Figure 8.3: (a) Effect of temperature on the strength of various engineering ceramics. Note that much of the strength is maintained at high temperatures. (b) Effect of temperature on the modulus of elasticity for various ceramics. These results can be compared with Fig. 2.6.

Unlike most metals and thermoplastics, ceramics generally lack impact toughness and thermal-shock resistance, because of their inherent lack of ductility; once initiated, a crack propagates rapidly. Crack growth can be somewhat retarded by using a stabilizing platelet in their structure, thereby improving fatigue and toughness.

In addition to undergoing fatigue failure under cyclic loading, ceramics exhibit a phenomenon called **static fatigue**, also exhibited by glasses. When subjected to a static tensile load over time, these materials may suddenly fail, a phenomenon that occurs in environments where water vapor is present. Static fatigue, which does not occur in a vacuum or in dry air, has been attributed to a mechanism similar to the stress–corrosion cracking of metals (Section 2.10.2).

Table 8.2: Properties of Various Ceramics at Room Temperature.

Material	Symbol	Transverse rupture strength (MPa)	Compressive strength (MPa)	Elastic modulus (GPa)	Hardness (HK)	Poisson's ratio, ν	Density (kg/m^3)
Aluminum oxide	Al_2O_3	140–240	1000–2900	310–410	2000–3000	0.26	4000–4500
Cubic boron nitride	cBN	725	7000	850	4000–5000	—	3480
Diamond	—	1400	7000	830–1000	7000–8000	—	3500
Silica, fused	SiO_2	—	1300	70	550	0.25	—
Silicon carbide	SiC	100–750	700–3500	240–480	2100–3000	0.14	3100
Silicon nitride	Si_3N_4	480–600	—	300–310	2000–2500	0.24	3300
Titanium carbide	TiC	1400–1900	3100–3850	310–410	1800–3200	—	5500–5800
Tungsten carbide	WC	1030–2600	4100–5900	520–700	1800–2400	—	10,000–15,000
Partially stabilized zirconia	PSZ	620	—	200	1100	0.30	5800

Note: These properties vary widely depending on the condition of the material.

Ceramic components subjected to tensile stresses may be *prestressed*, in much the same way that concrete is prestressed by steel bars. Prestressing the shaped ceramic components subjects them to compressive stresses, with methods that include:

- Heat treatment and chemical tempering (Section 18.4)
- Laser treatment of surfaces (Section 34.8)
- Coating with ceramics that have different thermal-expansion coefficients (Section 3.6)
- Surface-finishing operations, such as grinding, in which compressive residual stresses are induced on the surfaces (Section 26.3).

Major advances have been made in improving the toughness and other properties of ceramics, including the development of **machinable** and **grindable** ceramics. Among these advances are the proper selection and processing of raw materials, the control of purity and structure, and the use of reinforcements, with particular emphasis on advanced methods of stress analysis during the design of ceramic components.

8.3.2 Physical Properties

Most ceramics have a relatively low specific gravity, ranging from about 3 to 5.8 for oxide ceramics as compared to 7.86 for iron (Table 3.1). They have very high melting or decomposition temperatures.

The thermal conductivity of ceramics varies by as much as three orders of magnitude, depending on their composition, whereas in metals it varies by only one order. As with other materials, the thermal conductivity of ceramics decreases with increasing temperature and porosity, because air is a poor thermal conductor. The thermal conductivity, k, is related to porosity by the expression

$$k = k_o\left(1 - P\right), \tag{8.3}$$

where k_o is the thermal conductivity at zero porosity and P is the porosity, as a fraction of the total volume.

Thermal expansion and thermal conductivity induce internal stresses that can then lead to thermal shock or to thermal fatigue in ceramics. The tendency toward **thermal cracking** (called **spalling** when a small piece or a layer from the surface breaks off) is lower with the combination of low thermal expansion and high thermal conductivity. For example, fused silica has high thermal-shock resistance, because of its virtually zero thermal expansion.

The *optical* properties of ceramics can be controlled by using various formulations and by controlling their structure. These methods make possible the imparting of different degrees of transparency and translucency, and of different colors. For example, single-crystal sapphire is completely transparent, zirconia is white, and fine-grained polycrystalline aluminum oxide is translucent gray. Porosity influences the optical properties of ceramics in much the same way as air trapped in home-made ice cubes, making them less transparent and giving a white appearance. Although ceramics are basically resistors, they can be made *electrically conducting* by alloying them with certain elements in order to make the ceramic behave like a semiconductor or even like a superconductor.

8.3.3 Applications

Ceramics have numerous consumer and industrial applications. Various types of ceramics are used in the electrical and electronics industries, because they have high electrical resistivity, high dielectric strength (voltage required for electrical breakdown per unit thickness), and magnetic properties suitable for such applications as magnets for speakers.

The capability of ceramics to maintain their strength and stiffness at elevated temperatures makes them suitable for high-temperature applications. The higher operating temperatures made possible by the use of ceramic components results in more efficient combustion of fuel and the reduction of emissions in automobiles. Currently, internal combustion engines are only about 30% efficient, but with the use of ceramic components, the operating performance can be improved by at least 30%.

The ceramics that are being used successfully, especially in automotive gas-turbine engine components, such as rotors, are: silicon nitride, silicon carbide, and partially stabilized zirconia. Other attractive properties of ceramics are their low density and high elastic modulus. They enable product weight to be reduced and allow the inertial forces generated by moving parts to be lower. Ceramic turbochargers, for example, are about 40% lighter than conventional ones.

High-speed components for machine tools (Part IV) also are candidates for ceramics (Section 25.3). The high elastic modulus of ceramics makes them attractive for improving the stiffness of machines, while reducing the weight. Their high resistance to wear also makes them suitable for applications such as cylinder liners, bushings, seals, bearings, and liners for gun barrels. Coating metal with ceramics is another application, often done to reduce wear, prevent corrosion, or provide a thermal barrier.

Case Study 8.2 Ceramic Ball and Roller Bearings

Silicon-nitride ceramic ball and roller bearings are used when high temperature, high speed, or marginally lubricated conditions exist. The bearings can be made entirely from ceramics, or just the ball and rollers are ceramic and the races are metal, in which case they are referred to as *hybrid bearings* (Fig. 8.4). Examples of their applications include high-performance machine tool spindles, metal-can seaming heads, high-speed flow meters, and bearings for motorcycles, go karts and snowmobiles.

The ceramic spheres have a diametral tolerance of 0.13 μm and a surface roughness of 0.02 μm. They have high wear resistance, high fracture toughness, low density, and perform well with little or no lubrication. The balls have a coefficient of thermal expansion one-fourth that of steel, and they can withstand temperatures of up to 1400°C. Produced from titanium and carbon nitride by powder-metallurgy techniques, the full-density titanium carbonitride (TiCN) or silicon nitride (Si_3N_4) bearing-grade material can be twice as hard as chromium steel and 40% lighter. Components up to 300 mm in diameter have been produced.

(a) (b)

Figure 8.4: A selection of ceramic bearings and races. *Source:* Courtesy of The Timken Company.

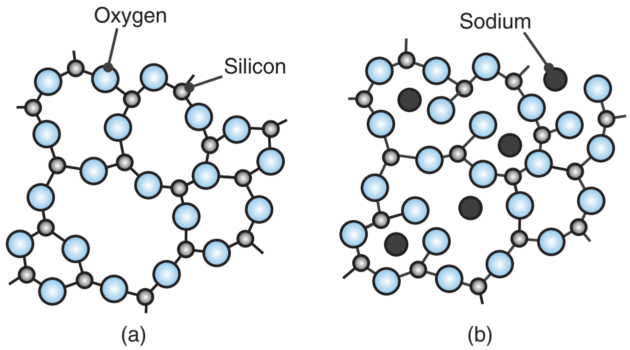


Figure 8.5: Schematic illustration of the structure of silica glass. (a) Pure silica glass, in the form of $(SiO_2)_n$ random structure; (b) partially depolymerized glass. Note that a fourth bond for each silicon is outside the plane shown.

8.4 Glasses

Glass is an *amorphous solid* with the structure of a liquid, as shown in Fig. 8.5, a condition that is obtained by *supercooling* (cooling at a rate too high to allow crystals to form). Technically, glass is defined as an inorganic product of fusion that has cooled to a rigid state without crystallizing. Glass has no distinct melting or freezing point, its behavior is thus similar to that of amorphous alloys (see *metallic glasses*, Section 6.15) and amorphous polymers (Section 7.2.2).

Glass beads first were produced in about 2000 B.C., and the art of glassblowing began in about 200 B.C. Silica was used for all glass products until the late 1600s; rapid developments in glasses then began in the early 1900s. There are about 750 different types of commercially available glasses, with applications ranging from window glass to glass for containers, cookware, lighting, and mobile phones, and to glasses with special mechanical, electrical, high-temperature, chemical inertness , corrosion, and optical characteristics. Special glasses are used in *fiber optics* (for communication by light with little loss in signal power) and in **glass fibers** with very high strength (for use in reinforced plastics, Section 9.2.1).

All glasses contain at least 50% silica, which is known as a **glass former**. The composition and properties of glasses can be modified by the addition of oxides of aluminum, sodium, calcium, barium, boron, magnesium, titanium, lithium, lead, and potassium. Depending on their function, these oxides are known as **intermediates** or **modifiers**.

8.4.1 Types of Glasses

Almost all *commercial glasses* are categorized by the following types (Table 8.3):

- **Soda-lime glass** (the most common type)
- **Lead-alkali glass**
- **Borosilicate glass**
- **Aluminosilicate glass**
- **96%-silica glass**
- **Fused silica glass**

Glasses are also classified as colored, opaque (white and translucent), multiform (variety shapes), optical, photochromatic (darkens when exposed to light, as in some sunglasses), photosensitive (changing from

Table 8.3: Properties of Various Glasses.

Property	Soda-lime glass	Lead-alkali glass	Borosilicate glass	96% silica	Fused silica
Density	High	Highest	Medium	Low	Lowest
Strength	Low	Low	Moderate	High	Highest
Resistance to thermal shock	Low	Low	Good	Better	Best
Electrical resistivity	Moderate	Best	Good	Good	Good
Hot workability	Good	Best	Fair	Poor	Poorest
Heat treatability	Good	Good	Poor	None	None
Chemical resistance	Poor	Fair	Good	Better	Best
Impact-abrasion resistance	Fair	Poor	Good	Good	Best
Ultraviolet-light transmission	Poor	Poor	Fair	Good	Good
Relative cost	Lowest	Low	Medium	High	Highest

clear to opaque), fibrous (drawn into long fibers, as in fiberglass), and foam or cellular (containing bubbles, thus a good thermal insulator). Glasses also can be referred to as **hard** or **soft**, usually in the sense of a thermal rather than mechanical property (see also hardness of glasses, Section 8.4.2). Thus, a soft glass softens at a lower temperature than does a hard glass. Soda-lime and lead-alkali glasses are considered soft, and the rest as hard.

8.4.2 Mechanical Properties

The behavior of glass, like that of most ceramics, is generally regarded as perfectly elastic and brittle. The modulus of elasticity for commercial glasses typically ranges from 55 to 90 GPa, and their Poisson's ratio from 0.16 to 0.28. The hardness of glasses, as a measure of resistance to scratching, ranges from 5 to 7 on the Mohs scale, which is equivalent to a range from around 350 to 500 HK (see Fig. 2.15).

Glass in **bulk** form generally has a strength lower than 140 MPa. The relatively low strength of bulk glass is attributed to the presence of small flaws and microcracks on its surface, some or all of which may be introduced during normal handling of the glass by inadvertently abrading it. These defects reduce the strength of glass by two to three orders of magnitude, compared to its ideal (defect free) strength. Glasses can be strengthened by thermal or chemical treatments to obtain high strength and toughness (Section 18.4).

The strength of glass theoretically can reach 35 GPa. When molten glass is drawn into fibers (**fiberglass**), its tensile strength ranges from about 0.2 to 7 GPa, with an average value of about 2 GPa. These fibers are stronger than steel, and are used to reinforce plastics in such applications as boats, automobile bodies, furniture, and sporting equipment (Tables 2.2 and 9.2).

8.4.3 Physical Properties

Glasses are characterized by low thermal conductivity and high electrical resistivity and dielectric strength. Their coefficient of thermal expansion is lower than those for metals and plastics, and may even approach zero; titanium silicate glass (a clear, synthetic high-silica glass), for example, has a near-zero coefficient of thermal expansion. Fused silica (a clear, synthetic amorphous silicon dioxide of very high purity) also has a near-zero coefficient of expansion. The optical properties of glasses, such as reflection, absorption, transmission, and refraction, can be modified by varying their composition and treatment. Glasses generally are resistant to chemical attack, and are ranked by their resistance to corrosion by acids, alkalis, or water.

8.5 Glass Ceramics

Although glasses are amorphous, **glass ceramics** have a high crystalline component to their microstructure. Glass ceramics, such as *Pyroceram* (a trade name) contain large proportions of several oxides; thus, their properties are a combination of those for glass and those for ceramics. Most glass ceramics are stronger than glass. These products are first shaped and then heat treated, whereby **devitrification** (recrystallization) of the glass occurs. Glass ceramics are generally white or gray in color.

The hardness of glass ceramics ranges approximately from 520 to 650 HK. Because they have a near-zero coefficient of thermal expansion, they also have high thermal-shock resistance, and because of the absence of porosity usually found in conventional ceramics, they are also strong.

The properties of glass ceramics can be improved by modifying their composition and by heat-treatment techniques. First developed in 1957, glass ceramics are typically used for cookware, heat exchangers in gas-turbine engines, radomes (housings for radar antennas), and electrical and electronics components. Some electric stoves use glass-ceramic cooktops and infrared halogen or radiant-heating coils; the glass ceramic is advantageous because of its high hardness and low thermal conductivity.

8.6 Graphite

Graphite is a crystalline form of **carbon,** and has a *layered structure*, with basal planes or sheets of close-packed carbon atoms (see Fig. 1.5); consequently, graphite is weak when sheared along the layers. This characteristic, in turn, gives graphite its low frictional properties behaving as a solid lubricant. However, its frictional properties are low only in an environment of air or moisture; in a vacuum, it is abrasive and thus a poor lubricant. Unlike with other materials, the strength and stiffness of graphite increase with temperature.

(a) (b)

Figure 8.6: (a) Various engineering components made of graphite. (b) Examples of graphite electrodes for electrical discharge machining. *Source:* Courtesy of (a) Poco Graphite, an Entegris Company (b) Graphel Corporation (www.graphel.com).

Amorphous graphite is known as **lampblack** (black soot) and is used as a pigment. Ordinary pencil *lead* is a mixture of graphite and clay.

Although brittle, graphite has high electrical and thermal conductivity and good resistance to thermal shock and to high temperature (although it begins to oxidize at 500°C). It is thus an important material for applications such as electrodes, heating elements, brushes for electric motors, high-temperature fixtures, furnace parts, mold materials (such as crucibles for melting and casting of metals), and seals (Fig. 8.6). An important use of graphite is as fibers in reinforced plastics and composite materials (Section 9.2). Another characteristic of graphite is its resistance to chemicals, thus it is also used in filters for corrosive fluids; its low absorption cross section and high scattering cross section for thermal neutrons make graphite also suitable for nuclear applications.

Carbon and Graphite Foams. These foams have high service temperatures, chemical inertness, low thermal expansion, and thermal and electrical properties that can be tailored to specific applications. Carbon foams are available in either graphitic or nongraphitic structures. Graphitic foams (typically produced from petroleum, coal tar, and synthetic pitch) have low density, high thermal and electrical conductivity; however, they have lower mechanical strength and are much more expensive than nongraphitic foams (produced from coal or organic resins), which are highly amorphous.

The foams have a cellular microstructure, with interconnected pores, thus their mechanical properties depend on density (see also Section 8.3). Blocks of foam can easily be machined into various complex shapes. Applications of carbon foams include their use as core materials for aircraft and ship interior panels, structural insulation, sound-absorption panels, substrates for spaceborne mirrors, lithium-ion batteries, and fire and thermal protection.

8.6.1 Fullerenes

Carbon molecules (typically C60) are produced in the shape of soccer balls, called **fullerenes** or **buckyballs**, after B. Fuller (1895–1983), the inventor of the geodesic dome. These chemically inert spherical molecules are produced from soot, and act much like solid lubricant particles. When mixed with metals, fullerenes can become superconductors at low temperatures (around 40 K). Despite their promise, no commercial applications of buckyballs currently exist.

8.6.2 Nanotubes

Carbon nanotubes can be thought of as tubular forms of graphite, and are of interest for the development of nanoscale devices (see also *nanomaterials*, Section 8.8). Nanotubes are produced most often by chemical vapor deposition (Section 34.6.2), or by laser ablation of graphite and by carbon-arc discharge.

Carbon nanotubes have exceptional strength, thus making them attractive as reinforcing fibers for composite materials. However, because they have very low adhesion with most materials, delamination with a matrix can limit their reinforcing effectiveness. It is difficult to disperse nanotubes properly because they have a tendency to clump and this limits their effectiveness as a reinforcement. Some products have used carbon nanotubes, such as bicycle frames specialty baseball bats, golf clubs, and tennis racquets. They provide only a fraction of the reinforcing material (by volume), graphite fibers playing the major role.

Another characteristic of carbon nanotubes is their very high electrical current carrying capability. They can be made as semiconductors or conductors, depending on the orientation of the graphite in the nanotube (see Fig. 8.7). Armchair nanotubes are theoretically capable of carrying a current density higher than 1000 times that for silver or copper, making them attractive for electrical connections in nanodevices (Section 29.6). Carbon nanotubes have been incorporated into polymers to improve their static-electricity discharge capability, especially in fuel lines for automotive and aerospace applications.

Carbon nanotubes are used in filter membranes, as magnetic shielding in military hardware, and in advanced batteries and capacitors where their high surface area provides better current carrying ability combined with mechanical and electrical stability. Among numerous proposed uses for carbon nanotubes are flat-panel displays, tissue engineering, electrical cables for nano-scale circuitry, catalysts, and X-ray and microwave generators.

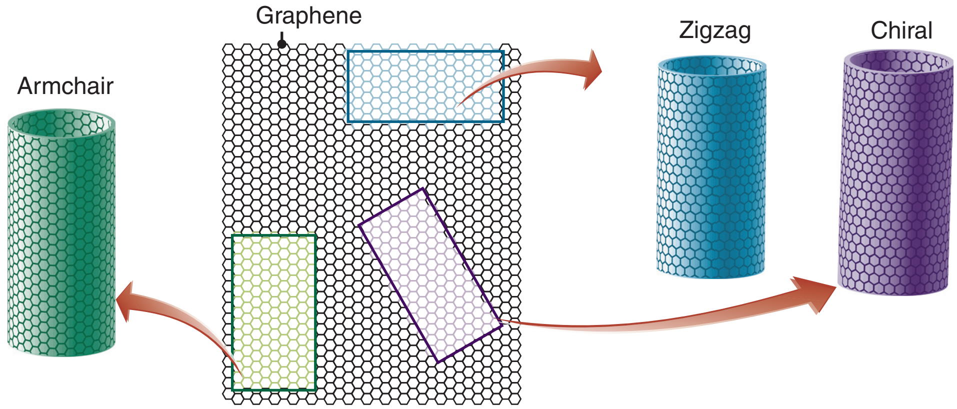


Figure 8.7: Forms of carbon nanotubes produced from a section of graphene: armchair, zigzag, and chiral. Armchair nanotubes are noteworthy for their high electrical conductivity, whereas zigzag and chiral nanotubes are semiconductors.

Annual carbon nanotube production worldwide is currently several thousand metric tons per year. They can be single-walled (SWNTs) or multi-walled (MWNTs) and can be doped with various species.

8.6.3 Graphene

Graphene can be considered to be a single sheet of graphite, or an unwrapped nanotube, as shown in Fig. 8.7; it is one of the most commonly encountered materials. Although flakes of graphene are produced whenever a pencil is abraded by paper, its direct observation, using transmission electron microscopy, dates to the early 1960s. Several methods have now been developed for producing graphene, including epitaxy (Section 28.5) on silicon carbide or metal substrates. or by chemical reduction of graphite. Currently, graphene is used as a cathode coating in some batteries, and is in development for numerous microelectronic and photonic applications (see Case Study 29.2).

8.7 Diamond

Diamond is a form of carbon, with a covalently bonded structure. It is the hardest substance known (7000–8000 HK); it is, however, brittle and begins to decompose in air at about 700°C, but resists higher temperatures in a nonoxidizing environment.

Synthetic (also called **industrial**) **diamond** was first produced in 1955. A common method of manufacturing it is to subject graphite to a hydrostatic pressure of 14 GPa and a temperature of 3000°C, referred to as *high-pressure, high-temperature* (HPHT) synthesis. An alternative is to produce diamonds through a chemical vapor deposition process (CVD, Section 34.6.2), whereby carbon is deposited onto a starting *seed* of diamond powder. The CVD process is used most often for synthetic gemstones.

Synthetic diamond has identical, and sometimes slightly superior, mechanical properties as natural diamond, because of its lack of impurities. The gemstones have a characteristic orange or yellow tint, due to impurities, resulting from the CVD process, whereas laser treatment of the diamond results changing the tint to pink or blue. However, since most of a gemstone's cost is attributed to grinding and finishing (Chapter 26) to achieve a desired shape, synthetic diamonds are only slightly less expensive than natural ones.

Synthetic diamond is available in a variety of sizes and shapes; for use in abrasive machining, the most common grit size is 0.01 mm in diameter. Diamond particles can be coated with nickel, copper, or titanium for improved performance in grinding operations. **Diamond-like carbon** also has been developed and is used as a diamond film coating, as described in Section 34.13.

Figure 8.8: A collection of synthetic diamonds. *Source:* Shutterstock/ijp2726.

In addition to its use in jewelry, gem-quality synthetic diamond has applications as heat sinks for computers, in telecommunications and integrated-circuit industries, and as windows for high-power lasers. Its electrical conductivity is 50 times higher than that of natural diamond, and it is 10 times more resistant to laser damage.

Because of its favorable characteristics, diamond has numerous important applications, such as the following:

- Cutting-tool materials, as a single crystal or in polycrystalline form
- Abrasives in grinding wheels, for hard materials
- Dressing of grinding wheels (i.e., sharpening of the abrasive grains)
- Die inserts for drawing wire less than 0.06 mm in diameter
- Coatings for cutting tools and dies.

8.8 Nanomaterials

Important developments continue to take place in the production of materials as particles, fibers, wire, tube, films, and composites, with features typically on the order of 1 nm to up to 100 nm. First investigated in the early 1980s and generally called *nanomaterials* or *nanostructured, nanocrystalline*, or *nanophase* materials, they have certain properties that are often superior to traditional materials. These characteristics include high strength, hardness, ductility, toughness, resistance to wear and corrosion, and suitable for structural (load bearing) and nonstructural applications in combination with unique electrical, magnetic, thermal, and optical properties.

The composition of a nanomaterial can be any combination of chemical elements; among the more important compositions are carbides, oxides, nitrides, metals and their alloys, organic polymers, semiconductors, and various composites. *Nanometal-polymer hybrid nanomaterials* have been developed for very lightweight components. More recent investigations include the development of *nanopaper*, with very high strength and toughness, produced from wood pulp with fibers rearranged into an entangled porous mesh.

Production methods for nanomaterials include inert-gas condensation, sputtering, plasma synthesis, electrode position, sol–gel synthesis, and mechanical alloying or ball milling. The synthesized powders are consolidated into bulk materials by various techniques, including compaction and sintering. Nanoparticles have a very high surface-area-to-volume ratio, thus affecting their behavior in processes such as diffusion and agglomeration (interaction of particles forming a cluster). Because the synthesis of nanomaterials is at atomic levels, their purity is on the order of 99.9999%, and their homogeneity and the uniformity of their microstructure are highly controlled. As a result, their mechanical, electrical, magnetic, optical, and chemical properties also can be controlled precisely. Nanomaterials are very expensive to produce and process into products, thus their cost-effectiveness is under continued study.

Applications of Nanomaterials. The unique properties of nanomaterials enable manufacturing of products that are strong, and light. The following are some current and potential applications for nanomaterials:

1. Cutting tools and inserts, made of nanocrystalline carbides and other ceramics.
2. Nanophase ceramics that are ductile and machinable.
3. Specialty bicycle frames, baseball bats, and tennis racquets, using carbon nanotubes (see also Section 8.6.2).
4. Next-generation computer chips, using nanocrystalline starting materials with very high purity, better thermal conductivity, and more durable interconnections.
5. Flat-panel displays for laptop computers and televisions, made by synthesizing nanocrystalline phosphorus to improve screen resolution.
6. Spark-plug electrodes, igniters and fuels for rockets, medical implants, high-sensitivity sensors, catalysts for elimination of pollutants, high-power magnets, and high-energy-density batteries.
7. Switches, valves, motor, and pumps.
8. Coatings made of nanomaterials are being investigated for improved wear, abrasion, corrosion resistance and thermal insulation; nanocrystalline materials; nanophase materials because of their lower thermal conductivity.

Health Hazards. Because of their extremely small size, nanoparticles can present various health *hazards* by their absorption through the skin, lungs, or the digestive track; they can also penetrate human cells. There is increasing evidence that nanoparticles can pollute the air, water, and the ground. Consequently, there is growing research on the risks of nanoparticles to humans and the environment.

Summary

- Ceramics, glasses, and various forms of carbon are of major importance in engineering applications and in manufacturing processes. Ceramics, which are compounds of metallic and nonmetallic elements, generally are characterized by high hardness, high compressive strength, high elastic modulus, low thermal expansion, high temperature resistance, good chemical inertness, low density, and low thermal and electrical conductivity. They are brittle and have low toughness.

- Ceramics are generally classified as either traditional ceramics or industrial (or high-tech) ceramics; the latter are particularly attractive for applications such as engine components, cutting tools, and components requiring resistance against wear and corrosion. Ceramics of importance in design and manufacturing are the oxide ceramics (alumina and zirconia), tungsten and silicon carbides, nitrides, and cermets.

- Glasses are supercooled liquids and are available in a wide variety of compositions and mechanical, physical, and optical properties. Glass ceramics are predominantly crystalline in structure, and have properties that are more desirable than those of glasses.

- Glasses in bulk form have relatively low strength, but they can be strengthened by thermal and chemical treatments. Glass fibers are used widely as a reinforcement in composite materials. Porous glass is used in biomedical and energy applications, because of its high surface area to volume ratio.

- Graphite, fullerenes, carbon nanotubes, graphene, and diamond are forms of carbon that display unique combinations of properties. Graphite has high-temperature use and electrical applications; graphite fibers are used to reinforce plastics and other composite materials.

- Diamond is used as cutting tools for precision machining operations, as dies for drawing of thin wire, and as abrasives for grinding wheels. Diamond-like carbon also has applications as a coating material for improved wear resistance.

- Nanomaterials have physical, mechanical, optical, chemical, and thermal properties, with several unique applications. Carbon nanotubes are of continued research interest, particularly because of their relevance to nanoscale electrical and electromechanical systems.

'

Key Terms

Alumina
Bioceramics
Buckyballs
Carbides
Carbon
Carbon foam
Carbon nanotubes
Ceramics
Cermets
Clay
Devitrification
Diamond
Diamond-like carbon
Feldspar
Flint
Fullerenes
Glass
Glass ceramics
Glass fibers
Glass former
Graphene
Graphite
Industrial ceramics
Industrial diamond
Nanoceramics
Nanophase ceramics
Nanotubes
Nitrides
Oxide ceramics
Partially stabilized zirconia
Porcelain
Porosity
Sialon
Silica
Static fatigue
Transformation-toughened zirconia
Zirconia

Bibliography

Bansal, N.P. (ed.), **Handbook of Ceramic Composites**, Springer, 2005.
Barsoum, M.W., **Fundamentals of Ceramics**, Institute of Physics Publishing, 2003.
Bhushan, B. (ed.), **Handbook of Nanotechnology**, 3rd ed., Springer, 2010.

Bhushan, B. (coordinating ed.), **Handbook of Nanomaterials Properties**, Springer, 2014.
Carter, C.B., and Norton, M.G., **Ceramic Materials: Science and Engineering**, 2nd ed., Springer, 2013.
Gogotsi, Y. and Presser, V., **Carbon Nanomaterials**, 2nd ed., CRC Press, 2013.
Harper, C.A. (ed.), **Handbook of Ceramics, Glasses, and Diamonds**, McGraw-Hill, 2001.
Hench, L.L., **A Introduction to Bioceramics,** 2nd ed., Imperial College Press, 2013.
Holand, W., and Beall, G.H., **Design and Properties of Glass Ceramics**, 2nd ed., Wiley-American Chemical Society, 2012.
Milne, W.I., and Gervasoni, J.L., **Graphene Science Handbook**, CRC Press, 2016.
O'Connell, M.J. (ed), **Carbon Nanotubes: Properties and Applications**, CRC Press, 2006.
Park, J., **Bioceramics: Properties, Characterization and Applications**, Springer, 2008.
Parinov, I.A., **Microstructure and Properties of High-Temperature Superconductors**, 2nd ed., Springer, 2013.
Pelleg, J., **Mechanical Properties of Ceramics**, Springer, 2014.
Prelas, M.A., Popovichi, G., and Bigelow, L.K. (eds.), **Handbook of Industrial Diamonds and Diamond Films**, Marcel Dekker, 1998.
Richerson, D.W. and Lee, W.E., **Modern Ceramic Engineering: Properties, Processing, and Use in Design**, 4th ed., Marcel Dekker, 2018.
Sharma, K.R., **Graphene Nanomaterials**, Momentum Press, 2014.
Shelby, J.E., **Introduction to Glass Science and Technology**, 2nd ed., Royal Society of Chemistry, 2005.
Skakalova, V., and Kaiser, L.B. (eds.) **Graphene: Properties, Preparation, Characterisation and Devices**, Woodhead Publishing, 2014.
Somiya, S. (ed.), **Handbook of Advanced Ceramics: Materials, Applications, Processing and Properties**, 2nd ed., Academic Press, 2013.
Vajtai, R. (ed.), **Handbook of Nanomaterials,** Springer, 2013.
Vollath, D., **Nanomaterials: An Introduction to Synthesis, Properties and Applications**, 2nd ed., Wiley-VCH, 2013.
Wachtman, J.B., Cannon, W.R., and Matthewson, M.J., **Mechanical Properties of Ceramics**, Wiley, 2009.

Review Questions

8.1. What is a ceramic?

8.2. List the major differences between the properties of ceramics and those of metals and plastics.

8.3. List the major types of ceramics that are useful in engineering applications.

8.4. What do the following materials typically consist of (a) carbides, (b) cermets, and (c) sialon?

8.5. What is porcelain?

8.6. What is glass? Why is it called a supercooled material?

8.7. How is glass different from a glass ceramic?

8.8. What is devitrification?

8.9. List the major types of glasses and their applications.

8.10. What is static fatigue? What is its significance?

8.11. Describe the major uses of graphite.

8.12. How are alumina ceramics produced?

8.13. What features of PSZ differentiate it from other ceramics?

8.14. What are buckyballs?

8.15. List the major uses of diamond.

8.16. What is a carbon nanotube? Explain why they are not as prevalent as other forms of carbon.

8.17. What is graphene? How is it related to graphite?

8.18. What do the terms “armchair,” “zigzag,” and “chiral” have in common?

8.19. How do platelets improve fracture strength of ceramics?

8.20. What are nanomaterials? Why are they useful?

Qualitative Problems

8.21. Explain why ceramics are weaker in tension than in compression.

8.22. What are the advantages of cermets? Suggest applications in addition to those given in this chapter.

8.23. Explain why the electrical and thermal conductivity of ceramics decreases with increasing porosity.

8.24. Explain why the mechanical property data given in Table 8.2 have such a broad range. What is the significance of this in engineering practice?

8.25. Describe the reasons that have encouraged the development of synthetic diamond.

8.26. Explain why the mechanical properties of ceramics generally differ from those of metals.

8.27. Explain how ceramics can be made tougher.

8.28. List and describe situations in which static fatigue can be important.

8.29. What properties are important in making heat-resistant ceramics for use on oven tops? Why?

8.30. A large variety of glasses are now available. Why is this so?

8.31. What is the difference between the structure of graphite and that of diamond? Is it important? Explain.

8.32. List and explain materials that are suitable for use as a coffee cup.

8.33. Aluminum oxide and PSZ are described as white in appearance. Can they be colored? If so, how would you accomplish this?

8.34. Why does the strength of a ceramic part depend on its size?

8.35. In old castles and churches in Europe, the glass windows display pronounced ripples and are thicker at the bottom than at the top. Explain.

8.36. Is a carbide an example of a composite material? Explain your answer.

8.37. Ceramics are hard and strong in both compression and shear. Why, then, are they not used as nails or other fasteners? Explain.

8.38. Perform an Internet search and determine the chemistry of glass used for (a) fiber-optic communication lines, (b) crystal glassware, and (c) high-strength glass fibers.

8.39. Investigate and list the ceramics used for high-temperature superconductor applications.

8.40. Explain why synthetic diamond gemstones are not appreciably less expensive than natural diamond gemstones.

8.41. Explain why ceramic glass is suitable for an electric stove cooktop.

Quantitative Problems

8.42. In a fully dense ceramic, $S_{\mathrm{ut,o}} = 250$ MPa and $E_o = 300$ GPa. What are these properties at 15% porosity for values of $n = 4$, 5, 6, and 7, respectively?

8.43. Plot the S_{ut}, E, and k values for ceramics as a function of porosity P, and describe and explain the trends that you observe in their behavior.

8.44. What would be the tensile strength and the modulus of elasticity of the ceramic in Problem 8.42 for porosities of 20% and 40%, for the four n values given?

8.45. Calculate the thermal conductivities for ceramics at porosities of 10%, 20%, and 40% for $k_o = 0.5$ W/mK.

8.46. A ceramic has $k_o = 0.75$ W/mK. If this ceramic is shaped into a cylinder with a porosity distribution given by $P = 0.1(x/L)(1 - x/L)$ where x is the distance from one end of the cylinder and L is the total cylinder length, plot the porosity as a function of distance, evaluate the average porosity, and calculate the average thermal conductivity.

8.47. It can be shown that the minimum weight of a column which will support a given load depends on the ratio of the material's stiffness to the square root of its density. Plot this property for a ceramic as a function of porosity.

Synthesis, Design, and Projects

8.48. Make a list of the ceramic parts that you can find around your house or in your car. Give reasons why those parts are made of ceramics.

8.49. Assume that you are working in technical sales and are fully familiar with all the advantages and limitations of ceramics. Which of the markets traditionally using nonceramic materials do you think ceramics can penetrate? What would you like to talk about to your potential customers during your sales visits? What questions do you think they may ask you about ceramics?

8.50. Describe applications in which a ceramic material with a near-zero coefficient of thermal expansion would be desirable.

8.51. The modulus of elasticity of ceramics is typically maintained at elevated temperatures. What engineering applications could benefit from this characteristic?

8.52. List and discuss the factors that you would take into account when replacing a metal component with a ceramic component in a specific product.

8.53. Obtain some data from the technical literature in the Bibliography, and quantitatively show the effects of temperature on the strength and the modulus of elasticity of several ceramics. Comment on how the shape of these curves differs from those for metals.

8.54. Conduct a literature search and write a three page paper summarizing the properties and potential applications of graphene.

8.55. It was noted in Section 8.4.1 that there are several basic types of glasses available. Make a survey of the technical literature, and prepare a table for these glasses, indicating various mechanical, physical, and optical properties.

8.56. Ceramic pistons are being considered for high-speed combustion engines. List the benefits and concerns that you would have regarding this application.

8.57. It has been noted that the strength of brittle materials (such as ceramics and glasses) is very sensitive to surface defects, such as scratches (known as *notch sensitivity*). Obtain several pieces of these materials, scratch them, and test them by carefully clamping them in a vise and bending them. Comment on your observations.

8.58. Electric space heaters for home use commonly utilize a ceramic filament as the heating element. List the required mechanical properties for this filament, explain why a ceramic is a suitable material, and perform an Internet search to determine the specific ceramic material actually utilized in this application.

8.59. Conduct a literature search and write a summary of the uses of graphene in modern batteries.

Chapter 9

Composite Materials: Structure, General Properties, and Applications

- With their high strength-to-weight and stiffness-to-weight ratios, composite materials are among the most important engineered materials.
- Composites are widely used as structural components, especially in the aerospace and automotive industries, where weight savings are a major consideration.
- This chapter describes the major types of composite materials, the characteristics of the commonly used reinforcing fibers, and their effect in improving mechanical properties.
- The role of the matrix is then described, and the three principal classes of matrix materials (plastic, metal, and ceramic) are examined.
- The chapter ends with the selection and applications of a variety of reinforced plastics and composites.

9.1 Introduction

A **composite material** is defined as a combination of two or more chemically distinct and insoluble phases with a recognizable interface, in such a manner that its properties and structural performance are superior to those of the constituents acting independently. These combinations are known as **polymer-matrix, metal-matrix**, and **ceramic-matrix composites**. As shown in Table 7.1, *fiber reinforcements* significantly improve the strength, stiffness, and creep resistance of plastics, particularly their strength-to-weight and stiffness-to-weight ratios. Composite materials have found increasingly wider applications in aircraft (Fig. 9.1), space vehicles, satellites, offshore structures, piping, electronics, automobiles, boats, and sporting goods.

The oldest example of composites, dating back to 4000 B.C., is the addition of *straw* to *clay* to make bricks for buildings. In this combination, the straws are the reinforcing fibers and the clay is the matrix. Another example of a composite material is reinforced concrete, developed in the 1800s. By itself, concrete is brittle and has little or no useful tensile strength; reinforcing steel rods (*rebar*) impart the necessary tensile strength to the concrete.

Composites include a wide variety of materials, such as cermets (Section 8.2.3), two-phase alloys (Section 4.2), natural materials such as wood and bone, and reinforced or combined materials such as steel-wire reinforced automobile tires or drive belts. This chapter describes the structure, properties, and applications of composite materials; processing and shaping of composite materials are described in Chapter 19.

9.2 The Structure of Reinforced Plastics

Reinforced plastics, also known as **polymer-matrix composites** (PMC) and **fiber-reinforced plastics** (FRP), consist of **fibers** (the discontinuous, or dispersed, phase) in a polymer **matrix** (the continuous phase), as

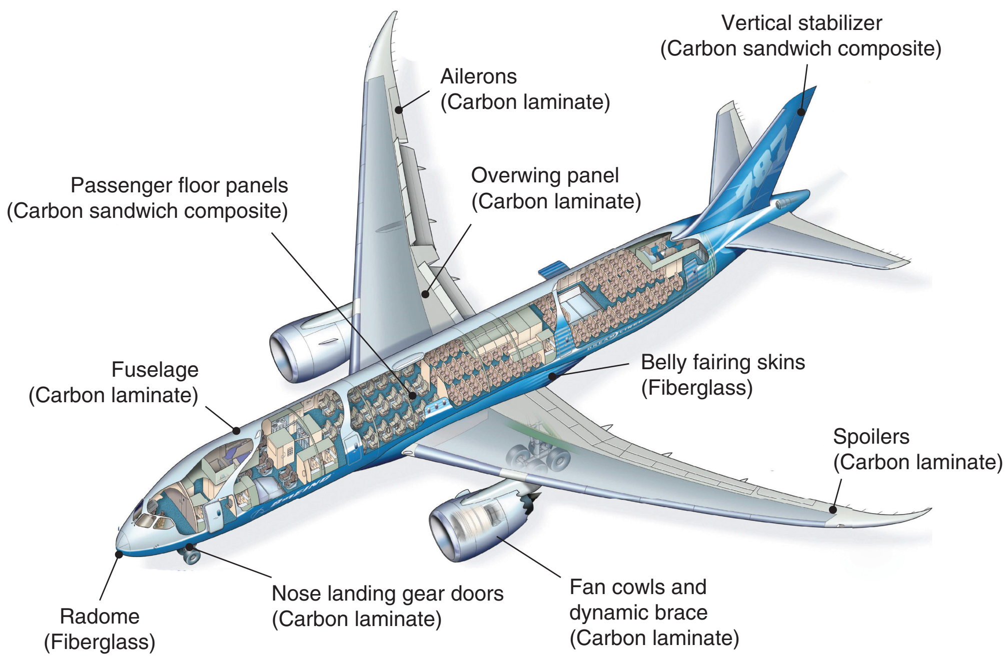


Figure 9.1: Application of advanced composite materials in the Boeing 787-D (Dreamliner), which is 50% composite material by weight. The reinforcement type is shown. *Source:* Graphic image Courtesy of FlightGlobal.

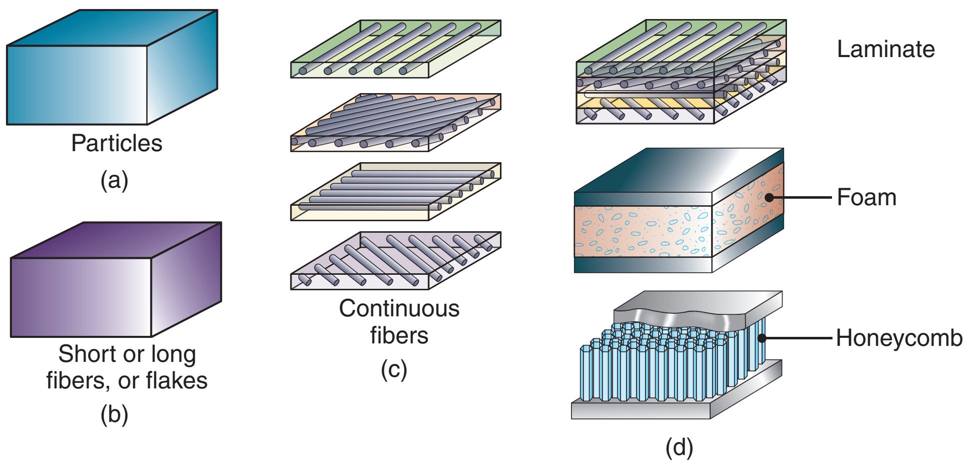


Figure 9.2: Schematic illustration of methods of reinforcing plastics (matrix) with (a) particles, (b) short or long fibers or flakes, and (c) continuous fibers. The laminate structures shown in (d) can be produced from layers of continuous fibers or sandwich structures using a foam or honeycomb core (see also Fig. 16.59).

shown in Fig. 9.2. These fibers are strong and stiff (Table 9.2) and have high specific strength (strength-to-weight ratio) and specific stiffness (stiffness-to-weight ratio), as shown in Fig. 9.3. In addition, reinforced-plastic structures have improved fatigue resistance, and higher toughness and creep resistance than those made of nonreinforced plastics.

The fibers in reinforced plastics have, by themselves, little structural value; they are stiff in their longitudinal direction but have no transverse stiffness or strength. Although the plastic matrix is less strong

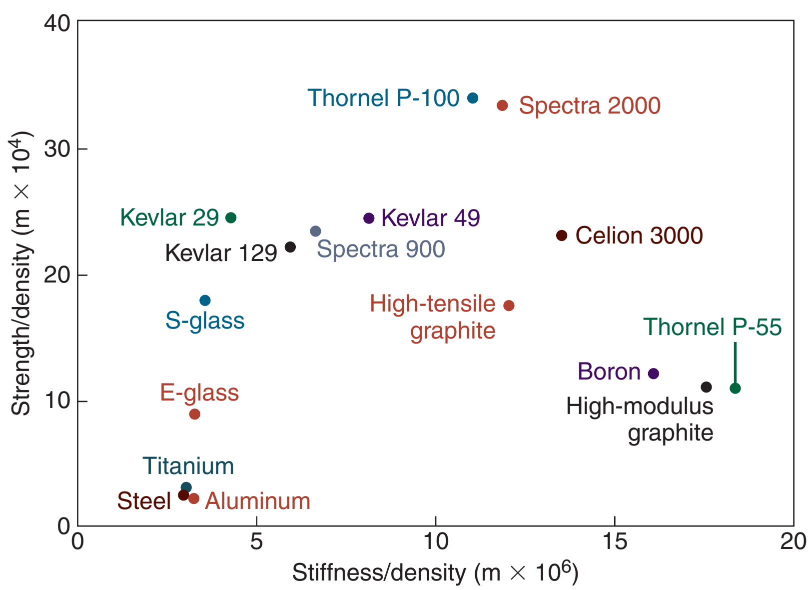


Figure 9.3: Specific tensile strength (tensile-strength-to-density ratio) and specific stiffness (modulus-of-elasticity-to-density ratio) for various fibers used in reinforced plastics. Note the wide range of specific strength and stiffness.

and less stiff than the fibers, it is tougher and often more chemically inert; thus, reinforced plastics combine the advantages of each of the two constituents. The percentage of fibers (by volume) in reinforced plastics typically ranges between 10 and 60%.

9.2.1 Reinforcing Fibers

Glass, carbon, ceramics, aramids, and boron are the most common reinforcing fibers for polymer-matrix composites (Table 9.1).

Glass Fibers. Glass fibers are the most widely used and the least expensive of all fibers. The composite material is called **glass-fiber reinforced plastic** (GFRP), and may contain between 30 and 60% glass fibers. The fibers are made by drawing molten glass through small openings in a platinum die (Section 18.3.4); the glass is then elongated, cooled, and wound on a roll. The fibers are later treated with *silane* (a silicon hydride), as described in Section 9.3. The principal types of glass fibers are:

- **E-type:** a calcium aluminoborosilicate glass, the type most commonly used
- **S-type:** a magnesia aluminosilicate glass, offering higher strength and stiffness, but at a higher cost
- **E-CR-type:** a high-performance glass fiber, with higher resistance to elevated temperatures and acid corrosion than does the E glass.

Carbon Fibers. Carbon fibers (Fig. 9.4a), although more expensive than glass fibers, have a combination of low density, high strength, and high stiffness; the composite is called **carbon-fiber reinforced plastic** (CFRP). Although the two words are often used interchangeably, the difference between *carbon* and *graphite* depends on the temperature at which it was processed and resulting microstructure. A typical carbon fiber contains amorphous (noncrystalline) carbon and graphite (crystalline carbon). These fibers are classified by their elastic modulus, which ranges from 35 to 800 GPa, as *standard, intermediate, high*, and *very high modulus*. Some trade names for carbon fibers are Celion and Thornel (Fig. 9.3). *Carbon nanotubes* have also been used as reinforcement in composite materials (Section 8.6.2).

All carbon fibers are made by **pyrolysis** of organic **precursors**. Pyrolysis is the process of inducing chemical changes by heat—for example, by burning a length of yarn, causing the material to carbonize and become black in color. A common precursor is *polyacrylonitrile* (PAN); *rayon* and *pitch* (the residue in

Table 9.1: Types and General Characteristics of Composite Materials.

Material	Characteristics
Fibers	
Glass	High strength, low stiffness, high density; lowest cost; E (calcium aluminoborosilicate) and S (magnesia aluminosilicate) types commonly used
Carbon	Available as high modulus or high strength; low cost; less dense than glass; sometimes used in combination with carbon nanotubes (see Section 8.6.2)
Boron	High strength and stiffness; highest density; highest cost; has tungsten filament at its center
Aramids (Kevlar)	Highest strength-to-weight ratio of all fibers; high cost
Other fibers	Nylon, silicon carbide, silicon nitride, aluminum oxide, boron carbide, boron nitride, tantalum carbide, steel, tungsten, molybdenum
Matrix materials	
Thermosets	Epoxy and polyester, with the former most commonly used; others are phenolics, fluorocarbons, polyethersulfone, silicon, and polyimides
Thermoplastics	Polyetheretherketone; tougher than thermosets, but lower resistance to temperature
Metals	Aluminum, aluminum–lithium, magnesium, and titanium; fibers are carbon, aluminum oxide, silicon carbide, and boron
Ceramics	Silicon carbide, silicon nitride, aluminum oxide, and mullite; fibers are various ceramics

Table 9.2: Typical Properties of Reinforcing Fibers.

Type	Tensile strength (MPa)	Elastic modulus (GPa)	Density (kg/m^3)	Relative cost
Boron	3500	380	2380	Highest
Carbon				
High strength	3000	275	1900	Low
High modulus	2000	415	1900	Low
Glass				
E-type	3500	73	2480	Lowest
S-type	4600	85	2540	Lowest
Kevlar				
29	2920	70.5	1440	High
49	3000	112.4	1440	High
129	3200	85	1440	High
Nextel				
312	1700	150	2700	High
610	2770	328	3960	High
Spectra				
900	2270	64	970	High
1000	2670	90	970	High
2000	3240	115	970	High
Alumina (Al_2O_3)	1900	380	3900	High
Silicon carbide	3500	400	3200	High

Note: These properties vary significantly depending on the material and method of preparation.

petroleum refining) also are used as precursors. With PAN, the fibers are partially cross-linked at a moderate temperature (in order to prevent melting during subsequent processing steps), and are elongated simultaneously. At this stage, the fibers are *carburized*, that is, they are exposed to elevated temperatures to expel the hydrogen (dehydrogenation) and the nitrogen (denitrogenation) from the PAN. The temperatures for carburizing range up to about 1500°C for carbon fibers and up to 3000°C for graphite fibers (graphitizing).

Conductive Graphite Fibers. These fibers are produced to make it possible to enhance the electrical and thermal conductivity of reinforced plastic components. The fibers are coated with a metal (usually nickel)

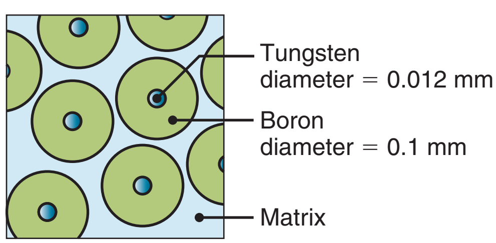


Figure 9.4: Cross section of boron fiber-reinforced composite material.

using a continuous electroplating process. The coating is typically 0.5-μm thick on a 7-μm-diameter graphite fiber core. Available in chopped or continuous form, the conductive fibers are incorporated directly into injection-molded plastic parts (Section 19.3). Applications include electromagnetic and radio-frequency shielding and lightning-strike protection.

Ceramic Fibers. Ceramic fibers are advantageous for high temperature applications and metal matrix composites (Section 9.5). The fibers have low elongation, low thermal conductivity, and good chemical resistance, in addition to being suitable for high-temperature applications. One family of ceramic fibers is *Nextel* (a trade name). These fibers are oval in cross section and consist of alumina, silica, and boric oxide. Typical mechanical properties are given in Table 9.2.

Polymer Fibers. Polymer fibers can be made of nylon, rayon, acrylics, or aramids; the most common are **aramid fibers**. Aramids (Section 7.6), such as **Kevlar**, are among the toughest fibers, with very high specific strength (Fig. 9.3). They can undergo some plastic deformation prior to fracturing, and thus have higher toughness than brittle fibers. However, aramids absorb moisture (*hygroscopy*), degrading their properties and complicating their application.

Another high-performance polyethylene fiber is *Spectra* (a trade name); it has ultra-high molecular weight and high molecular-chain orientation. Spectra, a bright white polyethylene, has better abrasion resistance and flexural-fatigue strength than aramid fibers, and at a similar cost. In addition, because of its lower density (970 kg/m^3), it has a higher specific strength and specific stiffness than aramid (Kevlar) fibers (Table 9.2). However, its low melting point and poor adhesion characteristics as compared to other polymers are major limitations to applications.

The manufacture of polymer fibers is described in Section 19.2.2.

Boron Fibers. These fibers consist of tungsten fibers with a layer of boron, deposited by chemical vapor-deposition techniques (Fig. 9.4b). Boron also can be deposited onto carbon fibers. These fibers have high strength and stiffness, both in tension and in compression, and resistance to high temperatures; however, because of the high density of tungsten, they are heavy and expensive.

Miscellaneous Fibers. Among other fibers used in composites are silicon carbide, silicon nitride, aluminum oxide, sapphire, steel, tungsten, molybdenum, boron carbide, boron nitride, and tantalum carbide. **Whiskers** also are used as reinforcing fibers (see also Section 22.10); they are tiny, needlelike single crystals that grow to 1–10 μm in diameter. They have high aspect ratios (ratio of fiber length to its diameter), ranging from 100 to 15,000. Because of their small size, whiskers are either free of imperfections or the imperfections they contain do not significantly affect their strength, which approaches the theoretical strength of the material. The elastic moduli of whiskers range between 400 and 700 GPa, and their tensile strength is on the order of 15 to 20 GPa, depending on the material.

9.2.2 Fiber Size and Length

Fibers are very strong and stiff in tension, because (a) the molecules in the fibers are oriented in the longitudinal direction and (b) their cross section is so small (usually less than 0.01 mm in diameter), that the probability is low for any significant defects to exist in the fiber. Glass fibers can have tensile strengths as high as 4600 MPa, whereas the strength of glass in bulk form (Section 8.4.2) is much lower (see Table 2.2).

Fibers are classified as **short** (**discontinuous**) or **long** (**continuous**). The designations short and long fiber are, in general, based on the following distinction: In a given type of fiber, if the mechanical properties improve as a result of increasing average fiber length, it is called a *short fiber*. If no such improvement in composite properties occurs, it is called a *long fiber*. Short fibers typically have aspect ratios between 20 and 60, and long fibers between 200 and 500.

Reinforcing elements in composites may also be in the shape of *chopped fibers, particles, flakes*, or as continuous *roving* (slightly twisted strands) fibers, *woven fabric* (similar to cloth), *yarn* (twisted strands), and *mats* of various combinations.

9.2.3 Matrix Materials

The matrix in reinforced plastics has three principal functions:

1. Support the fibers in place and transfer the stresses to them, so that the fibers can carry most of the load (see Example 9.1).

2. Protect the fibers against physical damage, and environmentally caused chemical degradation.

3. Slow the propagation of cracks in the composite by virtue of the higher ductility and toughness of the plastic matrix.

Matrix materials can be *thermoplastics* or *thermosets* such as epoxy, polyester, phenolic, fluorocarbon, polyethersulfone, or silicon. The most common are epoxies (80% of all reinforced plastics) and polyesters (less expensive than the epoxies). Polyimides, which resist exposure to temperatures in excess of 300°C, are available for use as a matrix with carbon fibers. Some thermoplastics, such as polyetheretherketone (PEEK), are also used as matrix materials. They generally have higher toughness than thermosets, but their resistance to temperature is lower, being limited to 100° to 200°C.

9.3 Properties of Reinforced Plastics

The mechanical and physical properties of reinforced plastics depend on the type, shape, and orientation of the reinforcing fiber, their length, and the volume fraction (percentage) of the reinforcing material. Short fibers are less effective than long fibers (Fig. 9.5), and their properties are strongly influenced by temperature and time under load. Long fibers transmit the load through the matrix better, and are less likely to pull out of the matrix (caused by shear failure of the fiber–matrix interface); thus, they are used in critical applications, particularly at elevated temperatures. The physical properties of reinforced plastics and their resistance to fatigue, creep, and wear depend greatly on the type and amount of reinforcement. Composites can be tailored to impart specific properties, such as permeability and dimensional stability, and make processing easier and reducing production costs.

Because the load is transmitted through the fiber–matrix interface, a critical factor in reinforced plastics is the strength of the interfacial bond. Weak bonding can cause **fiber pullout** and **delamination** of the composite, particularly under adverse environmental conditions. Adhesion at the interface can be improved by special surface treatments, such as coatings and coupling agents. Glass fibers, for example, are treated with **silane** for improved wetting and bonding between the fiber and the matrix. The importance of proper bonding can be appreciated by inspecting the fracture surfaces of reinforced plastics; note in Figs. 9.6a and b the separation between the fibers and the matrix.

Generally, the highest stiffness and strength in reinforced plastics are achieved when the fibers are aligned in the direction of the tension force; the composite is then highly anisotropic. As a result, properties such as creep resistance, thermal and electrical conductivity, and thermal expansion, also are anisotropic. The transverse properties of a unidirectionally reinforced structure are much lower than their longitudinal properties. For example, note how strong a fiber-reinforced packaging tape is when pulled in tension, yet how easily it can split when pulled in the width direction.

Because it is an *engineered material*, a part made of reinforced plastic can be given an optimal configuration for a specific service condition. If, for example, the part is to be subjected to forces in different directions, such as in thin-walled, pressurized vessels, (a) the fibers can be criss-crossed in the matrix or (b) the layers of fibers oriented in different directions can be built up into a laminate having improved properties in more than one direction (see *filament winding* and *tape laying*, Section 19.13).

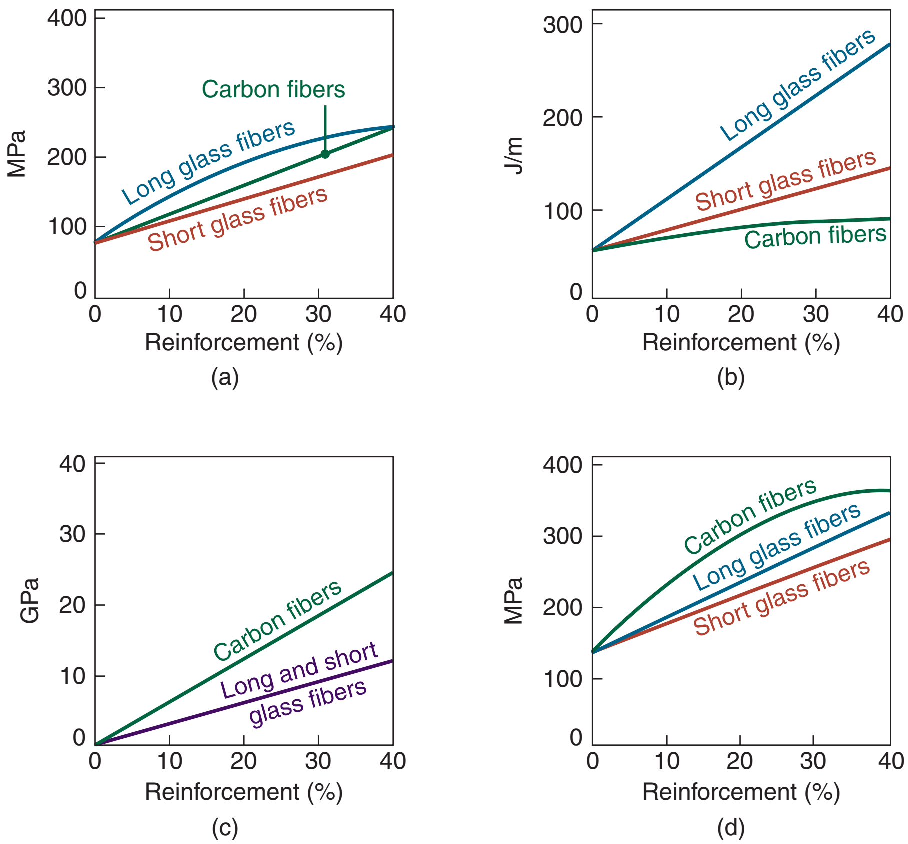


Figure 9.5: The effect of the type of fiber on various properties of fiber-reinforced nylon (6,6). *Source:* Courtesy of NASA.

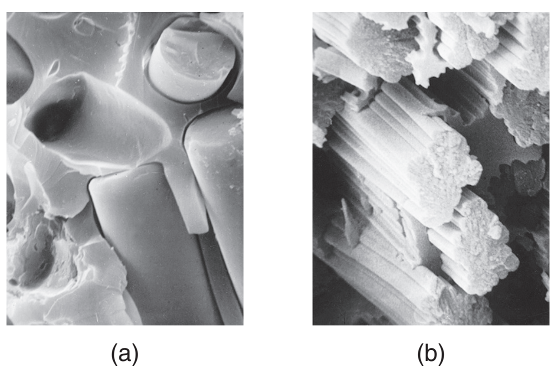


Figure 9.6: (a) Fracture surface of a glass fiber-reinforced epoxy composite. The fibers are 10 μm in diameter and have random orientation. (b) Fracture surface of a graphite fiber-reinforced epoxy composite. The fibers, 9 to 11 μm in diameter, are in bundles and are all aligned in the same direction. *Source:* After L.J. Broutman.

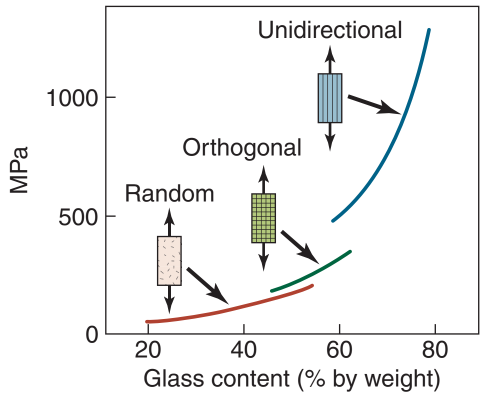


Figure 9.7: The tensile strength of glass-reinforced polyester as a function of fiber content and fiber direction in the matrix.

9.3.1 Strength and Elastic Modulus of Reinforced Plastics

The strength and elastic modulus of a reinforced plastic, with unidirectional fibers, can be determined in terms of the strengths and moduli of the fibers and the matrix, and in terms of the volume fraction of fibers in the composite. In the following equations, c refers to the composite, f to the fiber, and m to the matrix. The total load, P_c, on the composite is shared by the fiber (P_f) and the matrix (P_m). Thus,

$$P_c = P_f + P_m, \tag{9.1}$$

which can be written as

$$\sigma_c A_c = \sigma_f A_f + \sigma_m A_m, \tag{9.2}$$

where A_c, A_f, and A_m are the cross-sectional areas of the composite, the fiber, and the matrix, respectively; thus, $A_c = A_f + A_m$. Let's now denote x as the area fraction of the fibers in the composite. (Note that x also represents the volume fraction, because the fibers are uniformly longitudinal in the matrix.) Then Eq. (9.2) can be written as follows:

$$\sigma_c = x\sigma_f + (1-x)\sigma_m. \tag{9.3}$$

The fraction of the total load carried by the fibers can now be calculated. First, note that in the composite under a tensile load, the strains sustained by the fibers and the matrix are the same; that is, $e_c = e_f = e_m$. Next, recall from Eq. (2.3) that

$$e = \frac{\sigma}{E} = \frac{P}{AE}.$$

Consequently,

$$\frac{P_f}{P_m} = \frac{A_f E_f}{A_m E_m}. \tag{9.4}$$

Since the relevant quantities for a specific situation are known by using Eq. (9.1), the fraction P_f/P_c can be found. Then, using the foregoing relationships, the elastic modulus, E_c, of the composite can be calculated by replacing σ in Eq. (9.3) with E. Thus,

$$E_c = xE_f + (1-x)E_m. \tag{9.5}$$

Example 9.1 Calculation of Stiffness of a Composite and Load Supported by Fibers

Given: Assume that a graphite–epoxy reinforced plastic with longitudinal fibers contains 20% graphite fibers. The elastic modulus of the fibers is 300 GPa, and that of the epoxy matrix is 100 GPa.

Find: Calculate the elastic modulus of the composite and the fraction of the load supported by the fibers.

Solution: The data given are $x = 0.2$, $E_f = 300$ GPa, and $E_m = 100$ GPa. Using Eq. (9.5),

$$E_c = 0.2(300) + (1 - 0.2)100 = 60 + 80 = 140 \text{ GPa}.$$

From Eq. (9.4), the load fraction P_f/P_m is found to be

$$\frac{P_f}{P_m} = \frac{0.2(300)}{0.8(100)} = 0.75.$$

Because

$$P_c = P_f + P_m \qquad \text{and} \qquad P_m = \frac{P_f}{0.75},$$

it can be seen that,

$$P_c = P_f + \frac{P_f}{0.75} = 2.33P_f, \qquad \text{or} \qquad P_f = 0.43P_c.$$

Thus, the fibers support 43% of the load, even though they occupy only 20% of the cross-sectional area (and hence volume) of the composite.

9.4 Applications of Reinforced Plastics

The first engineering application of reinforced plastics was in 1907, for an acid-resistant tank made of a phenolic resin embedded with asbestos fibers. In the 1920s, *formica* (a trade name) was developed, and was used commonly for countertops. Epoxies were first used as a matrix material in the 1930s. Beginning in the 1940s, boats were made with fiberglass, and reinforced plastics were used for aircraft, electrical equipment, and sporting goods. Major developments in composites began in the 1970s, called **advanced composites**. Glass or carbon fiber-reinforced **hybrid** plastics were developed for high-temperature applications, with continuous use ranging up to about 300°C.

Reinforced plastics are typically used in commercial and military aircraft, rocket components, helicopter blades, automobile bodies, leaf springs, driveshafts, pipes, ladders, pressure vessels, sporting goods, helmets, boat hulls, and various other structures and components. About 50% (by weight) of the Boeing 787 Dreamliner is made of composites. By virtue of the resulting weight savings, reinforced plastics have reduced fuel consumption in aircraft by about 2%. The Airbus jumbo jet A380, with a capacity of up to 700 passengers, has horizontal stabilizers, ailerons, wing boxes and leading edges, secondary mounting brackets of the fuselage, and a deck structure made of composites with carbon fibers, thermosetting resins, and thermoplastics.

The contoured frame of the Stealth bomber is made of composites, consisting of carbon and glass fibers, epoxy-resin matrices, high-temperature polyimides, and other advanced materials. Boron fiber-reinforced composites are used in military aircraft, golf-club shafts, tennis rackets, fishing rods, and sailboards (Fig. 9.8). Another example is the development of a small, all-composite ship (twin-hull catamaran design) for the U.S. Navy, capable of speeds of 93 kph.

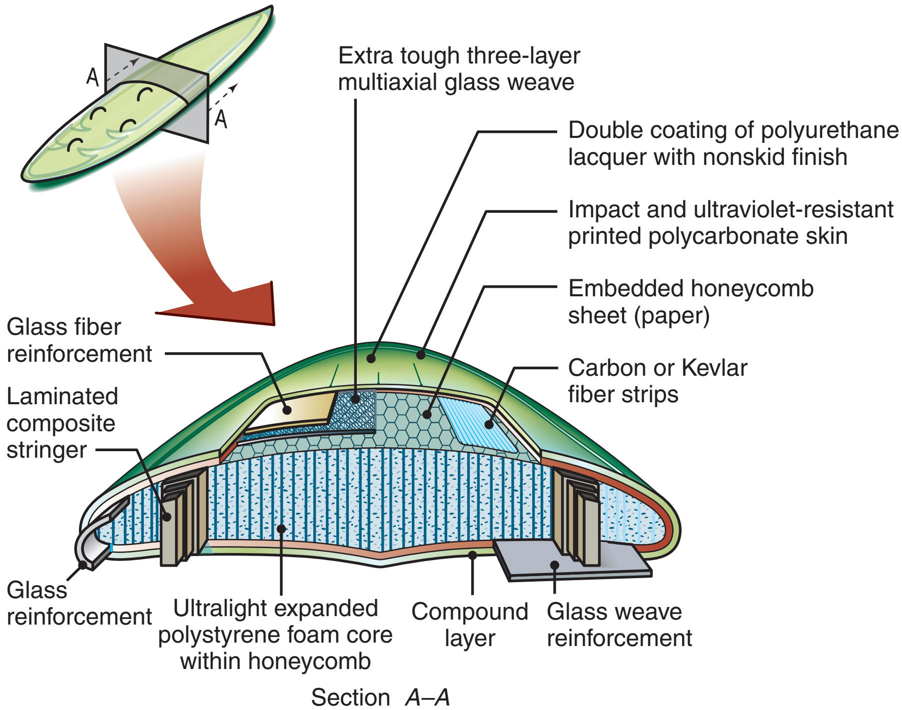


Figure 9.8: Cross section of a composite sailboard, an example of advanced materials construction. *Source:* K. Easterling, *Tomorrow's Materials*, 2nd ed., Institute of Metals, 1990.

Case Study 9.1 Composite Military Helmets and Body Armor

Personal protective equipment, in the form of body armor and helmets, are widespread for military and police applications. Body armor relies on high-strength woven fibers to prevent the penetration of projectiles. To stop a bullet, a composite material must first plastically deform it or flatten it, a process that occurs when the bullet's tip comes into contact with as many individual fibers of the composite as possible, without the fibers being pushed aside. The momentum associated with projectiles is felt by the user of the armor, but successful designs will contain bullets and shrapnel, preventing serious and fatal injuries.

There are two main types of body armor: (a) *soft armor*, which relies upon several layers of high-strength, woven fibers, and is designed mainly to contain handgun bullets and (b) *hard armor*, which utilizes a metal, ceramic, or polymer plate, in addition to the woven fiber; it is intended to provide protection against rifle rounds and shrapnel. A schematic of a body armor is shown in Fig. 9.9.

Several types of fiber meshes have been developed to be used in body armor applications. Different suppliers employ different combinations of fiber meshes, and may include additional layers to provide protection against blunt trauma. The first fiber used for flexible body armor was Kevlar 29 (an aramid), which has been improved through the years. Others include Kevlar 49, Kevlar 129, and Kevlar Protera—where tensile strength and energy-absorbing capabilities have been improved through the development of advanced spinning processes to produce the fibers. Aramid fibers are used very commonly in flexible body armor; other designs include over a thousand finely spun filaments that interact with each other to dissipate the impact energy.

Spectra fiber is used to make body armor; a layer of Spectra Shield composite consists of two unidirectional layers of Spectra fiber, arranged to cross each other at 0- and 90-degree angles, and held in place by a flexible resin. Both the fiber and the resin layers are sealed between two thin sheets of polyethylene film. Hard armor uses several designs, but typically it consists of steel, ceramic (usually aluminum oxide and silica), or polyethylene plates that are strategically located to prevent the penetration of ballistic particles to critical areas.

Source: Courtesy of Pinnacle Armor, Allied Signal Corp., and CGS Gallet SA.

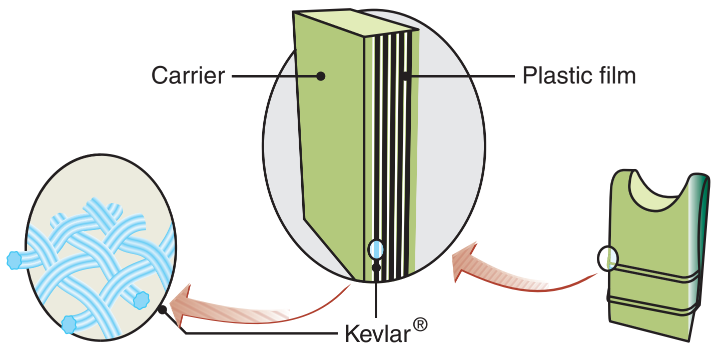


Figure 9.9: Schematic illustration of body armor, showing the layers of woven fibers.

9.5 Metal-matrix Composites

Matrix materials in *metal-matrix composites* (MMC) are usually aluminum, aluminum–lithium alloy (lighter than aluminum), magnesium, copper, titanium, or superalloys (Fig. 9.10). Fiber materials are graphite, aluminum oxide, silicon carbide, boron, molybdenum, or tungsten. The elastic modulus of nonmetallic fibers ranges between 200 and 400 GPa, with tensile strengths in the range from 2000 to 3000 MPa. The advantages

Figure 9.10: Brake disc: An example of metal-matrix composite parts. *Source:* Shutterstock/Andrei Kholmov.

Table 9.3: Metal-matrix Composite Materials and Applications.

Fiber	Matrix	Applications
Graphite	Aluminum	Satellite, missile, and helicopter structures
	Magnesium	Space and satellite structures
	Lead	Storage-battery plates
	Copper	Electrical contacts and bearings
Boron	Aluminum	Compressor blades and structural supports
	Magnesium	Antenna structures
	Titanium	Jet-engine fan blades
Alumina	Aluminum	Superconductor restraints in fission power reactors
	Lead	Storage-battery plates
	Magnesium	Helicopter transmission structures
Silicon carbide	Aluminum, titanium	High-temperature structures
	Superalloy (cobalt base)	High-temperature engine components
Molybdenum, tungsten	Superalloy	High-temperature engine components

of a metal matrix over a polymer matrix are higher elastic modulus, toughness, ductility, and higher resistance to elevated temperatures. The limitations are higher density and a greater difficulty in processing the composite parts. Typical compositions and applications for metal-matrix composites are given in Table 9.3.

Case Study 9.2 Aluminum-matrix Composite Brake Calipers

A trend in automobile design and manufacture is the increased effort toward lighter weight designs in order to realize improved performance and/or fuel economy; this trend can be seen in the development of metal-matrix composite brake calipers. Traditional brake calipers are made of cast iron, and can weigh around 3 kg each in a small car, and up to 14 kg in a truck. The cast-iron caliper could be redesigned completely, using aluminum to achieve weight savings, but it would require a larger volume since the nominal strength of aluminum is lower than the cast iron, and the space available between the wheel and the rotor is very constrained.

A new brake caliper was designed, using an aluminum alloy locally reinforced with precast composite inserts using continuous ceramic fiber. The fiber is a nanocrystalline alumina, with a diameter of 10 to 12 μm and a fiber volume fraction of 65%. The fiber and the composite properties are summarized in Table 9.4. Finite element analysis confirmed the placement and amount of reinforcement, leading to a design that exceeded minimum design requirements, and also matched deflections of cast-iron calipers in a packaging-constrained environment. The new brake caliper is shown in Fig. 9.11. It has a weight savings of 50%, with the added benefits of corrosion resistance and ease of recyclability.

Table 9.4: Summary of Fiber and Composite Properties for an Automotive Brake Caliper.

Property	Alumina fiber	Alumina-reinforced composite material
Tensile strength	3100 MPa	1.5 GPa
Elastic modulus	380 GPa	270 GPa
Density	3.9 g/cm^3	3.48 g/cm^3

Figure 9.11: Aluminum-matrix composite brake caliper using nanocrystalline alumina fiber reinforcement. *Source:* Courtesy of 3M Specialty Materials Division.

9.6 Ceramic-matrix Composites

Ceramic-matrix composites (CMC) are characterized by their resistance to high temperatures and corrosive environments. As described in Section 8.3.1, ceramics are strong and stiff; they resist high temperatures, but generally lack toughness. Matrix materials that retain their strength up to 1700°C are silicon carbide, silicon nitride, aluminum oxide, and mullite (a compound of aluminum, silicon, and oxygen). Carbon/carbon-matrix composites retain much of their strength up to 2500°C, although they lack oxidation resistance at high temperatures. Fiber materials are usually carbon and aluminum oxide. Applications of CMC include jet and automotive engine components, deep-sea mining equipment, pressure vessels, structural components, cutting tools, and dies for the extrusion and drawing of metals.

9.7 Other Composites

Composites also may consist of *coatings* of various types, applied on base metals or substrates (Chapter 34). Examples are:

- Plating of aluminum or other metals over plastics, generally for decorative purposes
- Enamels, for wear resistance, hardness, and decorative purposes
- Vitreous (glasslike) coatings on metal surfaces for various functional or ornamental purposes.

Composites are made into cutting tools and dies, such as cemented carbides and cermets. Other composites are grinding wheels, made of aluminum oxide, silicon carbide, diamond, or cubic-boron-nitride abrasive particles, all held together with various organic, inorganic, or metallic binders (Section 26.2.2). A composite used in machine-tool beds for some precision grinders consists of granite particles in an epoxy matrix; it has high strength, good vibration-damping capacity (better than gray cast iron), and good frictional characteristics.

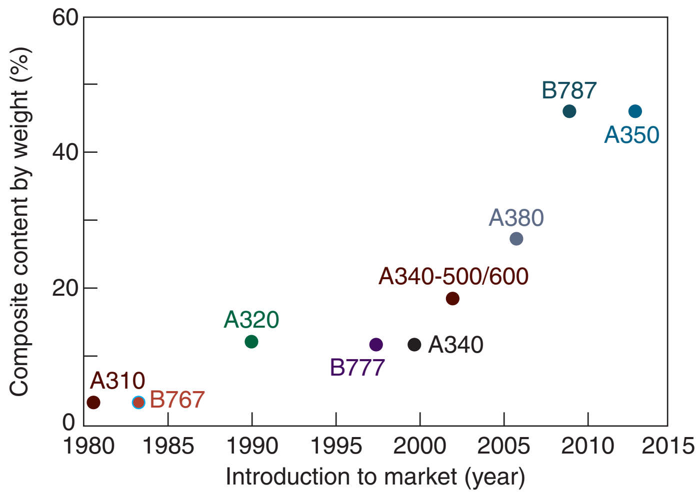


Figure 9.12: Composite content in selected commercial aircraft as a function of time (measured by date first introduced into the market). A = Airbus, B = Boeing.

Case Study 9.3 Composites in the Aircraft Industry

Any design changes that would lead to increased efficiency and fuel economy continue to be aggressively pursued by aircraft manufacturers. One area where this effect is most dramatic is the increased composite content in commercial aircraft, as shown in Fig. 9.12.

In addition to the amount of composite materials used, there are several design innovations in the types and applications of composite materials, including the following:

- GLARE is a GLAss-REinforced aluminum consisting of several layers of glass fiber-reinforced polymer and sandwiched between thin sheets of aluminum. It is used on the upper fuselage of the Airbus A380 and the leading edges of the tail plane, and has been credited with over 500 kg of weight savings, as compared to previously used materials. GLARE also provides improved fatigue strength and corrosion resistance.

- The Boeing 787 Dreamliner has an all-composite fuselage, constructed mainly from carbon-fiber reinforced plastic. In addition to weight savings, the fuselage is constructed in one piece and joined end to end, eliminating the need for an estimated 50,000 fasteners. Composites make up around 50% of the weight of the Dreamliner, as compared to 12% on the 777 aircraft, first introduced in 1994.

Summary

- Composites are an important class of engineered materials, with numerous attractive properties. Three major categories are fiber-reinforced plastics, metal-matrix composites, and ceramic-matrix composites. They have a wide range of applications in the aircraft, aerospace, and transportation industries, sporting goods, and structural components.

- In fiber-reinforced plastics, the fibers are usually glass, graphite, aramids, or boron. Polyester and epoxies commonly are used as the matrix material. These composites have particularly high toughness and high strength-to-weight and stiffness-to-weight ratios.
- In metal-matrix composites, the fibers are graphite, boron, aluminum oxide, silicon carbide, molybdenum, or tungsten. Matrix materials generally consist of aluminum, aluminum–lithium alloy, magnesium, copper, titanium, or superalloys.
- In ceramic-matrix composites, the fibers are usually carbon and aluminum oxide, and the matrix materials are silicon carbide, silicon nitride, aluminum oxide, carbon, or mullite (a compound of aluminum, silicon, and oxygen).
- In addition to the type and quality of the materials used, important factors in the structure and properties of composite materials are the size and length of the fibers, their volume percentage compared with that of the matrix, the strength of the bond at the fiber–matrix interface, and the orientation of the fibers in the matrix.

Key Terms

Advanced composites
Ceramic matrix
Composite materials
Delamination
Engineered materials
Fiber pullout
Fibers
Hybrid
Matrix
Metal matrix
Polymer matrix
Precursor
Pyrolysis
Reinforced plastics
Silane
Whiskers

Bibliography

Agarwal, B.D., Broutman, L.J., and Chandrashekhara, K., **Analysis and Performance of Fiber Composites**, 3rd ed., Wiley, 2006.
Antoniac, I.V. (ed.), **Handbook of Bioceramics and Biocomposites**, Springer, 2016.
ASM Handbook, Vol. 21: **Composites**, ASM International, 2001.
Bansal, N.P., and Lamon, G., **Ceramic Matrix Composites**, Wiley, 2014.
Campbell, F.C., **Structural Composite Materials**, ASM International, 2010.
Cantor, B., Dunne, F.P.E., and Stone, I.C. (eds.), **Metal and Ceramic Matrix Composites**, CRC Press, 2003.
Chawla, N., and Chawla, K.K., **Composite Materials: Science and Engineering**, 2nd ed., Springer, 2013.
—Metal Matrix Composites, Springer, 2006.
Chung, D.D.L., **Composite Materials: Science and Applications**, Butterworth-Heinemann, 2nd ed., 2016.
Daniel, I.M., and Ishai, O., **Engineering Mechanics of Composite Materials**, 2nd ed., Oxford, 2005.
Davim, J.P., and Jarzabek, D.M., **Micromechanics of Metal Matrix Composites**, Springer, 2016.
Dvorak, G., **Micromechanics of Composite Materials**, Springer, 2013.
Gay, D., **Composite Materials: Design and Applications**, 3rd ed., CRC Press, 2014.
Gibson, R.F., **Principles of Composite Material Mechanics**, 4th ed., CRC Press, 2016.
Krenker, W. (ed.), **Ceramic Matrix Composites: Fiber Reinforced Materials and Their Applications**, Wiley, 2008.
Loos, M., **Carbon Nanotube Reinforced Composites**, Elsevier, 2013.

Low, I.M. (ed.), **Advances in Ceramic Matrix Composites**, Woodhead Publishing, 2013.
Nishida, Y., **Introduction to Metal Matrix Composites**, Springer, 2013.
Sapuan, S.M., **Composite Materials: Concurrent Engineering Approach**, Butterworth Heinemann, 2017.
Shimpi, N.G., **Biodegradable and Biocompatible Polymer Composites**, Woodhead Publishing, 2017.
Strong, A.B., **Fundamentals of Composites Manufacturing: Materials, Methods and Applications**, 2nd ed., Society of Manufacturing Engineers, 2007.
Zweben, C.H., and Beaumont, P. (eds.), **Comprehensive Composite Materials II**, Elsevier, 2017.

Review Questions

9.1. Distinguish between composites and metal alloys.

9.2. What is a fiber? What is a matrix?

9.3. Describe the functions of the matrix and the reinforcing fibers. What fundamental differences are there in the characteristics of the two materials?

9.4. Name the reinforcing fibers generally used to make composites. Which type of fiber is the strongest? Which type is the weakest?

9.5. What is the range in length and diameter of typical reinforcing fibers?

9.6. List the important factors that determine the properties of reinforced plastics.

9.7. Comment on the advantages and limitations of metal-matrix composites, reinforced plastics, and ceramic-matrix composites.

9.8. What are the most commonly used matrix materials? Why?

9.9. Describe the advantages of hybrid composites over other composites.

9.10. What material properties are improved by the addition of reinforcing fibers?

9.11. Describe the purpose of the matrix material.

9.12. What are the most common types of glass fibers?

9.13. Explain the difference between a carbon fiber and a graphite fiber.

9.14. How can a graphite fiber be made electrically and thermally conductive?

9.15. What is a whisker? What is the difference between a whisker and a fiber?

9.16. Explain the composition of boron fibers. Why are they heavy?

9.17. Give a succinct definition of fiber, yarn, and fabric.

Qualitative Problems

9.18. How do you think the use of straw mixed with clay originally came about in making brick for dwellings?

9.19. What products have you personally seen that are made of reinforced plastics? How can you tell?

9.20. Describe applications that are not well suited for composite materials.

9.21. Is there a difference between a composite material and a coated material? Explain.

9.22. Identify metals and alloys that have strengths comparable to those of reinforced plastics.

9.23. What limitations or disadvantages do composite materials have? What suggestions would you make to overcome the limitations?

9.24. Give examples of composite materials other than those stated in this chapter.

9.25. Explain why the behavior of the materials depicted in Fig. 9.5 is as shown.

9.26. Explain why fibers are so capable of supporting a major portion of the load in composite materials.

9.27. Do metal-matrix composites have any advantages over reinforced plastics? Explain.

9.28. Give reasons for the development of ceramic-matrix composites. Name some applications and explain why they should be effective.

9.29. Explain how you would go about determining the hardness of reinforced plastics and of composite materials. Are hardness measurements on these types of materials meaningful? Does the size of the indentation make any difference? Explain.

9.30. How would you go about trying to determine the strength of a fiber?

9.31. Glass fibers are said to be much stronger than bulk glass. Why is this so?

9.32. Describe situations in which a glass could be used as a matrix material.

9.33. When the American Plains states were settled, no trees existed for the construction of housing. Pioneers cut bricks from sod—basically, prairie soil as a matrix and grass and its root system as reinforcement. Explain why this approach was successful. Also, if you were a pioneer, would you stack the bricks with the grass horizontally or vertically? Explain.

9.34. By incorporating small amounts of a blowing agent, it is possible to manufacture hollow polymer fibers with gas cores. List possible applications for such fibers.

9.35. Referring to Fig. 9.2c, would there be an advantage in using layers of cloth (woven fibers) instead of continuous fiber stacks without weaving? Explain.

9.36. Is it possible to design a composite material that has a Poisson's ratio of zero in a desired direction? Explain. Can a composite material be designed that has a thermal conductivity of zero in a desired direction? Explain.

Quantitative Problems

9.37. Calculate the average increase in the properties of the plastics given in Table 7.1 as a result of their reinforcement, and describe your observations.

9.38. In Example 9.1, what would be the percentage of the load supported by the fibers if their strength were 1000 MPa and the matrix strength were 250 MPa? What would be the answer if the fiber stiffness were doubled and the matrix stiffness were halved?

9.39. Calculate the percent increase in the mechanical properties of reinforced nylon from the data shown in Fig. 9.5.

9.40. Calculate the elastic modulus and load supported by fibers in a composite with an epoxy matrix ($E = 100$ GPa), made up of 30% fibers made of (a) high-modulus carbon fiber and (b) Kevlar 29 fibers.

9.41. For a composite material consisting of high modulus carbon fibers ($E = 415$ GPa) and an epoxy matrix ($E = 100$ GPa), determine the volume fraction of fibers needed to produce a composite material with a stiffness equal to that of steel.

9.42. Plot E/ρ and $E/\rho^{0.5}$ for the composite materials listed in Table 9.2, and compare your results with the properties of the materials described in Chapters 4 through 8 (see also Table 9.1).

9.43. Calculate the stress in the fibers and in the matrix in Example 9.1. Assume that the cross-sectional area is 300 m^2 and $P_c = 2250$ N.

9.44. Repeat the calculations in Example 9.1 (a) if Nextel 610 fiber is used and (b) if Spectra 2000 is used.

9.45. Refer to the properties listed in Table 7.1. If acetal is reinforced with E-type glass fibers, what is the range of fiber content in glass-reinforced acetal?

9.46. Plot the elastic modulus and strength of an aluminum metal-matrix composite with high-modulus carbon fibers as a function of fiber content.

9.47. For the data in Example 9.1, what should be the fiber content so that the fibers and the matrix fail simultaneously? Use an allowable fiber stress of 250 MPa and a matrix strength of 50 MPa.

9.48. It is desired to obtain a composite material with a target stiffness of 5 GPa. If a high strength carbon fiber is to be used, determine the required fiber volume if the matrix is (a) nylon, (b) polyester, (c) acetal, and (d) polyethylene.

9.49. A rectangular cantilever beam 150 mm high, 30 mm wide, and 1.5 m long is subjected to a concentrated load of 60 kg at its end. (a) Consider a polymer reinforced by high modulus carbon fibers, with a fiber volume ratio of $x = 15\%$. What is the maximum deflection of the beam if the matrix material is polyester? (b) Obtain the deflection of the beam if aluminum or steel was used, using the same beam dimensions. (c) What fiber volume ratio is needed to produce the same deflection as the aluminum or steel beams? (d) Determine the weight of the beams considered in parts (b) and (c) and compare them.

9.50. In Example 9.1, assume the strength of the fibers is 2800 MPa, and the load to be supported is 1.5 kN. (a) If the length of the tension member is 1.5 m, what is the weight of the composite?

9.51. In Example 9.1, graphite fibers are used to reinforce the composite material. If the fibers were produced from a different form of carbon, namely diamond, what percentage of load would be carried by the fibers? What volume percentage of fibers would be needed to maintain 43% of the load support by the fibers?

9.52. Consider a composite consisting of reinforcing fibers ($E_f = 300$ GPa) in an epoxy matrix ($E = 100$ GPa). If the allowable fiber stress is 225 MPa and the matrix strength is 90 MPa, what should be the fiber content so that the fibers and matrix fail simultaneously?

Synthesis, Design, and Projects

9.53. What applications for composite materials can you think of other than those given in Section 9.4? Why do you think your applications would be suitable for these materials?

9.54. Using the information given in this chapter, develop special designs and shapes for possible new applications of composite materials.

9.55. Would a composite material with a strong and stiff matrix and a soft and flexible reinforcement have any practical uses? Explain.

9.56. Make a list of products for which the use of composite materials could be advantageous because of their anisotropic properties.

9.57. Inspect Fig. 9.1 and explain what other components of an aircraft, including the cabin, could be made of composites.

9.58. Name applications in which both specific strength and specific stiffness are important.

9.59. What applications for composite materials can you think of in which high thermal conductivity would be desirable? Explain.

9.60. As with other materials, the mechanical properties of composites are obtained by preparing appropriate specimens and then testing them. Explain what problems you might encounter in preparing specimens for testing in tension. Suggest methods for making appropriate specimens, including fashioning their shape and how they would be clamped into the jaws of testing machines.

9.61. Developments are taking place in techniques for three-dimensional reinforcement of composites. Describe (a) applications in which strength in the thickness direction of the composite is important and (b) your ideas on how to achieve this strength. Include simple sketches of the structure utilizing such reinforced plastics.

9.62. Design and describe a test method to determine the mechanical properties of reinforced plastics in their thickness direction. (Note, for example, that plywood is not particularly strong in its thickness direction.)

9.63. As described in this chapter, reinforced plastics can be adversely affected by the environment—in particular, moisture, chemicals, and temperature variations. Design and describe test methods to determine the mechanical properties of composite materials subjected to these environmental conditions.

9.64. Comment on your observations on the design of the sailboard illustrated in Fig. 9.8.

9.65. Make a survey of various sports equipment and identify the components made of composite materials. Explain the reasons for and the advantages of using composites in these specific applications.

9.66. Several material combinations and structures were described in this chapter. In relative terms, identify those that would be suitable for applications involving one of the following: (a) very low temperatures, (b) very high temperatures, (c) vibrations, and (d) high humidity.

9.67. Obtain a textbook on composite materials and investigate the effective stiffness of a continuous fiber-reinforced polymer. Plot the stiffness of such a composite as a function of orientation with respect to the fiber direction.

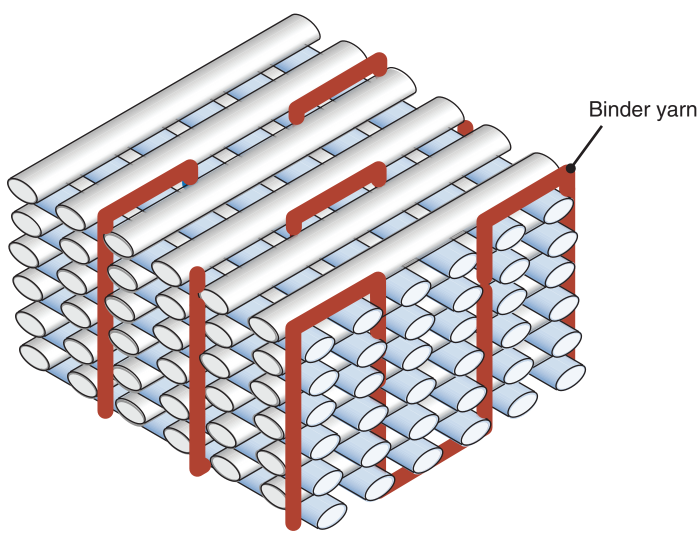
9.68. Derive a general expression for the coefficient of thermal expansion for a continuous fiber-reinforced composite in the fiber direction.

9.69. It is possible to make fibers or whiskers with a varying cross section, or a "wavy" fiber. What advantages would such fibers have?

9.70. Describe how you can produce some simple composite materials using raw materials available around your home. Explain.

9.71. *Gel spinning* is a specialized process used in making fibers with high strength or special properties. Search the technical literature and write a brief paper on this subject.

9.72. The sketch shows a section of a three-dimensional weave that uses a binder yarn to tie layers of fibers together. Conduct a literature search and determine the advantages and disadvantages of using three-dimensional weaves as reinforcements in composite materials.



9.73. Review the functions of the matrix material in Section 9.2.3. Could diamond be a useful matrix material in a fiber-reinforced composite material? Explain.

9.74. Glass fibers are used as reinforcements in composite materials. Is there any benefit in using the glass fibers to transmit light, as in a fiber optic cable? List potential applications for this feature.

PART II
Metal-casting Processes and Equipment

Several methods are available to shape metals into products. One of the oldest processes is **casting**, which basically involves pouring molten metal into a mold cavity; upon solidification, the metal takes the shape of the cavity. Two examples of cast parts are shown in Fig. II.1.

Casting was first used around 4000 B.C. to make ornaments, arrowheads, and various other simple objects. The process is now capable of producing intricate shapes, in one piece, and including those with internal cavities, such as engine blocks. Figure II.2 shows cast components in a typical automobile, a product that was used in the introduction to Part I to illustrate the selection and use of a variety of materials. The casting processes developed over the years are shown in Fig. II.3.

As in all manufacturing operations, each casting process has its own characteristics, applications, advantages, limitations, and costs involved. Casting is most often selected over other manufacturing methods for the following reasons:

- Casting can produce complex shapes and can incorporate internal cavities or hollow parts.
- Very large parts can be produced in one piece.
- Casting can utilize materials that are difficult or uneconomical to process by other methods, such as hard metals that are difficult to machine or plastically deform.
- The casting process is less expensive than other manufacturing processes for the particular application being considered.

(a) (b)

Figure II.1: (a) Examples of cast parts. (b) A tree of rings produced through investment casting. *Source:* (a) Shutterstock/Mr.1 (b) Courtesy of Romanoff, Inc.

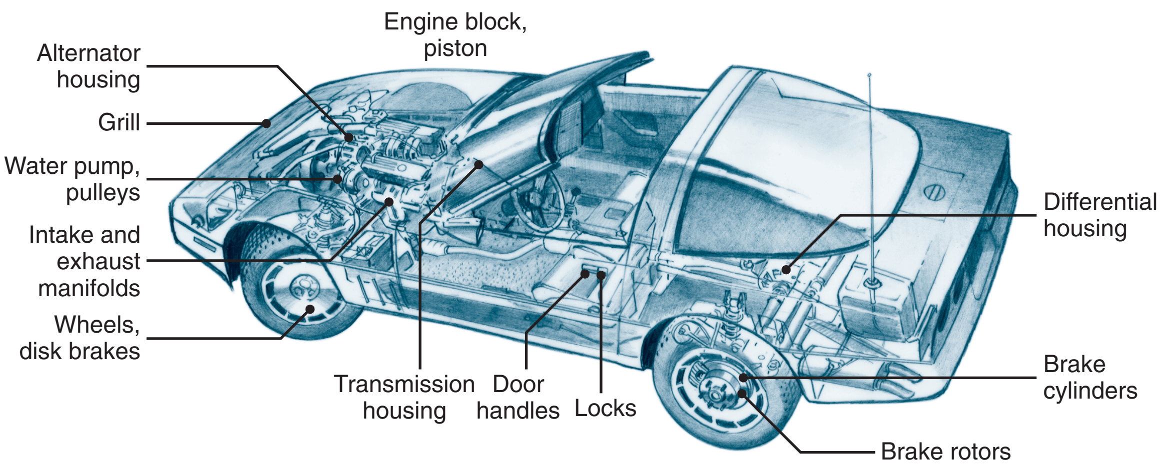


Figure II.2: Cast parts in a typical automobile.

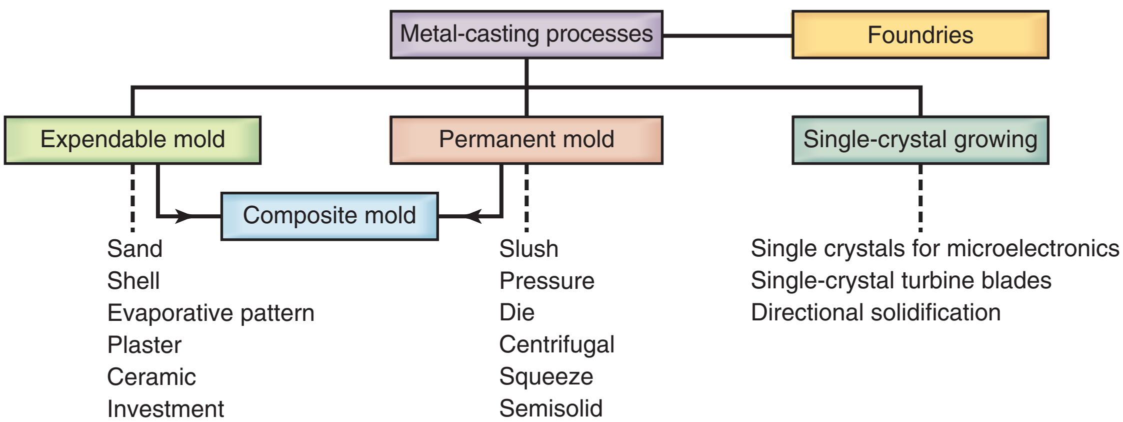


Figure II.3: Outline of metal-casting processes described in Part II.

Almost all metals can be cast in the *final shape* desired, or nearly so, often requiring only minor finishing operations. This capability places casting among the most important *net-shape manufacturing* technologies, along with net-shape forging (Chapter 14), stamping of sheet metal (Chapter 16), and powder metallurgy and metal-injection molding (Chapter 17). With modern processing techniques and control of chemical composition, mechanical properties of castings can equal those made by other manufacturing processes.

Chapter 10

Fundamentals of Metal Casting

Example:

- First used about 6000 years ago, casting continues to be an important manufacturing process for producing very small, very large, and complex parts.
- The first topic described is solidification of molten metals, including the differences between solidification of pure metals and alloys.
- Fluid flow in casting is then described, with Bernoulli's and the continuity equations being applied to establish a framework for analyzing molten metal flow into the cavities of a mold.
- The importance of turbulent versus laminar flow is introduced.
- Heat transfer and shrinkage of castings are also described, including Chvorinov's rule for solidification time.
- The chapter ends with a description of the causes of porosity in castings and common methods of reducing them to improve the properties of castings.

10.1 Introduction

The **casting** process basically involves (a) pouring molten metal into a mold containing a cavity that produces the desired part shape, (b) allowing it to solidify, and (c) removing the part from the mold. As with all other manufacturing processes, an understanding of the underlying science is essential for producing high quality, economical castings, and for establishing proper techniques for mold design and casting practice.

Important considerations in casting operations are:

- Flow of the molten metal into the mold cavity, and design of gating systems or pathways for molten metal to fill the cavity
- Solidification and cooling of the metal in the mold
- Influence of the mold material.

This chapter describes relationships among various relevant factors involved in casting. The flow of molten metal into the mold cavity is first described in terms of mold design and fluid-flow characteristics. Solidification and cooling of metals in the mold are affected by several factors, including the metallurgical and thermal properties of the metal and the type of mold because it affects the rate of cooling. The chapter ends with a description of the factors influencing defect formation in castings.

Metal-casting processes, design considerations, and casting materials are described in Chapters 11 and 12. The casting of ceramics and plastics, which involve methods and procedures somewhat similar to those for metal, are described in Chapters 18 and 19, respectively.

10.2 Solidification of Metals

After molten metal is poured into a **mold**, a sequence of events takes place during solidification and cooling of the metal to ambient temperature. These events greatly influence the size, shape, uniformity, and chemical composition of the grains formed throughout the casting, which, in turn, influence the overall properties of the casting. The significant factors affecting these events are the type of metal cast, the thermal properties of both the metal and the mold, the geometric relationship between volume and surface area of the casting, and the shape of the mold.

10.2.1 Pure Metals

Because pure metal has a clearly defined melting, or freezing, point, it solidifies at a constant temperature, as shown in Fig. 10.1. Pure aluminum, for example, solidifies at 660°C, iron at 1537°C, and tungsten at 3410°C (see also Table 3.1). After the molten metal temperature drops to its freezing point, its temperature remains constant while the *latent heat of fusion* is given off. The *solidification front* (the solid–liquid interface) moves through the molten metal from the mold walls in toward the center. The solidified metal, now called the *casting*, is then removed from the mold and allowed to cool to ambient temperature.

As shown in Fig. 10.1b and described in greater detail in Section 10.5.2, metals generally shrink when they solidify (Table 10.1) and shrink further while cooling. This behavior is an important consideration, because shrinkage can lead to microcracking and associated porosity, which can adversely affect the mechanical properties of the casting.

As an example of the grain structure that develops in a casting, Fig. 10.2a shows a cross section of a box-shaped mold. At the mold walls, which are at ambient temperature at first or typically are much cooler than the molten metal, the metal cools rapidly, producing a solidified **skin**, or *shell*, of fine equiaxed grains. The grains generally grow in a direction opposite to that of the heat transfer out through the mold. Those grains that have favorable orientation grow preferentially, and are called **columnar grains** (Fig. 10.3). Those grains that have substantially different orientations are blocked from further growing. As the driving force of the heat transfer decreases away from the mold walls, the grains become equiaxed and coarse. This sequence

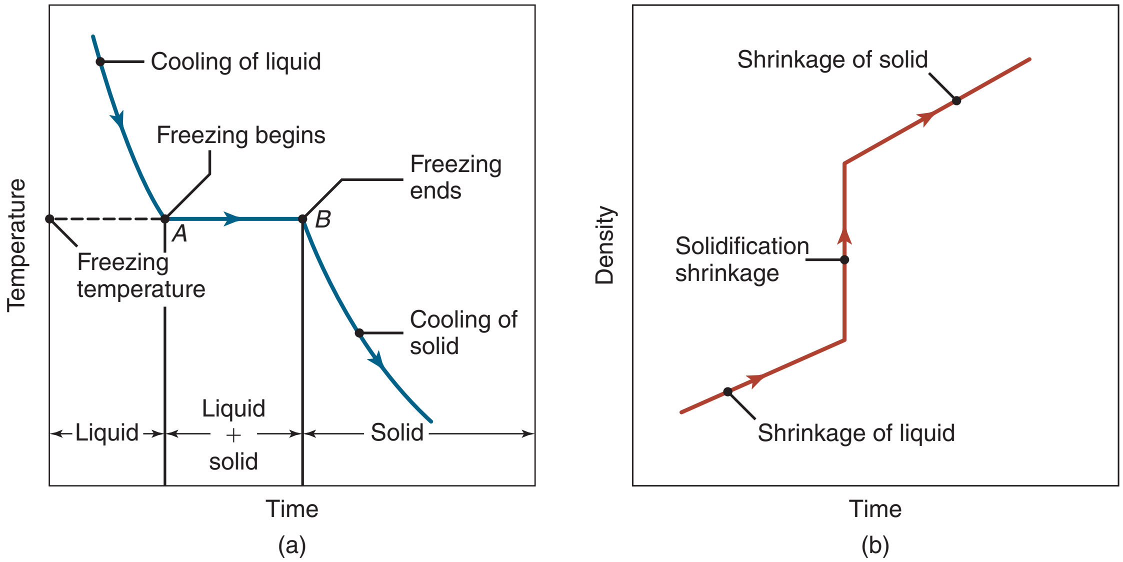


Figure 10.1: (a) Temperature as a function of time for the solidification of pure metals; note that freezing takes place at a constant temperature. (b) Density as a function of time.

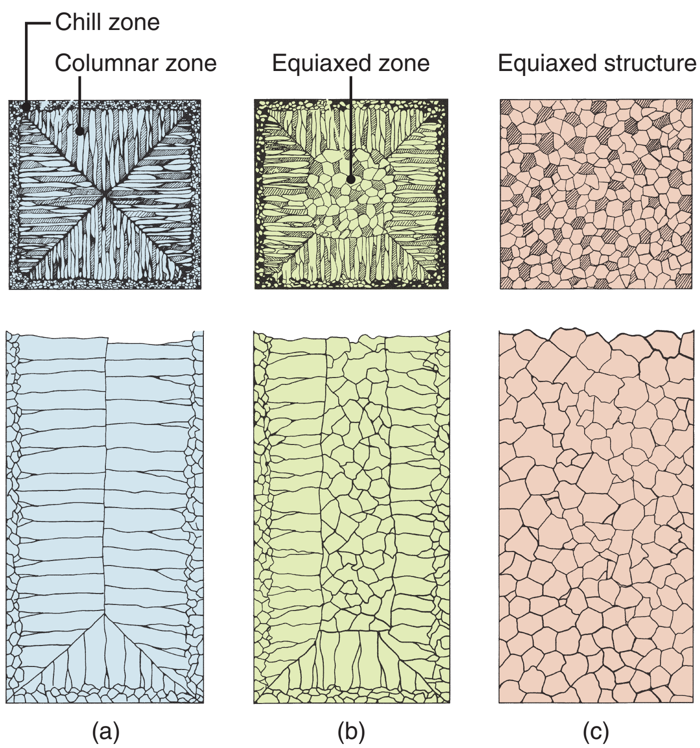


Figure 10.2: Schematic illustration of three cast structures of metals solidified in a square mold: (a) pure metals; (b) solid–solution alloys; and (c) structure obtained by using nucleating agents. *Source:* After G.W. Form, J.F. Wallace, J.L. Walker, and A. Cibula.

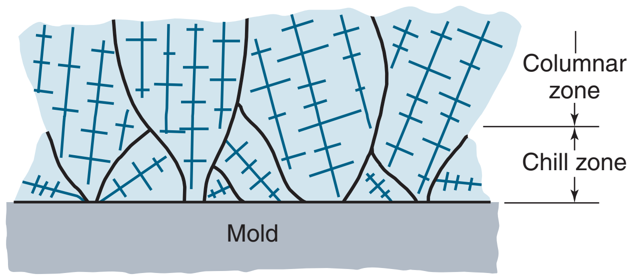


Figure 10.3: Development of a preferred texture at a cool mold wall; note that only favorably oriented grains grow away from the surface of the mold.

of grain development is known as **homogenous nucleation**, meaning that the grains (crystals) grow upon themselves, starting at the mold wall.

10.2.2 Alloys

Solidification in alloys begins when the temperature drops below the *liquidus*, T_L, and is complete when it reaches the *solidus*, T_S (Fig. 10.4). Within this temperature range, the alloy is in a *mushy* or *pasty* state, consisting of columnar **dendrites** (from the Greek *dendron*, meaning akin to, and *drys*, meaning tree). Note in the figure that the spaces between the dendrite arms are taken up by the liquid metal. Dendrites have three-dimensional *arms* and branches (*secondary arms*), which eventually interlock, as can be seen in Fig. 10.5. The study of dendritic structures, although complex, is important, because such structures can contribute to various detrimental factors, including compositional variations, segregation, and microporosity within a cast part.

The width of the **mushy zone**, where both liquid and solid phases are present, is an important factor during solidification. This zone is described in terms of a temperature difference, known as the **freezing range**, as

$$\text{Freezing range} = T_L - T_S. \tag{10.1}$$

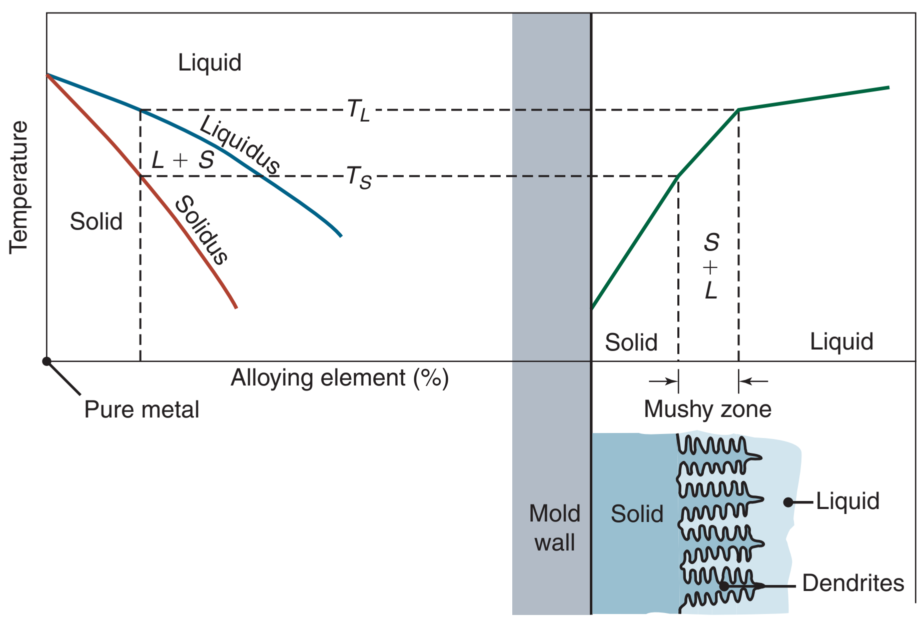


Figure 10.4: Schematic illustration of alloy solidification and temperature distribution in the solidifying metal. Note the formation of dendrites in the mushy zone.

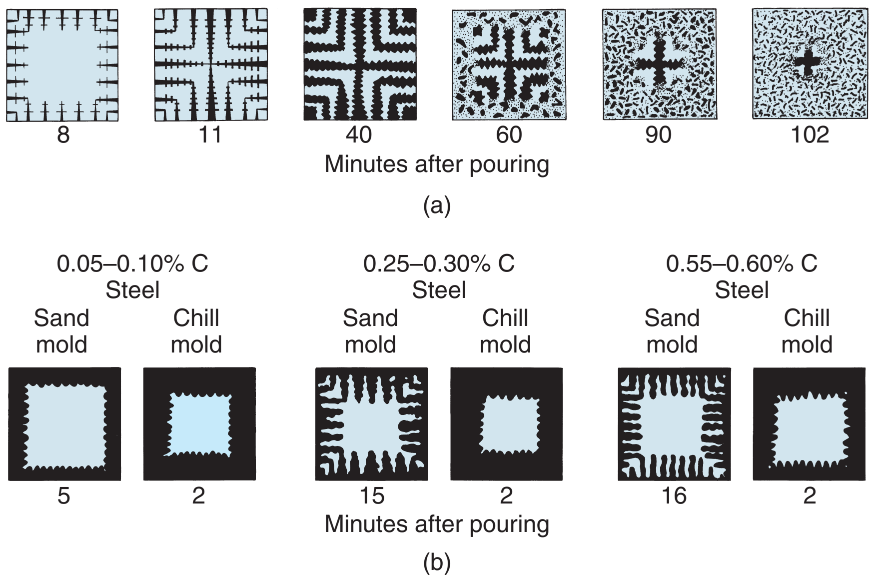


Figure 10.5: (a) Solidification patterns for gray cast iron in a 180-mm square casting. Note that after 11 minutes of cooling, dendrites begin to reach each other, but the casting is still mushy throughout. It takes about 2 hours more for this casting to solidify completely. (b) Solidification of carbon steels in sand and chill (metal) molds. Note the difference in solidification patterns as the carbon content of the metal increases. *Source:* After H.F. Bishop and W.S. Pellini.

It can be noted in Fig. 10.4 that pure metals have a freezing range that approaches zero and that the solidification front moves as a plane without developing a mushy zone. Eutectics (Section 4.3) solidify in a similar manner, with an essentially plane front; the structure developed upon solidification depends on the composition of the eutectic. In alloys with a nearly symmetrical phase diagram (see Fig. 4.4), the structure is generally lamellar, with two or more solid phases present, depending on the alloy system. When the volume fraction of the minor phase of the alloy is less than about 25%, the structure generally becomes fibrous. These conditions are particularly significant for cast irons.

For alloys, a *short freezing range* generally involves a temperature difference of less than 50°C, and a *long freezing range* more than 110°C. Ferrous castings typically have narrow mushy zones, whereas aluminum and magnesium alloys have wide mushy zones; consequently, these alloys are in a mushy state throughout most of their solidification cycle.

Effects of Cooling Rates. Slow cooling rates, on the order of 10^2 K/s, or long local solidification times result in *coarse* dendritic structures, with large spacing between dendrite arms. For higher cooling rates, on the order of 10^4 K/s, or short local solidification times, the structure becomes *finer*, with smaller dendrite arm spacing. For still higher cooling rates, on the order of from 10^6 to 10^8, the structures developed are *amorphous*, as described in Section 6.15.

The structures developed and the resulting grain sizes have an influence on the properties of the casting. As grain size decreases, the strength and ductility of the cast alloy increase, microporosity (*interdendritic shrinkage voids*) in the casting decreases, and the tendency for the casting to crack (*hot tearing*, see Fig. 10.14) during solidification decreases. Lack of uniformity in grain size and grain distribution produce castings that have *anisotropic properties*.

10.2.3 Structure–Property Relationships

Because all castings are expected to meet design and service requirements, the relationships between properties and the structures developed during solidification are important. This section describes these

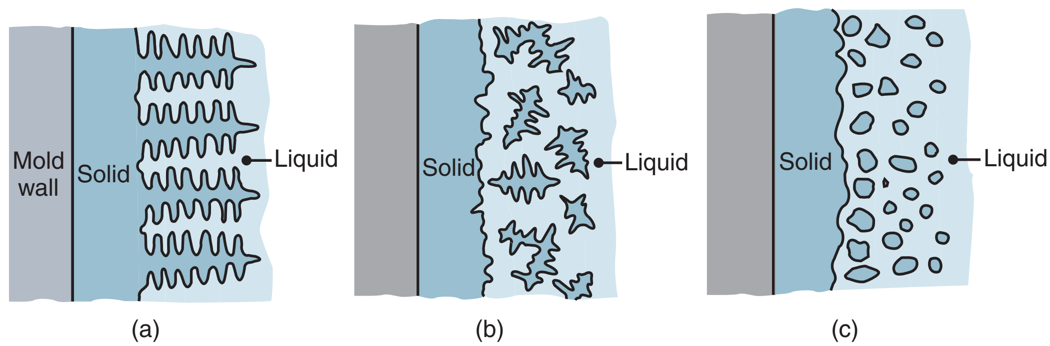


Figure 10.6: Schematic illustration of three basic types of cast structures: (a) columnar dendritic; (b) equiaxed dendritic; and (c) equiaxed nondendritic. *Source:* Courtesy of D. Apelian.

relationships in terms of dendrite morphology and the concentration of alloying elements in various regions within a casting.

The compositions of dendrites and the liquid metal are given by the *phase diagram* of the particular alloy. When the alloy is cooled very slowly, each dendrite develops a uniform composition; however, under the normally faster cooling rates encountered in practice, **cored dendrites** are formed. These dendrites have a surface composition different from that at their centers, a difference referred to as *concentration gradient*. The surface of the dendrite has a higher concentration of alloying elements than at its core, due to solute rejection from the core toward the surface during solidification of the dendrite (**microsegregation**). The darker shading in the interdendritic liquid near the dendrite roots, shown in Fig. 10.6, indicates that these regions have a higher solute concentration; microsegregation in these regions is much more pronounced than in others.

There are several types of **segregation**. In contrast to microsegregation, **macrosegregation** involves differences in composition throughout the casting itself. In situations where the solidification front moves away from the surface of a casting as a plane (Fig. 10.7), lower melting-point constituents in the solidifying alloy are driven toward the center (**normal segregation**). Consequently, such a casting has a higher concentration of alloying elements at its center than at its surfaces. In dendritic structures such as those found in solid–solution alloys (see Fig. 10.2b), the opposite occurs; that is, the center of the casting has a lower concentration of alloying elements (**inverse segregation**) than does at its surface. The reason is that liquid metal (having a higher concentration of alloying elements) enters the cavities developed from solidification shrinkage in the dendrite arms, which have solidified sooner. Another form of segregation is due to gravity; called **gravity segregation**, it involves a process whereby higher density inclusions, or compounds, sink while lighter elements (such as antimony in an antimony-lead alloy) float to the surface.

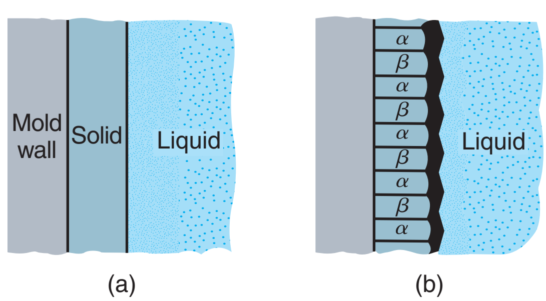


Figure 10.7: Schematic illustration of cast structures in (a) plane front, single phase, and (b) plane front, two phase. *Source:* Courtesy of D. Apelian.

A typical cast structure of a solid–solution alloy, with an inner zone of equiaxed grains, is shown in Fig. 10.2b. The inner zone can be extended throughout the casting, as shown in Fig. 10.2c, by adding an **inoculant** (*nucleating agent*) to the alloy. The inoculant induces nucleation of the grains throughout the liquid metal, called **heterogeneous nucleation**.

Because of the presence of *thermal gradients* in a solidifying mass of liquid metal, and due to gravity and the resulting density differences, *convection* has a strong influence on the structures developed. Convection involves heat transfer by the movement of matter; in a casting, it usually is associated with the flow of the liquid metal. Convection promotes the formation of an outer chill zone, refines grain size, and accelerates the transition from columnar to equiaxed grains. The structure shown in Fig. 10.6b also can be obtained by increasing convection within the liquid metal, whereby dendrite arms separate (**dendrite multiplication**). Conversely, reducing or eliminating convection results in coarser and longer columnar dendritic grains.

The dendrite arms are not particularly strong and can be broken up by agitation or by mechanical vibration in the early stages of solidification (as in **semisolid metal forming** and **rheocasting**, described in Section 11.4.7). This process results in finer grain size, with equiaxed nondendritic grains distributed more uniformly throughout the casting (Fig. 10.6c). A side benefit is the *thixotropic* behavior of alloys (that is, the viscosity decreases when the liquid metal is agitated), leading to improved castability of the metal. Another form of semisolid metal forming is **thixotropic casting**, where a solid billet is first heated to a semisolid state and then injected into a die-casting mold (Section 11.4.5).

10.2.4 Freeze Casting

Dendrite formation can be used to produce porous metals, ceramics, or polymers through the **freeze casting** process. In this approach, a slurry (suspension or mixture of particles in a liquid, usually water) is frozen. As the liquid solidifies, the particles are not soluble in the solid and therefore segregate at the solidification front. Eventually, the fluid freezes fully, with the carrier fluid and particles as separate phases.

The fluid can then be removed by lowering the pressure (in a vacuum), as shown in Fig. 10.8, and then raising the temperature. The result is a porous metal with a microstructure that is derived from the dendritic structure of the carrier fluid (Fig. 10.9). This structure generally has to be sintered to develop strength (Section 17.4).

10.3 Fluid Flow

To emphasize the importance of fluid flow in casting, consider a basic gravity casting system, as shown in Fig. 10.10. The molten metal is poured through a **pouring basin** or **cup**; it then flows through the **gating system** (consisting of sprue, runners, and gates) into the mold cavity. As also illustrated in Fig. 11.3, the **sprue** is a tapered vertical channel through which the molten metal flows downward in the mold. **Runners** are the channels that carry the molten metal from the sprue into the mold cavity or they connect the sprue to the **gate** (that portion of the runner through which the molten metal enters the mold cavity). **Risers**, also called **feeders**, serve as reservoirs of molten metal; they supply sufficient molten metal necessary to prevent porosity due to shrinkage during solidification.

Although such a gating system appears to be relatively simple, successful casting requires proper design and control of the solidification process to ensure adequate fluid flow in the system. For example, an important function of the gating system in sand casting is to trap contaminants (such as oxides and other inclusions) and remove them from the molten metal by having the contaminants adhere to the walls of the gating system, thereby preventing them from reaching the mold cavity. Furthermore, a properly designed gating system helps avoid or minimize such problems as premature cooling, turbulence, and gas entrapment. Even before it reaches the mold cavity, the molten metal must be handled carefully to avoid the formation of oxides on molten-metal surfaces from exposure to the environment or the introduction of impurities into the molten metal.

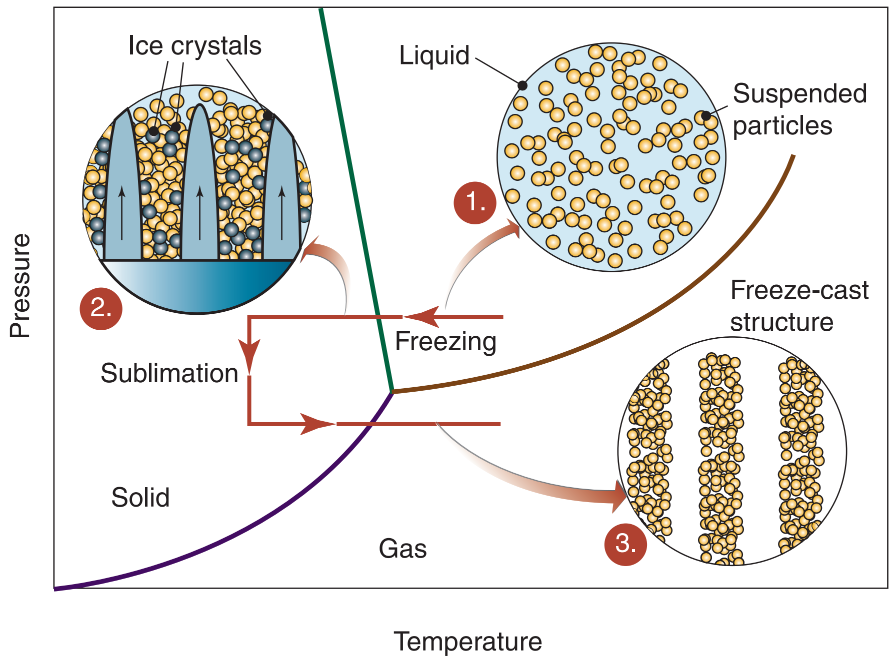


Figure 10.8: Freeze casting. (1) A slurry of a carrier liquid (commonly water) and insoluble particles is first reduced in temperature. (2) The liquid freezes in directional or dendritic fashion, forcing the particles away from the solidified volume where the particles are insoluble. (3) Decreasing pressure and increasing temperature evaporates the carrier fluid (*freeze drying*), leaving behind the particles that can be fused through sintering operations.

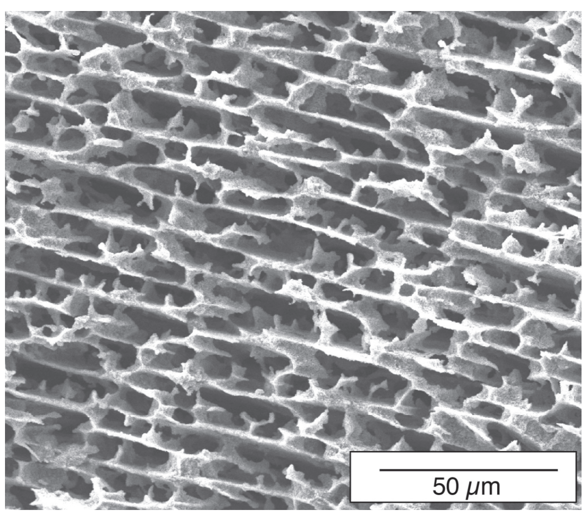


Figure 10.9: Microstructure after freeze casting. The specimen shown is titanium oxide (TiO_2) with pure water as a freezing agent. *Source:* Courtesy S. Naleway and T. Ogden, University of Utah.

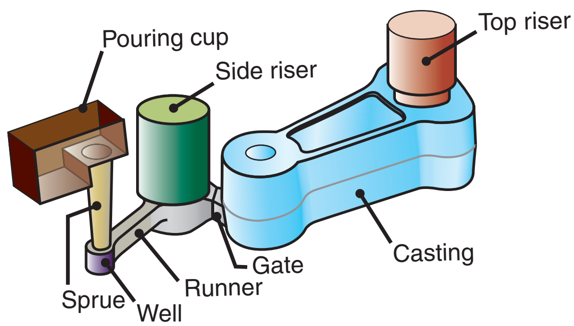


Figure 10.10: Schematic illustration of a typical riser-gated casting. Risers serve as reservoirs, supplying molten metal to the casting as it shrinks during solidification.

Two basic principles of fluid flow are relevant to gating design: Bernoulli's theorem and the law of mass continuity.

Bernoulli's Theorem. This theorem is based on the principle of the conservation of energy, and it relates pressure, velocity, the elevation of the fluid at any location in the system, and the frictional losses in a fluid system. The Bernoulli equation is

$$h + \frac{p}{\rho g} + \frac{v^2}{2g} = \text{constant}, \tag{10.2}$$

where h is the elevation above a certain reference level, p is the pressure at that elevation, v is the velocity of the liquid at that elevation, ρ is the density of the fluid (assuming that it is incompressible), and g is the gravitational constant. Conservation of energy requires that the following relationship be satisfied:

$$h_1 + \frac{p_1}{\rho g} + \frac{v_1^2}{2g} = h_2 + \frac{p_2}{\rho g} + \frac{v_2^2}{2g} + f, \tag{10.3}$$

where the subscripts 1 and 2 represent two different locations in the system and f represents the frictional loss in the liquid as it travels through the system. The frictional loss includes such factors as energy loss at the liquid-mold wall interfaces and turbulence in the liquid.

Mass Continuity. The law of mass continuity states that, for incompressible liquids and in a system with impermeable walls, the rate of flow is constant. Thus,

$$Q = A_1 v_1 = A_2 v_2, \tag{10.4}$$

where Q is the volume rate of flow (such as m^3/s), A is the cross-sectional area of the liquid stream, and v is the average velocity of the liquid in that cross section. The subscripts 1 and 2 refer to two different locations in the system. According to this law, the flow rate must be maintained everywhere in the system. The wall permeability is important, because otherwise some liquid will escape through the walls (as occurs in sand molds); thus, the flow rate will decrease as the liquid moves through the system. Coatings are often used to inhibit such behavior in sand molds. A small amount of permeability is, however, useful to allow escape of gases and can aid in heat transfer.

Sprue Design. An application of the two principles just described is the traditional *tapered* design of sprues (shown in Fig. 10.10). Note that in a free-falling liquid (such as water from a faucet), the cross-sectional area of the stream decreases as the liquid gains velocity. Thus, if a sprue has a constant cross-sectional area and molten metal is poured into it, regions may develop where the liquid loses contact with the sprue walls. As a result, **aspiration** (a process whereby air is entrapped in the liquid) may take place. One of two basic alternatives is used to prevent aspiration: a tapered sprue is used to prevent molten metal separation from the sprue wall, or straight-sided sprues are supplied with a **choking** mechanism at the bottom, consisting of either a choke core or a runner choke, as shown in Fig. 11.3. The choke slows the flow sufficiently to prevent aspiration in the sprue.

The specific shape of a tapered sprue that prevents aspiration can be determined from Eqs. (10.3) and (10.4). Assuming that the pressure at the top of the sprue is equal to the pressure at the bottom, and that there are no frictional losses, the relationship between height and cross-sectional area at any point in the sprue is given by the parabolic relationship

$$\frac{A_1}{A_2} = \sqrt{\frac{h_2}{h_1}}, \tag{10.5}$$

where, for example, the subscript 1 denotes the top of the sprue and 2 denotes the bottom. The distances h_1 and h_2 are measured from the liquid level in the pouring cup or basin (Fig. 10.10), so that h_2 is larger than h_1. Moving downward from the top, the cross-sectional area of the sprue must therefore decrease. The area at the bottom of the sprue, A_2, is selected to allow for desired flow rates, as described below, and the profile is produced according to Eq. (10.5).

Depending on the assumptions made, expressions other than Eq. (10.5) can also be obtained. For example, assume a certain molten-metal velocity, V_1, at the top of the sprue. Then, using Eqs. (10.3) and (10.4), an expression can be obtained for the ratio A_1/A_2 as a function of h_1, h_2, and V_1.

Modeling. Another application of the foregoing equations is in the *modeling of mold filling* in casting. For example, consider the situation shown in Fig. 10.10, where molten metal is poured into a pouring cup or basin; it flows through a sprue to a runner and a gate and fills the mold cavity. If the pouring basin has a much larger cross-sectional area than the sprue bottom, then the velocity of the molten metal at the top of the pouring basin is very low, and it can be taken to be zero. If frictional losses are due to a viscous dissipation of energy, then f in Eq. (10.3) can be taken to be a function of the vertical distance, and is often approximated as a linear function. Therefore, the velocity of the molten metal leaving the gate is obtained from Eq. (10.3) as

$$v = c\sqrt{2gh},$$

where h is the distance from the sprue base to the liquid metal height and c is a friction factor. For frictionless flow, c equals unity and for flows with friction, c is always between 0 and 1. The magnitude of c varies with mold material, runner layout, and channel size, and it can include energy losses due to turbulence, as well as to viscous effects.

If the liquid level has reached a height of x at the gate, then the gate velocity is

$$v = c\sqrt{2g}\sqrt{h - x}.$$

The flow rate through the gate will be the product of this velocity and the gate area according to Eq. (10.4). The shape of the casting will determine the height as a function of time. Integrating Eq. (10.4) gives the mean fill time and flow rate, and dividing the casting volume by the mean flow rate gives the mold fill time.

Simulation of mold filling assists designers in the specification of the runner diameter, as well as the size and number of sprues and pouring basins. To ensure that the runners remain open, the *fill time* must be a small fraction of the solidification time, but the velocity should not be so high as to erode the mold (referred to as mold wash) or to result in too high of a Reynolds number (see below). Otherwise, turbulence and associated air entrainment will result. Several computational tools are now available to evaluate gating designs and to assist in the sizing of components, such as Magmasoft, Flow 3D Cast, Wincast, ProCast, Quikcast, SolidCast, SUTCast, and PASSAGE/PowerCAST.

Flow Characteristics. An important consideration of fluid flow in gating systems is **turbulence**, as opposed to *laminar flow* of fluids. Turbulence is flow that is highly chaotic; such flow can lead to aspiration in casting systems. The *Reynolds number*, Re, is used to quantify this aspect of fluid flow. It represents the ratio of the *inertia* to the *viscous* forces in fluid flow and is defined as

$$\text{Re} = \frac{vD\rho}{\eta}, \tag{10.6}$$

where v is the velocity of the liquid, D is the diameter of the channel, and ρ and η are the density and viscosity of the liquid, respectively. The higher the Reynolds number, the greater the tendency for turbulent flow.

In gating systems, Re typically ranges from 2000 to 20,000, where a value of up to 2000 represents laminar flow; between 2000 and 20,000, it represents a mixture of laminar and turbulent flow. Such a mixture is generally regarded as harmless in gating systems. However, Re values in excess of 20,000 represent severe turbulence, resulting in significant air entrainment and the formation of *dross* (the scum that forms on the surface of molten metal) from the reaction of the liquid metal with air and other gases. Techniques for minimizing turbulence generally involve avoidance of sudden changes in flow direction and in the shape of channel cross sections in gating system design.

Dross or slag can be eliminated only by *vacuum casting* (see Section 11.4.2). Conventional atmospheric casting relieves the problem of dross or slag by (a) skimming, (b) using properly designed pouring basins and runner systems, (c) tapping the molten metal from below the surface, such as in pressure casting (Fig. 11.18), or (d) using filters, which also can eliminate turbulent flow in the runner system. Filters are typically made of ceramics, mica, or fiberglass; their proper location and placement are important for effective filtering of dross and slag.

10.4 Fluidity of Molten Metal

The capability of molten metal to fill mold cavities is called *fluidity*; it consists of two basic factors: (a) the molten metal and (b) casting parameters. The characteristics of the molten metal that influence fluidity are:

Viscosity. As viscosity and its sensitivity to temperature increase, fluidity decreases.

Surface Tension. A high surface tension of the liquid metal reduces fluidity; also, oxide films on the surface of the molten metal have a significantly adverse effect on fluidity. For example, an oxide film on the surface of pure molten aluminum triples the surface tension.

Inclusions. Because they are insoluble, inclusions can have a significant effect on fluidity. This effect can easily be verified by observing the viscosity of a liquid (such as oil) with and without sand particles in it; a liquid with sand in it has a higher viscosity and thus lower fluidity.

Solidification Pattern of the Alloy. The manner in which solidification takes place (Section 10.2) can influence fluidity. Fluidity is inversely proportional to the freezing range (see Eq. 10.1): The shorter the range, as in pure metals and eutectics, the higher the fluidity. Conversely, alloys with long freezing ranges, such as solid–solution alloys, have lower fluidity.

The following casting parameters influence fluidity, fluid flow, and thermal characteristics of the system:

Mold Design. The design and dimensions of the sprue, runners, and risers all influence fluidity.

Mold Material and Its Surface Characteristics. The higher the thermal conductivity of the mold and the rougher its surfaces, the lower is the fluidity of the molten metal. Although heating the mold improves fluidity, it slows down solidification of the metal; thus, the casting develops coarse grains and hence has lower strength.

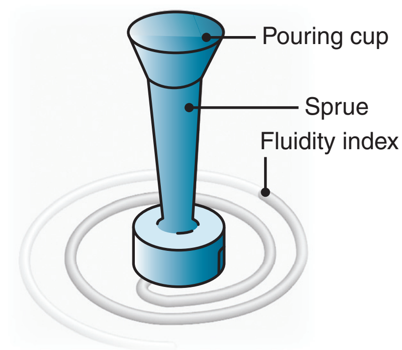


Figure 10.11: A test method for fluidity using a spiral mold. The fluidity index is the length of the solidified metal in the spiral passage. The greater the length of the solidified metal, the greater is the metal's fluidity.

Degree of Superheat. *Superheat* (defined as the increment of temperature of an alloy above its melting point) improves fluidity by delaying solidification. The **pouring temperature** often is specified instead of the degree of superheat, because it is more easily measured and controlled.

Rate of Pouring. The slower the rate of pouring molten metal into the mold, the lower the fluidity because of the higher rate of cooling when poured slowly.

Heat Transfer. This factor directly affects the viscosity of the liquid metal (see below).

Castability. Although complex, this term is generally used to describe the ease with which a metal can be cast to produce a part with good quality. It includes not only fluidity, but is also affected by casting practices.

10.4.1 Tests for Fluidity

Several tests have been developed to quantify fluidity, although none has been accepted universally. In one common test, the molten metal is made to flow along a channel that is at room temperature (Fig. 10.11); the distance the metal flows before it solidifies and stops flowing is a measure of its fluidity. Obviously, the length is a function of the thermal properties of the metal and the mold, as well as of the design of the channel.

10.5 Heat Transfer

The heat transfer during the complete cycle (from pouring, to solidification, and to cooling to room temperature) is an important consideration in metal casting. Heat flow at different locations in the system is a complex phenomenon and depends on several factors related to the material cast and the physical properties of the mold and the processing parameters. For instance, in casting thin sections, the metal flow rates must be high enough to avoid premature chilling and solidification of the metal. On the other hand, the flow rate must not be so high as to cause excessive turbulence, with its detrimental effects on the casting operation.

A typical temperature distribution at the mold-liquid metal interface is shown in Fig. 10.12. Heat from the liquid metal is given off through the mold wall and to the surrounding air. The temperature drop at the air–mold and mold–metal interfaces is caused by the presence of boundary layers and imperfect contact at these interfaces. The shape of the curve depends on the thermal properties of the molten metal and the mold.

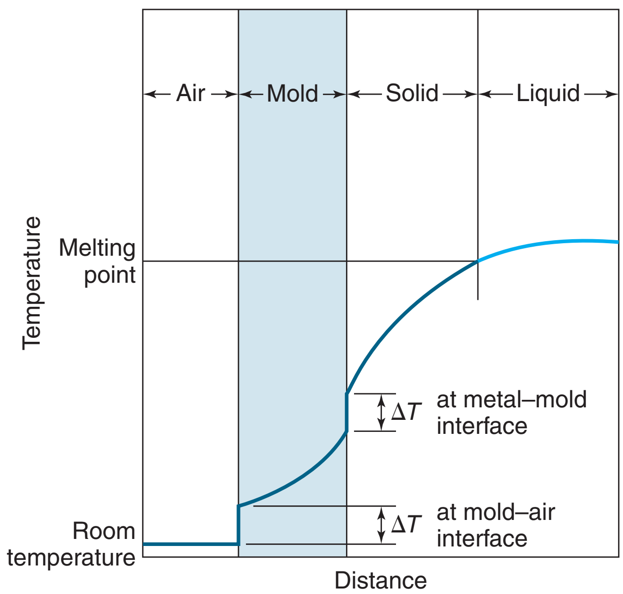


Figure 10.12: Temperature distribution at the interface of the mold wall and the liquid metal during the solidification of metals in casting.

10.5.1 Solidification Time

During the early stages of solidification of the metal, a thin skin begins to form at the relatively cool mold walls; as time passes, the thickness of the skin increases (Fig. 10.13). With flat mold walls, the thickness is proportional to the square root of time; thus, doubling the time will make the skin $\sqrt{2} = 1.41$ times or 41% thicker.

The **solidification time** is a function of the volume of a casting and its surface area (*Chvorinov's rule*):

$$\text{Solidification time} = C\left(\frac{\text{Volume}}{\text{Surface area}}\right)^n, \tag{10.7}$$

where C is a constant reflecting (a) the mold material, (b) the metal properties, including latent heat, and (c) the temperature. The parameter n has a value between 1.5 and 2, but usually taken as 2. Thus, for example, a large solid sphere will solidify and cool to ambient temperature at a much slower rate than will a smaller solid sphere. Note that the volume of a sphere is proportional to the cube of its diameter, and its surface area

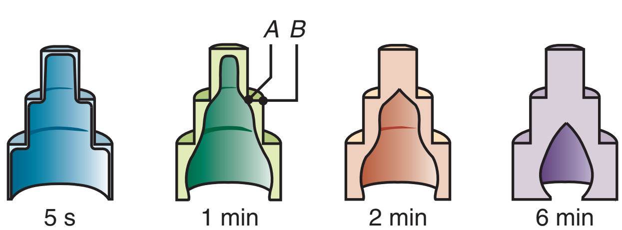


Figure 10.13: Solidified skin on a steel casting. The remaining molten metal is poured out at the times indicated in the figure. Hollow ornamental and decorative objects are made by a process called *slush casting*, which is based on this principle. *Source:* After H.F. Taylor, J. Wulff, and M.C. Flemings.

is proportional to the square of its diameter. Similarly, it can be shown that molten metal in a cube-shaped mold will solidify faster than in a spherical mold of the same volume (see Example 10.1).

The effects of mold geometry and elapsed time on skin thickness and shape are shown in Fig. 10.13. As illustrated, the unsolidified molten metal has been poured from the mold at different time intervals ranging from 5 seconds to 6 minutes. As expected, the skin thickness increases with elapsed time, and the skin is thinner at internal angles (location A in the figure) than at external angles (location B). The latter condition is caused by slower cooling at internal angles than at external angles.

Example 10.1 Solidification Times for Various Shapes

Given: Three metal pieces being cast have the same volume, but different shapes: One is a sphere, one a cube, and the other a cylinder with its height equal to its diameter. Assume that $n = 2$.

Find: Which piece will solidify the fastest, and which one the slowest?

Solution: The volume of the piece is taken as unity; thus from Eq. (10.7),

$$\text{Solidification time} \propto \frac{1}{(\text{Surface area})^2}.$$

The respective surface areas are as follows:

Sphere:

$$V = \left(\frac{4}{3}\right)\pi r^3, \qquad r = \left(\frac{3}{4\pi}\right)^{1/3},$$

$$A = 4\pi r^2 = 4\pi\left(\frac{3}{4\pi}\right)^{2/3} = 4.84.$$

Cube:

$$V = a^3, \qquad a = 1, \qquad \text{and} \qquad A = 6a^2 = 6$$

Cylinder:

$$V = \pi r^2 h = 2\pi r^3, \qquad r = \left(\frac{1}{2\pi}\right)^{1/3},$$

$$A = 2\pi r^2 + 2\pi r h = 6\pi r^2 = 6\pi\left(\frac{1}{2\pi}\right)^{2/3} = 5.54.$$

The respective solidification times are therefore

$$t_{\text{sphere}} = 0.043C, \qquad t_{\text{cube}} = 0.028C, \qquad t_{\text{cylinder}} = 0.033C.$$

Hence, the cube-shaped piece will solidify the fastest, and the spherical piece will solidify the slowest.

10.5.2 Shrinkage

Because of their thermal expansion characteristics, metals usually shrink (contract) during solidification and while cooling to room temperature. *Shrinkage*, which causes dimensional changes and sometimes warping and cracking, is the result of the following three sequential events:

1. Contraction of the molten metal as it cools prior to its solidification
2. Contraction of the metal during phase change from liquid to solid
3. Contraction of the solidified metal (the casting) as its temperature drops to ambient temperature.

Table 10.1: Volumetric Solidification Contraction or Expansion for Various Cast Metals.

Contraction (%)		Expansion (%)	
Aluminum	7.1	Bismuth	3.3
Zinc	6.5	Silicon	2.9
Al–4.5% Cu	6.3	Gray iron	2.5
Gold	5.5		
White iron	4–5.5		
Copper	4.9		
Brass (70–30)	4.5		
Magnesium	4.2		
90% Cu–10% Al	4		
Carbon steels	2.5–4		
Al–12% Si	3.8		
Lead	3.2		

The largest amount of shrinkage occurs during the phase change of the material from liquid to solid; detrimental effects of this shrinkage can be reduced through the use of risers or by pressure-feeding of molten metal. The amount of contraction during solidification of various metals is shown in Table 10.1. Note that some metals, such as gray cast iron, expand. The reason for this expansion is that graphite has a relatively high specific volume, and when it precipitates as graphite flakes during solidification of the gray cast iron, it causes a net expansion of the metal. Shrinkage, especially that due to thermal contraction, is further described in Section 12.2.1 in connection with design considerations in casting.

10.6 Defects

Depending on factors such as the quality of raw materials, product design, and control of processing parameters, several defects can develop in castings, as illustrated in Figs. 10.14 and 10.15. While some defects affect only the appearance of the parts made, others can have major adverse effects on their structural

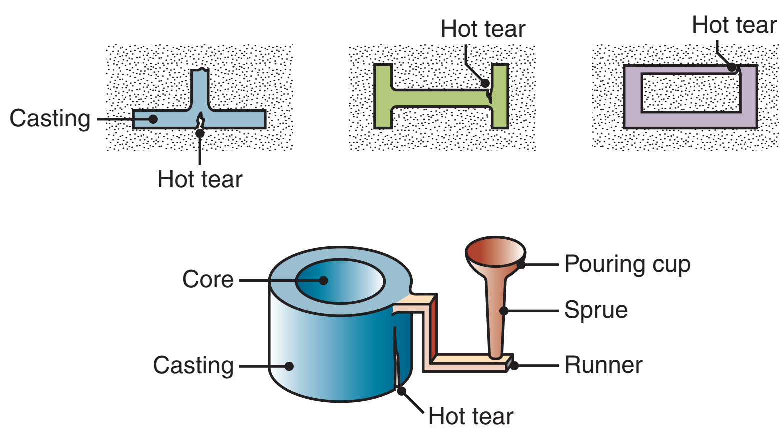


Figure 10.14: Examples of hot tears in castings. These defects occur because the casting cannot shrink freely during cooling, owing to constraints in various portions of the molds and cores. Exothermic (heat producing) compounds may be used as *exothermic padding* to control cooling at critical regions to avoid hot tearing.

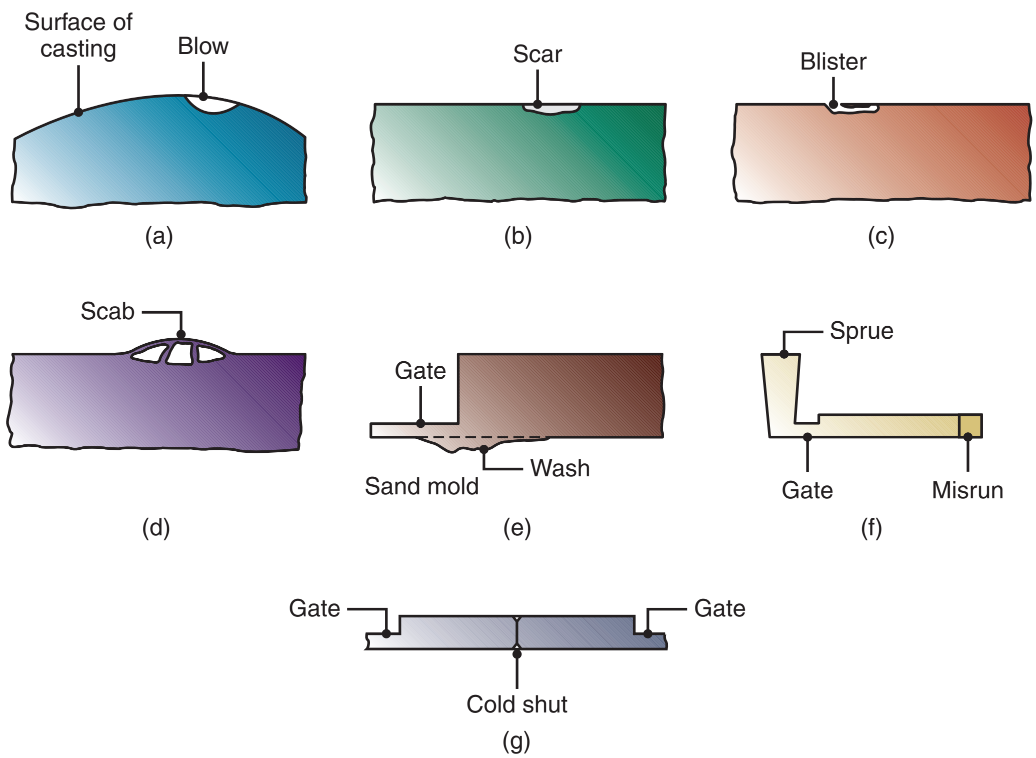


Figure 10.15: Examples of common defects in castings. These defects can be minimized or eliminated by proper design and preparation of molds and control of pouring procedures. *Source:* After J. Datsko.

integrity. The International Committee of Foundry Technical Associations has developed a standardized nomenclature, consisting of seven basic categories of casting defects, identified with boldface capital letters:

A—Metallic projections, consisting of fins, flash, or projections, such as swells and rough surfaces.

B—Cavities, consisting of rounded or rough internal or exposed cavities, including blowholes, pinholes, and shrinkage cavities (see *porosity*, Section 10.6.1).

C—Discontinuities, such as cracks, cold or hot tearing, and cold shuts. If the solidifying metal is constrained from shrinking freely, cracking and tearing may occur. Although several factors are involved in tearing, coarse grain size and the presence of low-melting-point segregates along the grain boundaries of the metal increase the tendency for hot tearing. *Cold shut* is an interface in a casting that lacks complete fusion, because of the meeting of two streams of liquid metal from different gates.

D—Defective surface, such as surface folds, laps, scars, adhering sand layers, and oxide scale.

E—Incomplete casting, such as *misruns* (due to premature solidification), insufficient volume of the metal poured, and *runout* (due to loss of metal from the mold after pouring). Incomplete castings also can result from the molten metal being at too low a temperature or from pouring the metal too slowly.

F—Incorrect dimensions or shape, due to such factors as improper shrinkage allowance, pattern-mounting error, irregular contraction, deformed pattern, or warped casting.

G—Inclusions, which form during melting, solidification, and molding. Generally nonmetallic, they are regarded as harmful because they act as stress raisers, reducing the strength of the casting. Inclusions may form during melting when the molten metal reacts with the environment (usually oxygen), with the crucible or the mold material. Chemical reactions among components in the molten metal itself may produce inclusions. Slags and other foreign material entrapped in the molten metal can become inclusions, although filtering the molten metal can remove particles as small as 30 μm. Spalling of the mold and core surfaces also can produce inclusions, thus indicating the importance of the quality of molds and of their proper maintenance.

10.6.1 Porosity

Porosity in a casting may be caused by *shrinkage*, entrained or dissolved *gases*, or both. Porous regions can develop in castings because of **shrinkage** of the solidified metal. Thin sections in a casting solidify sooner than thicker regions; consequently, molten metal flows into the thicker regions that have not yet solidified. Porous regions may develop at their centers because of contraction as the surfaces of the thicker regions begin to solidify first. *Microporosity* can develop when the liquid metal solidifies and shrinks between dendrites and between dendrite branches. Porosity is detrimental to the strength and ductility of a casting and its surface finish, potentially making the casting permeable, thus affecting the pressure tightness of a cast pressure vessel.

Porosity caused by shrinkage can be reduced or eliminated by various means:

- Adequate liquid metal should be provided to prevent cavities caused by shrinkage.
- Internal or external **chills**, as those used in sand casting (Fig. 10.16), are an effective means of reducing shrinkage porosity. The function of chills is to increase the rate of solidification in critical regions. Internal chills are usually made from the same material as the casting itself, and are left within the casting. Problems may arise that involve proper fusion of the internal chills with the casting; thus,

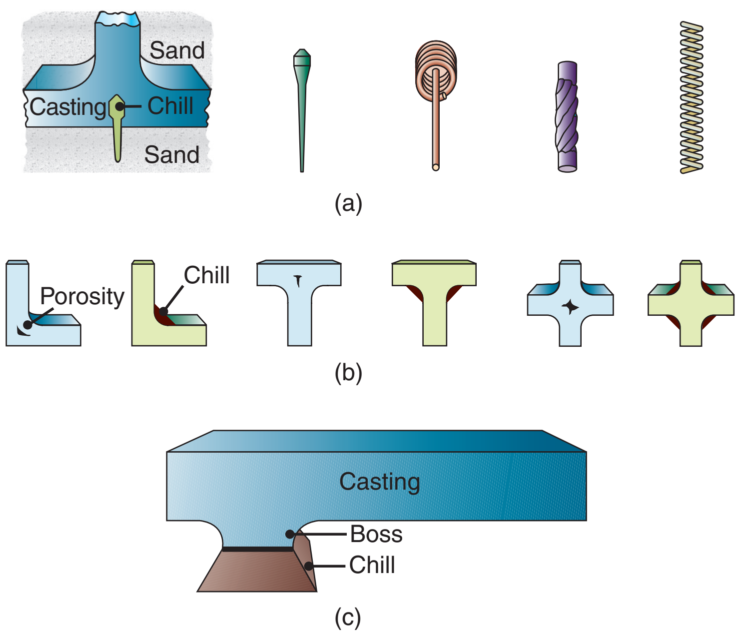


Figure 10.16: Various types of (a) internal and (b) external chills (dark areas at corners) used in castings to eliminate porosity caused by shrinkage. Chills are placed in regions where there is a larger volume of metal, as shown in (c).

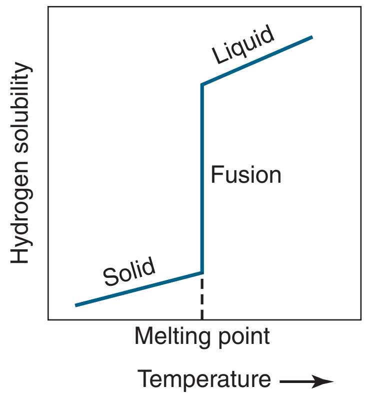


Figure 10.17: Solubility of hydrogen in aluminum. Note the sharp decrease in solubility as the molten metal begins to solidify.

foundries generally avoid using internal chills. External chills may be made from the same material as the casting or may be made of iron, copper, or graphite.

- Porosity in alloys can be reduced or eliminated by using high temperature gradients, that is, by increasing the cooling rate. For example, mold materials with higher thermal conductivity may be used.

- Subjecting the casting to hot isostatic pressing is another method of reducing porosity (see Section 17.3.2).

Gases are more soluble in liquid metals than in solid metals (Fig. 10.17); thus, when a metal begins to solidify, the dissolved gases are expelled. Gases also may be due to reactions of the molten metal with the mold materials. Gases either accumulate in regions of existing porosity (such as in interdendritic regions; see Section 10.2.3) or cause microporosity in the casting, particularly in cast iron, aluminum, and copper. Dissolved gases may be removed from the molten metal by *flushing* or *purging* with an inert gas or by melting and pouring the metal in a vacuum. If the dissolved gas is oxygen, the molten metal can be *deoxidized*.

Whether microporosity is a result of shrinkage or is caused by gases may be difficult to determine. If the porosity is spherical and has smooth walls (similar to the shiny holes in Swiss cheese), it is generally from gases. On the other hand, if the walls are rough and angular, porosity is likely from shrinkage between dendrites. Gross porosity is from shrinkage and usually is called a **shrinkage cavity**.

Summary

- Casting is a solidification process in which molten metal is poured into a mold and allowed to cool. The metal may flow through a variety of passages (pouring basins, sprues, runners, risers, and gating systems) before reaching the final mold cavity. Bernoulli's theorem, the continuity law, and the Reynolds number are the analytical tools used in designing castings, with the goals of achieving an appropriate flow rate and eliminating defects associated with fluid flow.

- Solidification of pure metals takes place at a constant temperature, whereas solidification of alloys occurs over a range of temperatures. Phase diagrams are important tools for identifying the solidification point or points for technologically important metals.

- The composition and cooling rates of the molten metal both affect the size and shape of the grains and the dendrites in the solidifying alloy. In turn, the size and structure of grains and dendrites influence properties of the solidified casting. Solidification time is a function of the volume and surface area of a casting (Chvorinov's rule).
- The grain structure of castings can be controlled by various means to obtain desired properties. Because most metals contract during solidification and cooling, cavities can form in the casting. Porosity caused by gases evolved during solidification can be a significant problem, particularly because of its adverse effect on the mechanical properties of castings. Various defects also can develop in castings from lack of control of material and process variables.
- Dimensional changes and cracking (hot tearing) are difficulties that can arise during solidification and cooling. Several basic categories of casting defects have been identified.
- Melting practices have a direct effect on the quality of castings, as do foundry operations such as pattern and mold making, pouring the molten metal, removing the cast parts from molds, cleaning, heat treatment, and inspection.

Key Terms

Aspiration
Bernoulli's theorem
Casting
Chills
Columnar dendrite
Columnar grain
Cored dendrite
Dendrite
Fluidity
Freeze casting
Freezing range
Gate
Gating system
Heterogeneous nucleation
Homogenous nucleation
Inoculant
Macrosegregation
Microsegregation
Mold
Mushy zone
Normal segregation
Porosity
Pouring basin
Reynolds number
Rheocasting
Riser
Runner
Segregation
Shrinkage
Skin
Solidification
Sprue
Turbulence

Bibliography

Analysis of Casting Defects, American Foundrymen's Society, 2002.
ASM Handbook, Vol. 1A: **Cast Iron Science and Technology**, ASM International, 2017.
ASM Handbook, Vol. 15: **Casting**, ASM International, 2008.
Campbell, J., **Castings**, Butterworth-Heinemann, 2nd ed., 2003.
Campbell, J., **Complete Casting Handbook: Metal Casting Processes, Techniques and Design**, Butterworth-Heinemann, 2011.

Cantor, B., and O'Reilly, K., **Solidification and Casting**, Taylor & Francis, 2002.
Casting, in *Tool and Manufacturing Engineers Handbook*, Volume II: *Forming*, Society of Manufacturing Engineers, 1984.
Glicksmann, M.E., **Principles of Solidification: An Introduction to Modern Casting and Crystal Growth Concepts**, Springer, 2010.
Poirer, D.R., and Poirer, E.J., **Heat Transfer Fundamentals for Metal Casting**, Wiley, 1998.
Reikher, A., and Barkhudarov, M., **Casting: An Analytical Approach**, Springer, 2008.
Sahoo, M., and Sahu, S., **Principles of Metal Casting**, 3rd ed., McGraw-Hill, 2014.
Steel Castings Handbook, 6th ed., Steel Founders' Society of America, 1995.
Stefanescu, D.M., **Science and Engineering of Casting Solidification**, 3rd ed., Springer, 2015.

Review Questions

10.1. Explain why casting is an important manufacturing process.

10.2. Why do most metals shrink when they are cast?

10.3. What are the differences between the solidification of pure metals and metal alloys?

10.4. What are dendrites? Why are they called so?

10.5. Describe the difference between short and long freezing ranges.

10.6. What is superheat? Is it important? What are the consequences of excessive superheat?

10.7. Define shrinkage and porosity. How can you tell whether cavities in a casting are due to porosity or to shrinkage?

10.8. What is the function of chills? What are they made of?

10.9. Why is the Reynolds number important in casting?

10.10. What is a sprue? What shape should a sprue have if a mold has no other choking means?

10.11. How is fluidity defined? Why is it important?

10.12. Explain the reasons for hot tearing in castings.

10.13. Why is it important to remove dross or slag during the pouring of molten metal into the mold? What methods are used to remove them?

10.14. Why is Bernoulli's equation important in casting?

10.15. Describe thixocasting and rheocasting.

10.16. What is Chvorinov's Rule?

10.17. How is a blister related to a scab?

Qualitative Problems

10.18. Is there porosity in a chocolate bar? In an ice cube? Explain.

10.19. Describe the stages involved in the contraction of metals during casting.

10.20. Explain the effects of mold materials on fluid flow and heat transfer in casting operations.

10.21. It is known that pouring metal at a high rate into a mold can have certain disadvantages. Are there any disadvantages to pouring it very slowly?

10.22. Describe the events depicted in Fig. 10.5.

10.23. Would you be concerned about the fact that portions of internal chills are left within the casting? Explain.

10.24. Review Fig. 10.10 and make a summary, explaining the purpose of each feature in shown and the consequences of omitting the feature from the mold design.

10.25. Make a sketch of volume vs. temperature for a metal that shrinks when it cools from the liquid state to room temperature. On the graph, mark the area where shrinkage is compensated by risers.

10.26. What practical demonstrations can you suggest to indicate the relationship of the solidification time to the volume and surface area of a casting?

10.27. Explain why a casting may have to be subjected to various heat treatments.

10.28. List and explain the reasons that porosity can develop in a casting.

10.29. Why does porosity have detrimental effects on the mechanical properties of castings? Would physical properties, such as thermal and electrical conductivity, also be adversely affected by porosity? Explain.

10.30. A spoked handwheel is to be cast in gray iron. In order to prevent hot tearing of the spokes, would you insulate the spokes or chill them? Explain.

10.31. Which of the following considerations are important for a riser to function properly? Must it: (a) have a surface area larger than the part being cast, (b) be kept open to atmospheric pressure, and/or (c) solidify first? Explain.

10.32. Explain why the constant C in Eq. (10.7) depends on mold material, metal properties, and temperature.

10.33. Are external chills as effective as internal chills? Explain.

10.34. Explain why, as shown in Table 10.1, gray cast iron undergoes expansion rather than contraction during solidification.

10.35. Referring to Fig. 10.13, explain why internal corners, such as A, develop a thinner skin than external corners, such as B, during solidification.

10.36. Note the shape of the two risers shown in Fig. 10.10, and discuss your observations with respect to Eq. (10.7).

10.37. Is there any difference in the tendency for shrinkage void formation in metals with short and long freezing ranges, respectively? Explain.

10.38. What is the influence of the cross-sectional area of the spiral channel shown in Fig. 10.11 on fluidity test results? What is the effect of sprue height? If this test is run with the entire test setup heated to elevated temperatures, would the results be more useful? Explain.

10.39. It has long been observed that (a) low pouring temperatures (*i.e.*, low superheat) promote the formation of equiaxed grains over columnar grains and (b) equiaxed grains become finer as the pouring temperature decreases. Explain these two phenomena.

10.40. In casting metal alloys, what would you expect to occur if the mold were agitated (vibrated) aggressively after the molten metal had been in the mold for a sufficient amount of time to form a skin?

10.41. If you inspect a typical cube of ice, you are likely to see air pockets and cracks in the cube. Some ice cubes, however, are tubular in shape and do not have noticeable air pockets or cracks in their structure. Explain this phenomenon.

10.42. How can you tell whether cavities in a casting are due to shrinkage or entrained air bubbles?

10.43. Describe the drawbacks to having a riser that is (a) too large and (b) too small.

10.44. Reproduce Fig. 10.2 for a casting that is spherical in shape.

10.45. List the process variables that affect the fluidity index as shown in Fig. 10.11.

10.46. Assume that you have a method of measuring porosity in a casting. Could you use this information to accurately predict the strength of the casting? Explain.

Quantitative Problems

10.47. Derive Eq. (10.5).

10.48. Sketch a plot of specific volume versus temperature for a metal that shrinks as it cools from the liquid state to room temperature. On the graph, mark the area where shrinkage is compensated for by risers.

10.49. A round casting is 0.3 m in diameter and 1 m in length. Another casting of the same metal is elliptical in cross section, with a major-to-minor axis ratio of 2, and has the same length and cross sectional area as the round casting. Both pieces are cast under the same conditions. What is the difference in the solidification times of the two castings?

10.50. A 75-mm thick square plate and a right circular cylinder with a radius of 100 mm and a height of 25 mm have the same volume. If each is to be cast with the use of a cylindrical riser, will each part require the same-size riser to ensure proper feeding? Explain.

10.51. Assume that the top of a round sprue has a diameter of 75 mm and is at a height of 150 mm from the runner. Based on Eq. (10.5), plot the profile of the sprue diameter as a function of its height. Assume that the sprue has a diameter of 5 mm at the bottom.

10.52. Pure aluminum is poured into a sand mold. The metal level in the pouring basin is 250 mm above the metal level in the mold, and the runner is circular with a 7.5 mm diameter. What is the velocity and rate of the flow of the metal into the mold? Is the flow turbulent or laminar?

10.53. A cylinder with a diameter of 50 mm and height of 75 mm solidifies in 3 minutes in a sand casting operation. What is the solidification time if the cylinder height is doubled? What is the time if the diameter is doubled?

10.54. The volume flow rate of metal into a mold is 0.05 m^3/s. The top of the sprue has a diameter of 30 mm, and its length is 200 mm. What diameter should be specified at the bottom of the sprue to prevent aspiration? What is the resultant velocity and Reynolds number at the bottom of the sprue if the metal being cast is aluminum with a viscosity of 0.004 Ns/m^2?

10.55. A rectangular mold with dimensions 120 mm $\times$ 240 mm $\times$ 480 mm is filled with aluminum with no superheat. Determine the final dimensions of the part as it cools to room temperature. Repeat the analysis for gray cast iron.

10.56. The constant C in Chvorinov's rule is given as 2.5 s/mm^2 and is used to produce a cylindrical casting with a diameter of 60 mm and height of 130 mm. Estimate the time for the casting to fully solidify. The mold can be broken safely when the solidified shell is at least 20 mm. Assuming that the cylinder cools evenly, how much time must transpire after pouring the molten metal before the mold can be broken?

10.57. A sprue is 300 mm long and has a diameter of 75 mm at the top. The molten metal level in the pouring basing (which is much larger than the top of the sprue) is taken to be 75 mm from the top of the sprue for design purposes. If a flow rate of 500 cm^3/s is to be achieved, what should be the diameter at the bottom of the sprue? Will the sprue aspirate? Explain.

10.58. Pure copper is poured into a sand mold. The metal level in the pouring basin is 250 mm above the metal level in the mold, and the runner is circular with a 10 mm diameter. What are the velocity and rate of the flow of the metal into the mold? Is the flow turbulent or laminar?

10.59. For the sprue described in Problem 10.58, what runner diameter is needed to ensure a Reynolds number of 2000? How long will a 250 cm^3 casting take to fill with such a runner?

10.60. How long would it take for the sprue in Problem 10.58 to feed a casting with a square cross section of 50 mm per side and a height of 100 mm? Assume that the sprue is frictionless.

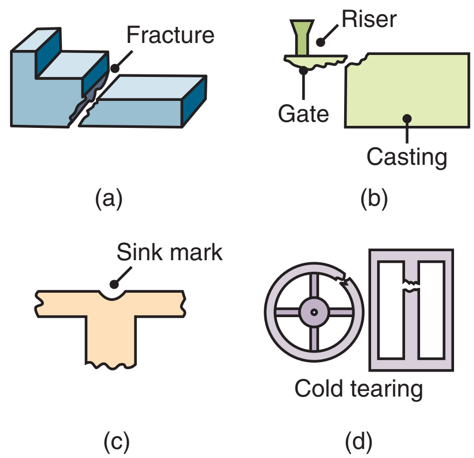
10.61. Assume that you are an instructor covering the topics described in this chapter and you are giving a quiz on the numerical aspects to test the understanding of the students. Prepare two quantitative problems and supply the answers to them.

10.62. When designing patterns for casting, pattern makers use special rulers that automatically incorporate solid shrinkage allowances into their designs. Therefore, a 100-mm patternmaker's ruler is longer than 100 mm. How long should a patternmaker's ruler be for (1) aluminum castings, (2) malleable cast iron, and (3) high manganese steel?

Synthesis, Design, and Projects

10.63. Can you devise fluidity tests other than that shown in Fig. 10.11? Explain the features of your test methods.

10.64. The illustration indicates various defects and discontinuities in cast products. Review each defect and offer solutions to avoid it.



10.65. The fluidity test shown in Fig. 10.11 illustrates only the principle of this test. Design a setup for such a test, showing the type of materials and the equipment to be used. Explain the method by which you would determine the length of the solidified metal in the spiral passage.

10.66. Utilizing the equipment and materials available in a typical kitchen, design an experiment to reproduce results similar to those shown in Fig. 10.13. Comment on your observations.

10.67. One method of relieving stress concentrations in a part is to apply a small, uniform plastic deformation to it. Make a list of your concerns and recommendations if such an approach is suggested for a casting.

10.68. Describe the effects on mold design, including the required change in the size of the risers, runners, chokes, and sprues, for a casting of a given shape that is to be doubled in volume.

10.69. Small amounts of slag often persist after skimming and are introduced into the molten-metal flow in casting. Recognizing that the slag is much less dense than the metal, design mold features that will remove small amounts of slag before the metal reaches the mold cavity.

10.70. Figure II.2 shows a variety of components in a typical automobile that are produced by casting. Think of other products, such as power tools and small appliances, and prepare an illustration similar to the figure.

10.71. Design an experiment to measure the constants C and n in Chvorinov's rule, Eq. (10.7). Describe the features of your design, and comment on any difficulties that might be encountered in running such an experiment.

Chapter 11

Metal-casting Processes and Equipment

Case Studies:

- Building upon the fundamentals of solidification, fluid flow, and heat transfer described in the preceding chapter, this chapter presents the principles of industrial casting processes.
- Casting processes are generally categorized as permanent-mold and expendable-mold processes; expendable-mold processes are further categorized as permanent-mold and expendable-pattern processes.
- The characteristics of each process are described, together with typical applications, advantages, and limitation.
- Special casting processes that produce single-crystal components as well as amorphous alloys are then described.
- The chapter ends with a description of inspection techniques for castings.

Typical products made by casting: Engine blocks, crankshafts, power tool housings, turbine blades, plumbing parts, zipper teeth, dies and molds, gears, railroad wheels, propellers, and office equipment.

Alternative processes: Forging, powder metallurgy, additive manufacturing, machining, and fabrication.

11.1 Introduction

Metal castings were first made during the period 4000 to 3000 B.C., using stone and metal molds for casting copper. A variety of casting processes have been developed over time, each with its own characteristics and applications to meet specific design requirements (Table 11.1, see also Fig. I.6a). A very wide variety of parts and components are made by casting, such as frying pans, jewelry, engine blocks, crankshafts, automotive components and powertrains (Fig. 11.1), agricultural and railroad equipment, pipes, plumbing fixtures, power-tool housings, gun barrels, orthopedic implants, and very large components for hydraulic turbines.

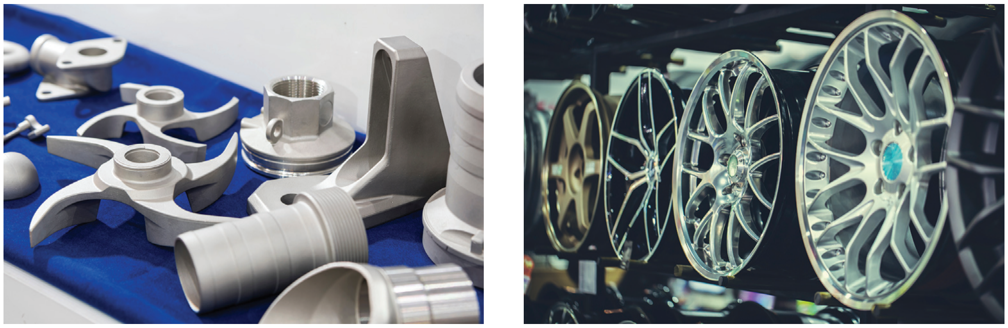
Four trends have had a major impact on the casting industry. (a) Mechanization and automation of the casting process, which has led to significant changes in the use of equipment and labor; advanced machinery and automated process-control systems have replaced or enhanced traditional methods of casting. (b) Increasing demand for high-quality castings with close dimensional tolerances. (c) Development of powerful modeling software, allowing predictive evaluation of dies and potential defects that result from poor design, as well as estimating a cast material's mechanical properties and its microstructure. This **Integrated Computer Materials Engineering** (ICME) trend has positively affected most manufacturing processes. (d) Additive manufacturing (Chapter 20) is still evolving, but has greatly aided mold manufacturing.

This chapter is organized around the major classifications of casting practices (see Fig. II.3 outlined in the Introduction to Part II), as they relate to mold materials, pattern production, molding processes, and methods of feeding the mold with molten metal. The major categories are:

1. **Expendable molds**, typically made of sand, plaster, ceramics, and similar materials, and generally mixed with various binders (*bonding agents*) for improved properties. A typical sand mold consists

Table 11.1: Summary of Casting Processes.

Process	Advantages	Limitations
Sand	Almost any metal can be cast; no limit to part size, shape, or weight; low tooling cost	Some finishing required; relatively coarse surface finish; wide tolerances
Shell mold	Good dimensional accuracy and surface finish; high production rate	Part size limited; expensive patterns and equipment
Evaporative pattern	Most metals can be cast, with no limit to size; complex part shapes	Patterns have low strength and can be costly for low quantities
Plaster mold	Intricate part shapes; good dimensional accuracy and surface finish; low porosity	Limited to nonferrous metals; limited part size and volume of production; mold-making time relatively long
Ceramic mold	Intricate part shapes; close-tolerance parts; good surface finish; low cooling rate	Limited part size
Investment	Intricate part shapes; excellent surface finish and accuracy; almost any metal can be cast	Part size limited; expensive patterns, molds, and labor
Permanent mold	Good surface finish and dimensional accuracy; low porosity; high production rate	High mold cost; limited part shape and complexity; not suitable for high-melting-point metals
Die	Excellent dimensional accuracy and surface finish; high production rate	High die cost; limited part size; generally limited to nonferrous metals; long lead time
Centrifugal	Large cylindrical or tubular parts with good quality; high production rate	Expensive equipment; limited part shape



(a) (b)

Figure 11.1: (a) Examples of stainless steel castings. Note the intricate part shapes. (b) Die-cast magnesium automobile wheels. *Source:* (a) Shutterstock/Mr.1 (b) Shutterstock/socrates471.

of 90% sand, 7% clay, and 3% water. As described in Section 8.2, these materials are *refractories*, that is, they are capable of withstanding the high temperatures of molten metals. After the casting has solidified, the mold is broken up to remove the casting, hence the word *expendable*.

The mold is produced from a *pattern*; in some processes, the mold is expendable but the pattern is reused to produce several molds; such processes are referred to as *expendable-mold, permanent-pattern casting processes*. On the other hand, investment casting requires a pattern for each mold produced, an example of an *expendable-mold, expendable-pattern process*.

2. **Permanent molds**, made of metals that maintain their strength at high temperatures. As the name implies, the molds are used repeatedly, and are designed in such a manner that the casting can be removed easily and the mold used for the next casting. Metal molds are better heat conductors than expendable nonmetallic molds (see Table 3.1), thus the casting is subjected to a higher rate of cooling during solidification. This in turn affects the microstructure and grain size within the casting.

3. **Composite molds**, made of two or more different materials (such as sand, graphite, and metal), combining the advantages of each material. These molds have a permanent and an expendable portion, and are used in some casting processes to improve mold strength, control the cooling rate, and optimize the overall economics of the casting operation.

The general characteristics of sand casting and other casting processes are summarized in Table 11.2. Almost all commercial metals can be cast; the surface finish obtained is largely a function of the mold material, and can be very good, although, as expected, sand castings generally have rough, grainy surfaces. Dimensional tolerances generally are not as good as those in machining and other net-shape processes. However, intricate shapes, such as engine blocks and turbocharger impellers, can be made by casting.

Because of their unique characteristics and applications, particularly in making microelectronic devices (Part V), basic crystal-growing techniques are included in this chapter, which concludes with a brief overview of modern foundries.

11.2 Expendable-mold, Permanent-pattern Casting Processes

The major categories of expendable-mold, permanent-pattern casting processes are sand, shell mold, plaster mold, ceramic mold, and vacuum casting.

Table 11.2: General Characteristics of Casting Processes.

	Sand	Shell	Evaporative pattern	Plaster	Investment	Permanent mold	Die	Centrifugal
Typical materials cast	All	All	All	Nonferrous (Al, Mg, Zn, Cu)	All	All	Nonferrous (Al, Mg, Zn, Cu)	All
Weight (kg):								
Minimum	0.01	0.01	0.01	0.01	0.001	0.1	¡0.01	0.01
Maximum	No limit	100+	100+	50+	100+	300	50	5000+
Typical surface finish (R_a in μm)	5–25	1–3	5–25	1–2	0.3–2	2–6	1–2	2–10
Porosity[1]	3–5	4–5	3–5	4–5	5	2–3	1–3	1–2
Shape Complexity[1]	1–2	2–3	1–2	1–2	1	2–3	3–4	3–4
Dimensional accuracy[1]	3	2	3	2	1	1	1	3
Section thickness (mm):								
Minimum	3	2	2	1	1	2	0.5	2
Maximum	No limit	—	—	—	75	50	12	100
Typical dimensional tolerance (mm)	1.6–4 mm (0.25 mm for small parts)	±0.003		±0.005 – 0.010	±0.005	±0.015	±0.001 – 0.005	0.015
Equipment	3–5	3	2–3	3–5	3–5	2	1	1
Pattern/die	3–5	2–3	2–3	3–5	2–3	2	1	1
Labor	1–3	3	3	1–2	1–2	3	5	5
Typical lead time[2]	Days	Weeks	Weeks	Days	Weeks	Weeks	Weeks to months	Months
Typical production rate[2] (parts/mold-hour)	1–20	5–50	1–20	1–10	1–1000	5–50	2–200	1–1000
Minimum quantity[2]	1	100	500	10	10	1000	10,000	10–10,000

Notes: 1. Relative rating, from 1 (best) to 5 (worst). Note that, for example, a die casting has relatively low porosity, mid-to-low shape complexity, high dimensional accuracy, high equipment and die costs, and low labor costs. These ratings are general, as significant variations can occur, depending on the particular production method.
2. Approximate values, without using rapid prototyping technologies. Minimum quantity is 1 when applying rapid prototyping.
Source: Data taken from J.A. Schey, *Introduction to Manufacturing Processes*, 3rd ed., McGraw-Hill, 2000.

11.2.1 Sand Casting

The traditional method of casting metals is using sand molds, and has been done for millennia. Sand casting is still the most prevalent form of casting, led by China (about 45 million metric tons per year), followed by India and the U.S. Typical applications of sand casting include machine bases, large turbine impellers, propellers, and plumbing fixtures. The capabilities of sand casting are given in Table 11.2.

Sand casting basically consists of (a) placing a pattern, having the shape of the part to be cast, in sand to make an imprint, (b) incorporating a gating system for molten metal flow, (c) removing the pattern and filling the mold cavity with molten metal, (d) allowing the metal to cool until it solidifies, (e) breaking away the sand mold, and (f) removing the casting (Fig. 11.2).

Alternatively, a pattern can be machined directly into a sand preform. Since the strength of sand arises from its binders, machining can take place at high removal rates and can produce molds of high quality. An important and more recent development in mold and pattern making is the application of **additive manufacturing** (see Chapter 20) to directly produce molds. In sand casting, for example, a pattern can be binder-jet printed with complex shapes and self-supported cores to produce hollow sections, greatly easing mold assembly. There are several rapid prototyping techniques applicable to casting, and can produce molds or patterns; they are best suited for small production runs.

Sands. Most sand-casting operations use silica sand (SiO_2) as the mold material, although alternative sands and binders are under development, because of challenging health concerns associated with silica exposure in foundries. Sand is inexpensive and is suitable as a mold material because of its high-temperature characteristics and high melting point. There are two general types of sand: **naturally bonded** (*bank sand*) and **synthetic** (*lake sand*). Because its composition can be controlled more accurately, synthetic sand is preferred by most foundries.

Several factors are important in the selection of sand for molds, and certain tradeoffs with respect to properties have to be considered. Sand having fine, round grains can be packed closely, forming a smooth mold surface. Although fine-grained sand enhances mold strength, the fine grains also lower mold *permeability*. Good permeability of molds and cores allows gases and steam evolved during the casting process to escape easily. The mold also should have good *collapsibility*, in order to allow the casting to shrink while it is cooling, and thus prevent defects in the casting, such as hot tearing and cracking, shown in Fig. 10.14.

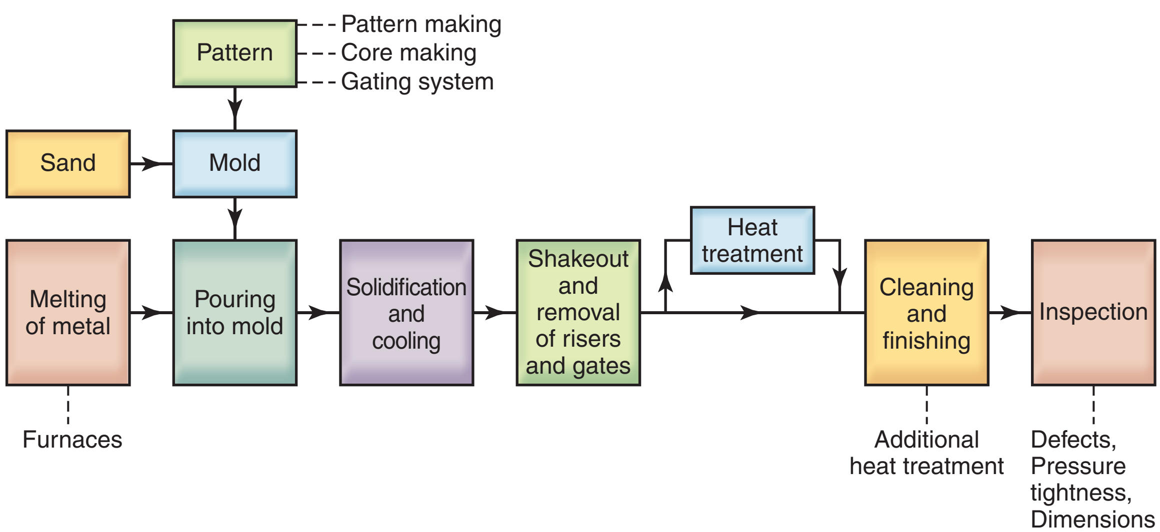


Figure 11.2: Outline of production steps in a typical sand-casting operation.

Types of Sand Molds. Sand molds (Fig. 11.3) are characterized by the types of sand and the methods used to produce them. There are three basic types of sand molds: (a) green-sand, (b) cold-box, and (c) no-bake molds. The most common mold material is **green molding sand**, a mixture of sand, clay, and water. The term green refers to the fact that the sand in the mold is moist or damp while the metal is being poured into it. Green-sand molding is the least expensive method of making molds, and the sand is recycled easily for subsequent reuse. In the *skin-dried* method, the mold surfaces are dried, either by storing the mold in air or by drying it with torches. Because of their higher strength, these molds are generally used for large castings.

In the **cold-box mold** process, various organic and inorganic *binders* are blended into the sand to bond the grains chemically for greater strength. These molds are more dimensionally accurate than green-sand molds, but are more expensive. In the no-bake mold process, a synthetic liquid resin is mixed with the sand, and the mixture hardens at room temperature. Because the bonding of the mold in this and in the cold-box process takes place without applying heat, they are called **cold-setting processes**.

Sand molds can be oven dried (*baked*) prior to pouring the molten metal; they then become stronger than green-sand molds and impart better dimensional accuracy and surface finish to the casting. However, this method has the drawbacks that (a) distortion of the mold is greater, (b) the castings are more susceptible to hot tearing, because of the lower collapsibility of the mold, and (c) production rate is lower, because of the significant drying time required.

The major features of molds in sand casting are:

1. The **flask**, which supports the mold itself. Two-piece molds consist of a **cope** on top and a **drag** on the bottom; the seam between them is the *parting line*. When more than two pieces are used in a sand mold, the additional parts are called *cheeks*.

2. A **pouring basin** or **pouring cup**, into which the molten metal is poured.

3. A **sprue**, through which the molten metal flows downward by gravity.

4. The **runner system**, which has channels that carry the molten metal from the sprue to the mold cavity. **Gates** are the inlets into the mold cavity.

5. **Risers**, which supply additional molten metal to the casting as it shrinks during solidification. Two types of risers, a *blind riser* and an *open riser*, are shown in Fig. 11.3.

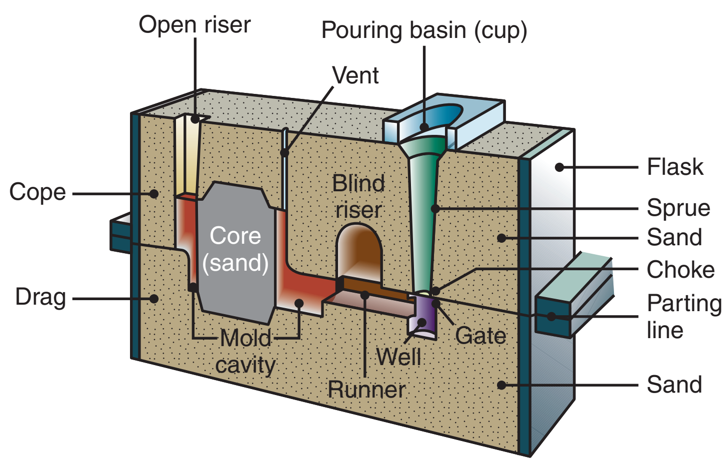


Figure 11.3: Schematic illustration of a sand mold, showing various features.

6. **Cores**, which are inserts made from sand and placed in the mold to form hollow regions or otherwise define the interior surface of the casting. Cores also are used on the outside of the casting to shape features, such as lettering and numbering.

7. **Vents**, which are placed in molds to carry off gases produced when the molten metal comes into contact with the sand in the mold and the core. They also exhaust air from the mold cavity as the molten metal flows into the mold.

Patterns. *Patterns* are used to mold the sand mixture into the shape of the casting, and may be made of wood, plastic, metal, or a combination of materials. Their selection depends on the size and shape of the casting, the dimensional accuracy and the quantity of castings required, and the molding process. Because patterns are used repeatedly to make molds, the strength and durability of the material selected for a pattern must reflect the number of castings that the mold is expected to produce. Patterns made of a combination of materials reduce wear in critical regions; they usually are also coated with a **parting agent** to facilitate the removal of the casting from the mold.

Patterns can be designed with a variety of features to fit specific applications and economic requirements. **One-piece patterns**, also called *loose* or *solid patterns*, are generally used for simpler shapes and low-quantity production; they generally are made of wood and are inexpensive. **Split patterns** have two pieces, made in such a way whereby each part forms a portion of the cavity for the casting; in this way, castings with complicated shapes can be produced. **Match-plate patterns** are a common type of mounted pattern in which two-piece patterns are constructed by securing each half of one or more split patterns to the opposite sides of a single plate (Fig. 11.4). In such constructions, the gating system can be mounted on the drag side of the pattern. This type is used most often in conjunction with molding machines and for large production runs for producing smaller castings.

Pattern design is a critical aspect of the total casting operation. The design should provide for **metal shrinkage**, permit proper metal flow in the mold cavity, and allow the pattern to be easily removed from the sand mold by means of a taper or *draft* (Fig. 11.5) or some other geometric feature (see also Chapter 12).

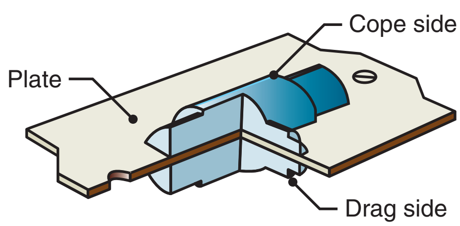


Figure 11.4: A typical metal match-plate pattern used in sand casting.

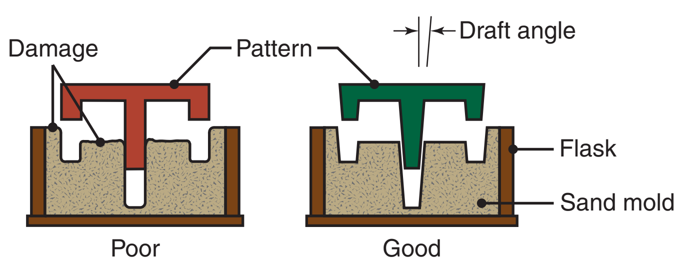


Figure 11.5: Taper on patterns for ease of removal from the sand mold.

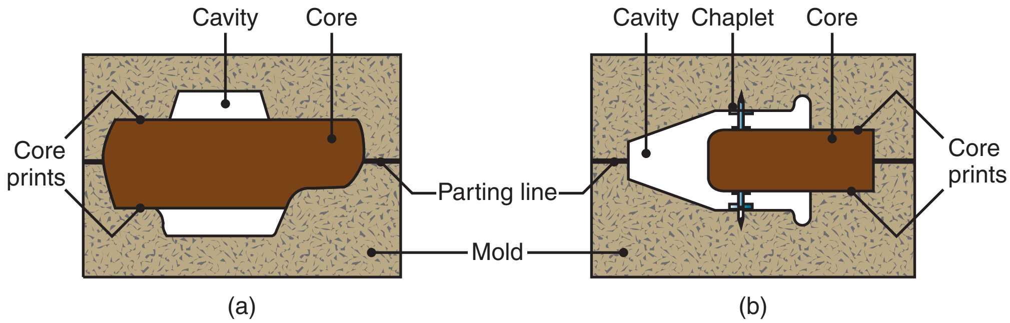


Figure 11.6: Examples of sand cores, showing core prints and chaplets to support the cores.

Cores. For castings with internal cavities or passageways, such as those found in automotive engine blocks or valve bodies, *cores* are placed in the mold cavity to form the interior surfaces of the casting. They are then removed from the finished part during shakeout and further processing. Like molds, cores must possess strength, permeability, collapsibility, and the ability to withstand heat; hence, they are made of sand aggregates. The core is anchored by **core prints**, geometric features added to the pattern in order to locate and support the core and to provide vents for the escape of gases (Fig. 11.6a). A common difficulty with cores is that, for some casting requirements (as in the case where a recess is required), they may lack sufficient structural support in the cavity. To keep the core from shifting, metal supports (**chaplets**) may be used to anchor the core in place (Fig. 11.6b).

Cores are generally made in a manner similar to that used in sand mold making; most are made using shell (see Section 11.2.2), no-bake, or cold-box processes. Cores are shaped in *core boxes* and used in much the same way that patterns are used to form sand molds.

Sand-molding Machines. The oldest known method of molding, which is still used for simple castings and for small production runs, is to compact the sand by hand hammering (*tamping*) or ramming it around the pattern. For most operations, the sand mixture is compacted around the pattern by *molding machines*. These machines manipulate the mold in a controlled manner, offer high-quality casting by improving the application and distribution of forces, and increase production rate.

In **vertical flaskless molding**, the pattern halves form a vertical chamber wall against which sand is blown and compacted (Fig. 11.7). The mold halves are then packed horizontally, with the parting line oriented vertically and moved along a pouring conveyor. The operation is simple and eliminates the need to handle flasks, allowing for very high production rates, particularly when other aspects of the operation (such as coring and pouring) are all automated.

Sandslingers fill the flask uniformly with sand under a high-pressure stream; often automated, they are used to fill large flasks. An impeller in the machine throws sand from its blades or cups at such high speeds that the machine not only places the sand but also rams it sufficiently for proper packing.

In **impact molding**, the sand is compacted by a controlled explosion or instantaneous release of compressed gases. This method produces molds with uniform strength and good permeability.

In **vacuum molding** (also known as the *V process*), shown in Fig. 11.8, the pattern is covered tightly with a thin sheet of plastic. A flask is placed over the covered pattern and is then filled with dry, binderless sand. A second sheet of plastic is then placed on top of the sand, and a vacuum action compacts the sand. Both halves of the mold are made in this manner and are subsequently assembled. During pouring of the molten metal, the mold remains under a vacuum, but not the casting cavity. When the hot metal has solidified, the vacuum is turned off and the sand falls away, releasing the casting.

As shown in Fig. 11.8, vacuum molding does not require a draft in the part, and can be very economical because of the low tooling costs, long pattern life, and absence of binders in the sand (also simplifying sand recovery and reuse). Vacuum molding produces castings with high surface detail and dimensional accuracy; it is suited especially well for large, relatively flat (plane) castings.

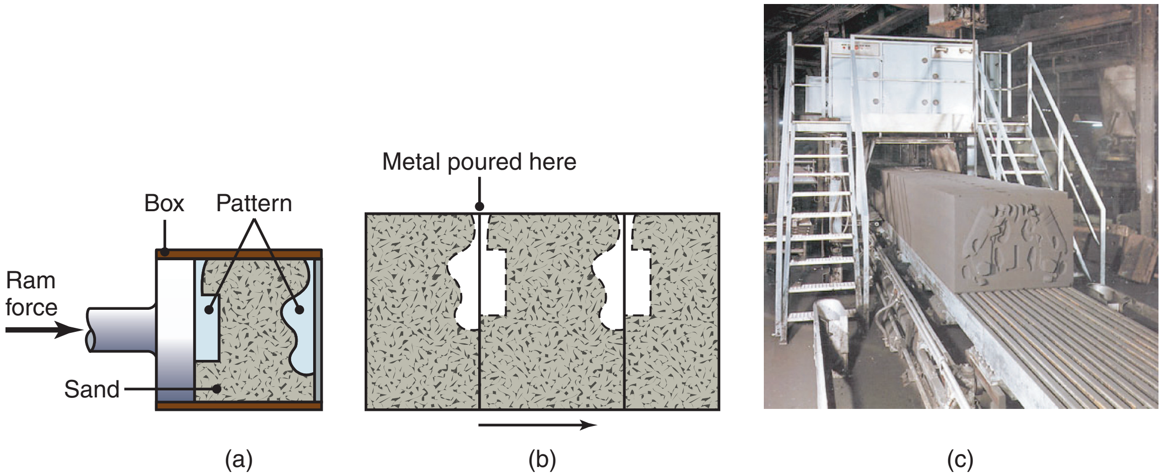


Figure 11.7: Vertical flaskless molding. (a) Sand is squeezed between two halves of the pattern. (b) Assembled molds pass along an assembly line for pouring. (c) A photograph of a vertical flaskless molding line. *Source:* Courtesy of American Foundry Society.

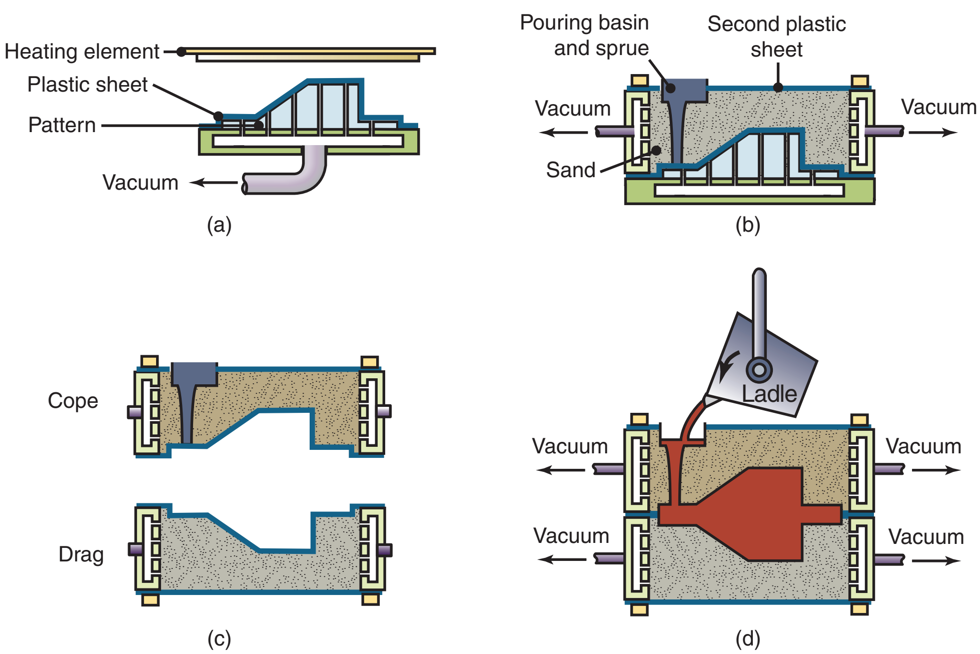


Figure 11.8: The vacuum molding process. (a) A plastic sheet is first thermoformed (see Section 19.6) over a pattern; (b) a vacuum flask is then placed over the pattern, a pouring basin/sprue insert is located, and the flask is filled with sand. A second sheet is located on the top of the sand mold, and a vacuum is applied to tightly compact the sand against the pattern. (c) A drag is also produced, along with cheeks, cores, etc., as in conventional sand casting. The cope and drag can be carefully transported without vacuum being applied. (d) After the mold halves are joined, a vacuum is applied to ensure mold strength, and molten metal is poured into the mold.

The Sand-casting Operation. After the mold has been shaped and the cores have been placed in their positions, the two mold halves (cope and drag) are closed, clamped, and weighted down, to prevent the separation of the mold sections under the pressure exerted when the molten metal is poured into the mold cavity. A complete sequence of operations in sand casting is shown in Fig. 11.9.

After solidification, the casting is shaken out of its mold, and the sand and oxide layers adhering to the casting are removed (by means of vibration, using a shaker, or by sand blasting). Castings are also cleaned

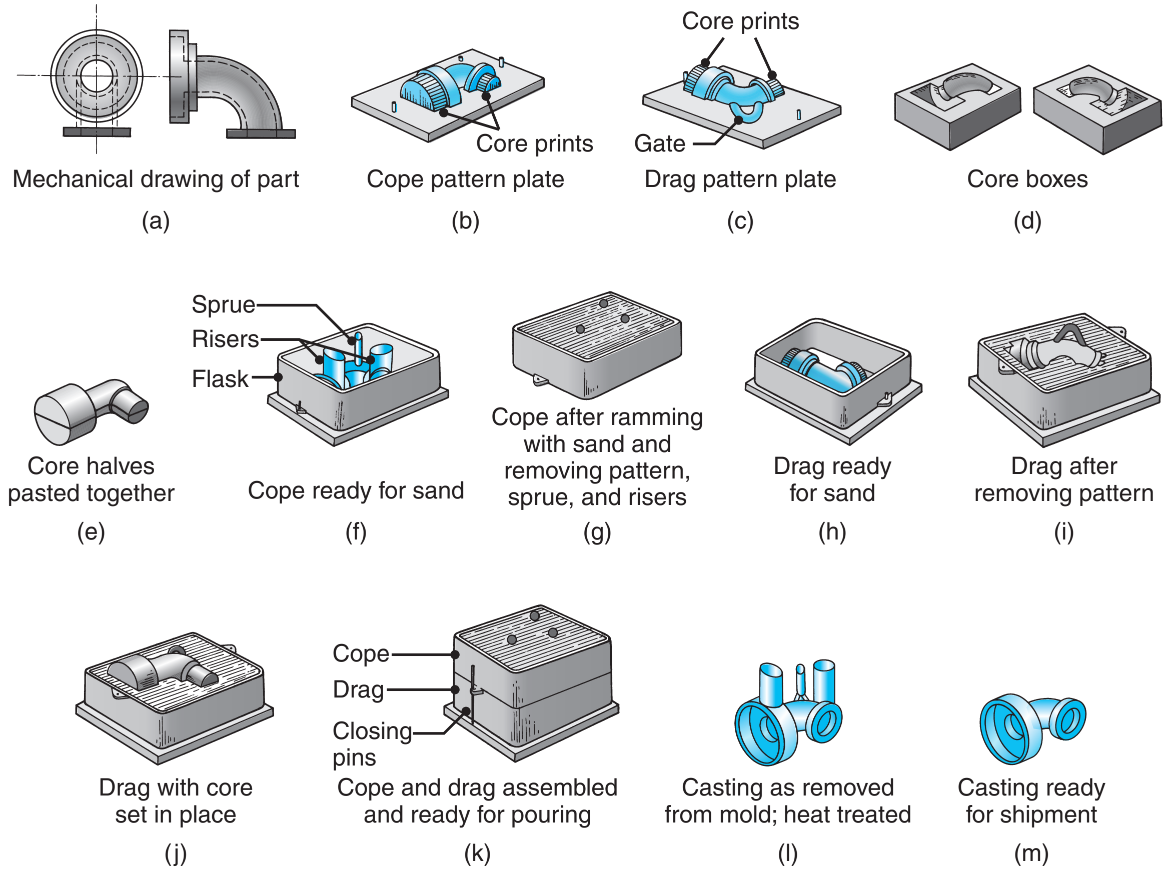


Figure 11.9: Schematic illustration of the sequence of operations for sand casting. (a) A mechanical drawing or CAD representation of the part is used to generate a design for the pattern. Considerations such as part shrinkage and draft must be included into the drawing. (b–c) Patterns have been mounted on plates equipped with pins for alignment. Note also the presence of core prints designed to hold the core in place. (d–e) Core boxes produce core halves, which are pasted together. The cores will be used to produce the hollow area of the part shown in (a). (f) The cope half of the mold is assembled by securing the cope pattern plate to the flask with aligning pins and attaching inserts to form the sprue and risers. (g) The flask is rammed with sand, and the plate and inserts are removed. (h) The drag half is produced in a similar manner, with the pattern inserted. A bottom board is placed below the drag and is aligned with pins. (i) The pattern, flask, and bottom board are inverted, and the pattern is withdrawn, leaving the appropriate imprint. (j) The core is set in place within the drag cavity. (k) The mold is closed by placing the cope on top of the drag and securing the assembly with pins. The flasks are then subjected to pressure to counteract buoyant forces in the molten metal, which might lift the cope. (l) After the metal solidifies, the casting is removed from the mold. (m) The sprue and risers are cut off and recycled, and the casting is cleaned, inspected, and heat treated (if necessary). *Source:* Courtesy of Steel Founders' Society of America.

by blasting with steel shot or grit (*shot blasting*; Section 26.8). The risers and gates are cut off either by oxyfuel-gas cutting, sawing, shearing, or abrasive wheels; they may also be trimmed in appropriate dies. Gates and risers on steel castings may also be removed with air carbon-arc cutting (Section 30.8) or torches. Castings may be further cleaned by electrochemical means or by pickling with chemicals to remove surface oxides (see Section 34.16).

The casting may subsequently be *heat treated* (Chapter 4) to improve certain properties required for its intended use; heat-treatment is particularly important for steel castings. *Finishing operations* may involve machining, straightening, or forging with dies (sizing) to obtain final dimensions. *Inspection* is an important final step, and is carried out to ensure that the casting meets all design and quality-control requirements.

Rammed-graphite Molding. In this process, rammed graphite (Section 8.6) is used to make molds for casting reactive metals, such as titanium and zirconium; sand cannot be used because these metals react vigorously with silica. The molds are packed like sand molds, air dried, baked at 175°C, fired at 870°C, and then stored under controlled humidity and temperature. The casting procedures are similar to those for sand molds.

Mold Ablation. Ablation can be used to improve the mechanical properties and production rates in sand casting. In this process, a sand mold is filled with molten metal, and the mold is then immediately sprayed with a liquid and/or gas solvent to progressively erode the sand. As the mold is exposed, the liquid stream causes rapid and directional solidification of the metal. With properly designed risers, mold ablation results in significantly lower porosity than conventional sand casting, leading to higher strength and ductility; it has therefore been applied to normally difficult-to-cast materials or for metal-matrix composites. Since ablation speeds up solidification and also removes cores, significant productivity improvements can also be achieved.

11.2.2 Shell Molding

Shell molding, first developed in the 1940s, has grown significantly because it can produce numerous types of castings with close dimensional tolerances and good surface finish, and at low cost. Shell-molding applications include small mechanical parts requiring high precision, such as gear housings, cylinder heads, and connecting rods. The process is also used widely in producing high-precision molding cores.

The capabilities of **shell-mold casting** are given in Table 11.2. In this process, a mounted pattern, made of a ferrous metal or aluminum, is (a) heated to a range of 175°–370°C, (b) coated with a parting agent (such as silicone), and (c) clamped to a box or chamber. The box contains fine sand, mixed with 2.5–4% of a thermosetting resin binder (such as phenol-formaldehyde), which coats the sand particles. Either the box is rotated upside down (Fig. 11.10) or the sand mixture is blown over the pattern, allowing it to form a coating.

The assembly is then placed in an oven for a short period of time to complete curing of the resin. In most shell-molding machines, the oven consists of a metal box, with gas-fired burners that swing over the shell mold and cure it. The shell hardens around the pattern and is removed by means of built-in ejector pins. Two half-shells are made in this manner and are bonded or clamped together to form a mold.

The thickness of the shell can be determined accurately by controlling the time that the pattern is in contact with the mold. In this way, the shell can be formed with the required strength and rigidity to hold the weight of the molten liquid. The shells are light and thin, usually 5 to 10 mm, and, consequently, their thermal characteristics are different from those for thicker molds.

Since a much smaller grain size is used in shell molding, shell sand has a much lower permeability than the sand for green-sand molding. The decomposition of the shell-sand binder produces a high volume of gas; consequently, unless the molds are vented properly, trapped air and gas can produce defects in shell molding of ferrous castings. The high quality of the finished casting can reduce cleaning, machining, and other finishing costs significantly. Complex shapes can be produced with less labor, and the process can be automated.

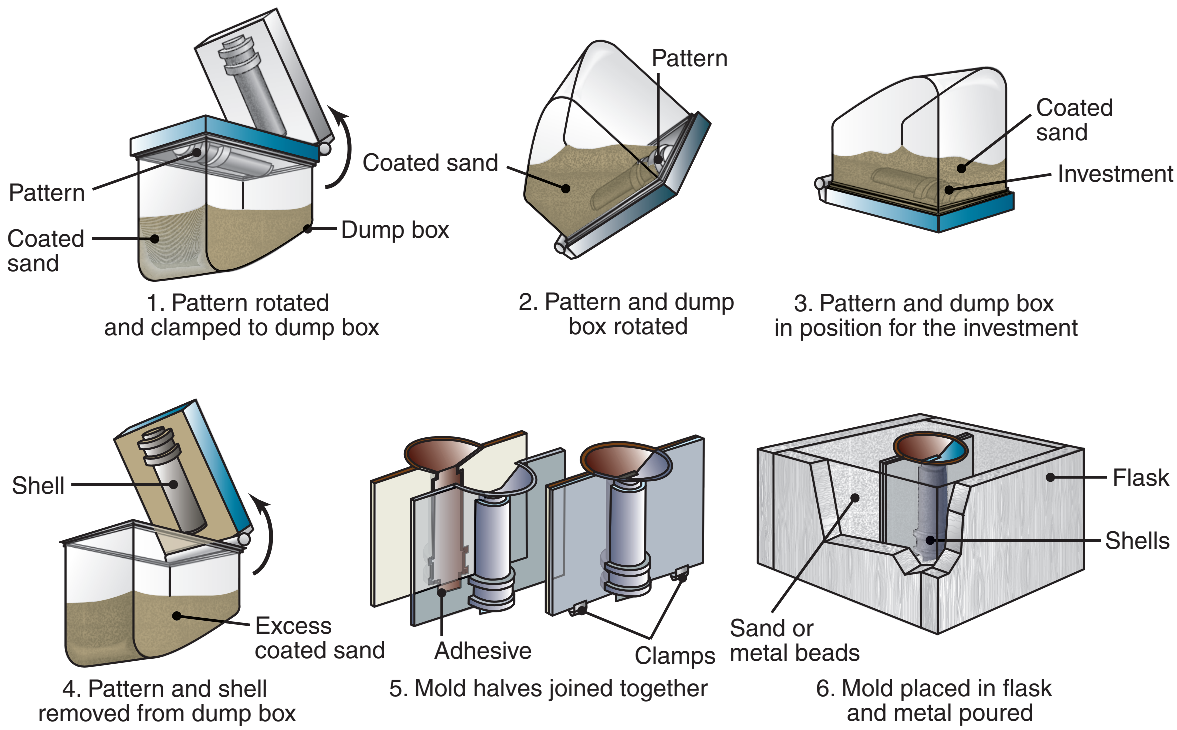


Figure 11.10: The shell-molding process, also called the dump-box technique.

11.2.3 Plaster-mold Casting

This process, and the ceramic-mold and investment casting processes (described in Sections 11.2.4 and 11.3.2) are known as **precision casting**, because of the high dimensional accuracy and good surface finish obtained. Typical parts made are lock components, gears, valves, fittings, tooling, and ornaments. They weigh in the range of 125–250 g, although parts as light as 1 g have been made. The capabilities of plaster-mold casting are given in Table 11.2.

In the *plaster-molding process*, the mold is made of plaster of paris (gypsum or calcium sulfate), with the addition of talc and silica powder to improve strength and to control the time required for the plaster to set. The three components are mixed with water and the slurry is poured over the pattern. After the plaster sets, usually within 15 minutes, it is removed and the mold is dried, at a typical temperature range of 120° to 260°C. The mold halves are then assembled to form the mold cavity and are preheated to about 120°C. The molten metal is then poured into the mold.

Because plaster molds have very low permeability, gases evolved during solidification of the metal cannot escape; consequently, the molten metal is poured either in a vacuum or under pressure. Mold permeability can be increased significantly by the *Antioch process*, in which the molds are dehydrated in an autoclave (pressurized oven) for 6 to 12 hours, and then rehydrated in air for 14 hours. Another method of increasing the permeability of the mold is to use foamed plaster, containing trapped air bubbles.

Patterns for plaster molding are generally made of aluminum alloys or thermosetting plastics, but brass or zinc alloys are also used. Since there is a limit to the maximum temperature that the plaster mold can withstand (generally about 1200°C), plaster-mold casting is used only for aluminum, magnesium, zinc, and some copper-based alloys. The castings have a good surface finish with fine details. Also, because plaster molds have lower thermal conductivity than other mold materials, the castings cool slowly; a more uniform grain structure is obtained and with less warpage. The wall thickness of the parts made can be as thin as 1 to 2.5 mm.

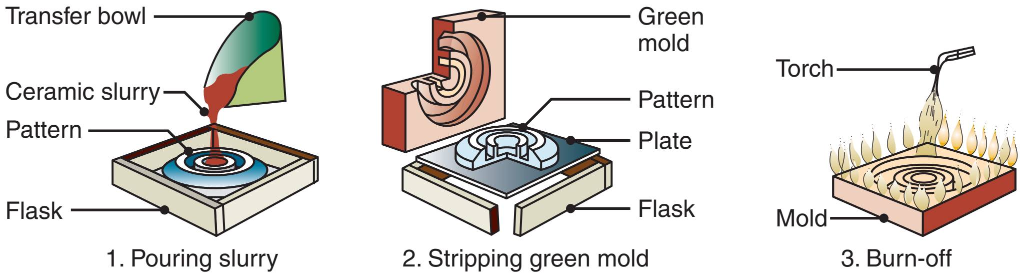


Figure 11.11: Sequence of operations in making a ceramic mold. *Source: Metals Handbook*, Vol. 5, 8th ed, ASM International, 1970.

11.2.4 Ceramic-mold Casting

This process, also called *cope-and-drag investment casting*, is similar to the plaster-mold process, except that it uses refractory mold materials suitable for high-temperature applications. A slurry is first produced from a mixture of fine-grained zircon ($ZrSiO_4$), aluminum oxide, fused silica, and bonding agents; this slurry is then poured over the pattern (Fig. 11.11) which has been placed in a flask. Typical parts made are impellers, cutters for machining operations, dies for metalworking operations, and molds for casting plastic and rubber components. Parts weighing as much as 700 kg have been cast by this process.

After setting, the molds (ceramic facings) are removed, dried, ignited to burn off volatile matter, and baked. The molds are then clamped firmly and used as all-ceramic molds. In the *Shaw process*, the ceramic facings are backed by fireclay (which resists high temperatures) to give strength to the mold. The facings are later assembled into a complete mold, ready to be used.

The high-temperature resistance of refractory molding materials allows the molds to be used for casting ferrous and other high-temperature alloys, stainless steels, and tool steels. Although the process is somewhat expensive, the castings have good dimensional accuracy and surface finish over a wide range of sizes and intricate shapes.

11.3 Expendable-mold, Expendable-pattern Casting Processes

Evaporative-pattern and investment casting are also referred to as *expendable-pattern* casting processes or *expendable mold–expendable pattern* processes. They are unique in that a mold and a pattern has to be produced for each casting, whereas the patterns in the processes described in the preceding section are all reusable. Typical applications of these processes are cylinder heads, engine blocks, crankshafts, brake components, and machine bases.

11.3.1 Evaporative-pattern Casting (Lost-foam Process)

The *evaporative-pattern casting* (EPC) process uses a polystyrene pattern, which evaporates upon contact with molten metal to form a cavity for the casting; this process is also known as *lost-foam casting*, or the *full-mold casting* (FMC) process. It has become one of the more important casting processes for ferrous and nonferrous metals, particularly for the automotive industry.

In this process, polystyrene beads, containing 5–8% pentane (a volatile hydrocarbon), are placed in a preheated die that is usually made of aluminum. Complex patterns may be made by bonding various individual pattern sections, using a hot-melt adhesive (Section 32.4.1). Polymethylmethacrylate (PMMA) and polyalkylene carbonate also may be used as pattern materials for ferrous castings.

The polystyrene expands and takes the shape of the die cavity; additional heat is applied to fuse and bond the beads together. The die is cooled and opened, and the polystyrene pattern is removed. The pattern is then coated with water-based refractory slurry, dried, and placed in a flask. The flask is filled with loose,

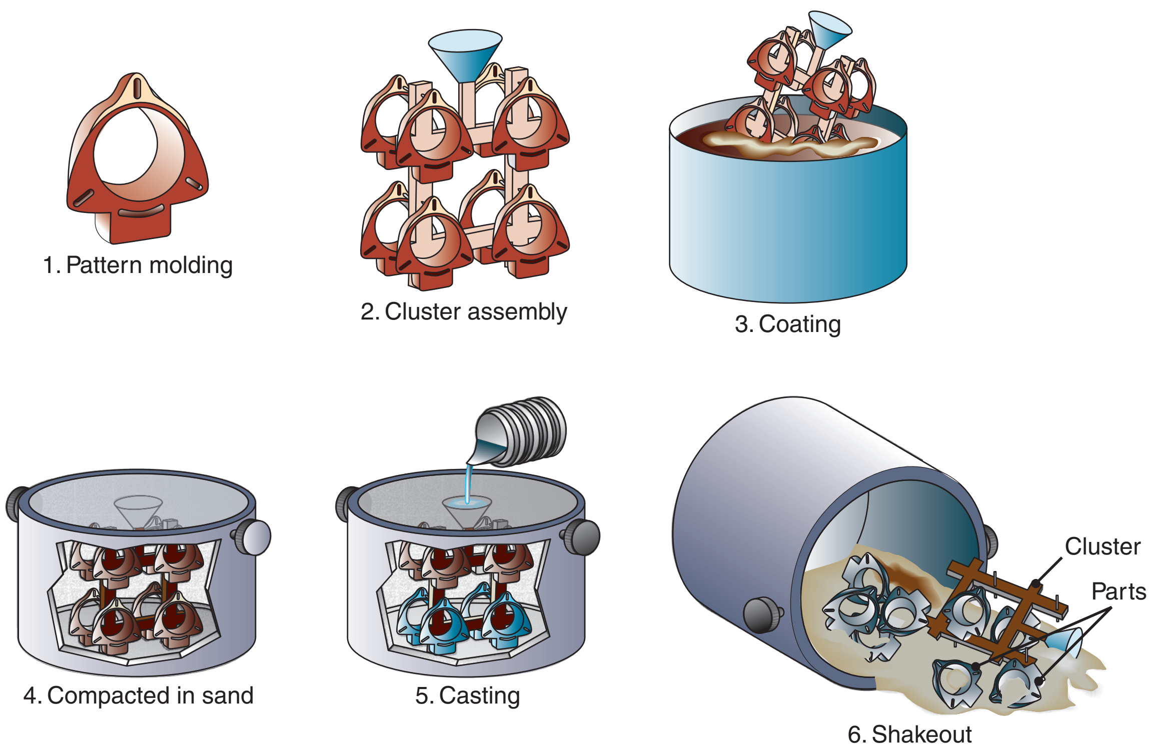


Figure 11.12: Schematic illustration of the expendable-pattern casting process, also known as lost-foam or evaporative-pattern casting.

fine sand, which surrounds and supports the pattern (Fig. 11.12), and may be dried or mixed with bonding agents to give it additional strength. The sand is compacted periodically, without removing the polystyrene pattern; then the molten metal is poured into the mold. The molten metal vaporizes the pattern and fills the mold cavity, completely replacing the space previously occupied by the polystyrene. Any degradation by-products from the polystyrene are vented into the surrounding sand.

Because the polymer requires considerable energy to degrade, large thermal gradients are present at the metal–polymer interfaces. In other words, the molten metal cools faster than it would if it were poured directly into an empty cavity; consequently, fluidity is less than in sand casting. This has important effects on the microstructure throughout the casting, and also leads to directional (columnar) solidification of the metal (see Section 10.2.3).

The evaporative-pattern process has several advantages over other casting methods:

- The process is relatively simple, because there are no parting lines, cores, or riser systems.
- The flasks used for the process are inexpensive.
- Polystyrene is inexpensive, and can be processed easily into patterns having complex shapes, various sizes, and fine surface detail.
- The casting requires minimal finishing and cleaning operations.
- The process can be automated and is economical for long production runs; however, the cost of producing the die and the need for two sets of tooling are significant factors to consider.

In a modification of the evaporative-pattern process, called the *Replicast® C-S process*, a polystyrene pattern is surrounded by a ceramic shell; then the pattern is burned out prior to pouring the molten metal into the mold. Its principal advantage over investment casting (which uses wax patterns, Section 11.3.2) is that carbon pickup into the metal is avoided. Further developments in EPC include the production of metal-matrix composites (Sections 9.5 and 19.14). During molding of the polymer pattern, fibers or particles are embedded throughout the part, which then become an integral part of the casting. Other techniques include the modification and grain refinement of the casting, by using grain refiners and modifier master alloys.

Case Study 11.1 Lost-foam Casting of Engine Blocks

One of the most important components in an internal combustion engine is the engine block. Industry trends have focused upon high-quality, low-cost and lightweight designs. Economic benefits can be gained through casting more complex geometries and by incorporating multiple components into one part. Recognizing that EPC can simultaneously satisfy all of these requirements, Mercury Castings built a lost-foam casting line to produce aluminum engine blocks and cylinder heads.

One example of a part produced through lost-foam casting is a 45-kW, three-cylinder engine block used for marine applications, such as an outboard motor on a small boat, and illustrated in Fig. 11.13c. Previously manufactured as eight separate die castings, the block was converted to a single 10-kg casting, with a weight and cost savings of 1 kg and $25, respectively, on each block. The casting chosen also allowed consolidation of the engine's cylinder head and the exhaust and cooling systems into the block, thus eliminating the associated machining operations and fasteners required in sand-cast or die-cast designs. Moreover, since the pattern contained holes, which could be cast without the need for cores, numerous drilling operations were eliminated.

Mercury Marine also was in the midst of developing a new V6 engine, utilizing a new corrosion-resistant aluminum alloy with increased wear resistance. This engine design also required the integration of the cylinder block and the engine head, featuring hollow sections for water jacket cooling that could not be cored out in die casting or semipermanent mold processes (which were used for other V6 blocks). Based on the success that the foundry had with the three-cylinder lost-foam block, engineers applied this process for casting the V6 die block (Fig. 11.13b). The new engine block involves only one casting, that is lighter and less expensive than the previous designs. Produced with an integrated cylinder head and exhaust and cooling system, this component is cast hollow to develop more efficient water jacket cooling of the engine during its operation.

The company also developed a pressurized lost-foam process. First, a foam pattern is made, placed in a flask, and surrounded by sand. Then the flask is inserted into a pressure vessel, where a robot pours molten aluminum onto the polystyrene pattern. A lid on the pressure vessel is closed, and a pressure of 1 MPa is applied to the casting until it solidifies, in about 15 minutes. The result is a casting with better dimensional accuracy, lower porosity, and improved strength compared to conventional lost-foam casting.

Source: Courtesy of Mercury Marine.

11.3.2 Investment Casting

The *investment-casting* process, also called the **lost-wax process**, was first used during the period from 4000 to 3000 B.C. Typical parts made are components for office equipment and mechanical components, such as gears, cams, valves, and ratchets. Parts up to 1.5 m in diameter and weighing as much as 1140 kg have been cast successfully. The capabilities of investment casting are given in Table 11.3.

The sequence of operations involved in investment casting is shown in Fig. 11.14. The pattern is made of wax or of a plastic, such as polystyrene, by molding or rapid-prototyping techniques (Chapter 20). It

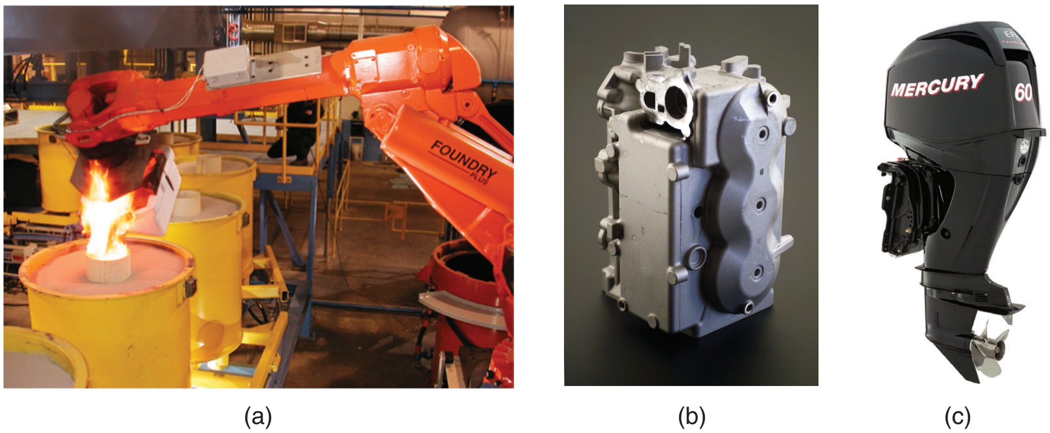


Figure 11.13: (a) Metal is poured into a mold for lost-foam casting of a 45-kW, three-cylinder marine engine; (b) finished engine block; (c) completed outboard motor. *Source:* Mercury Marine, a division of Brunswick Corporation.

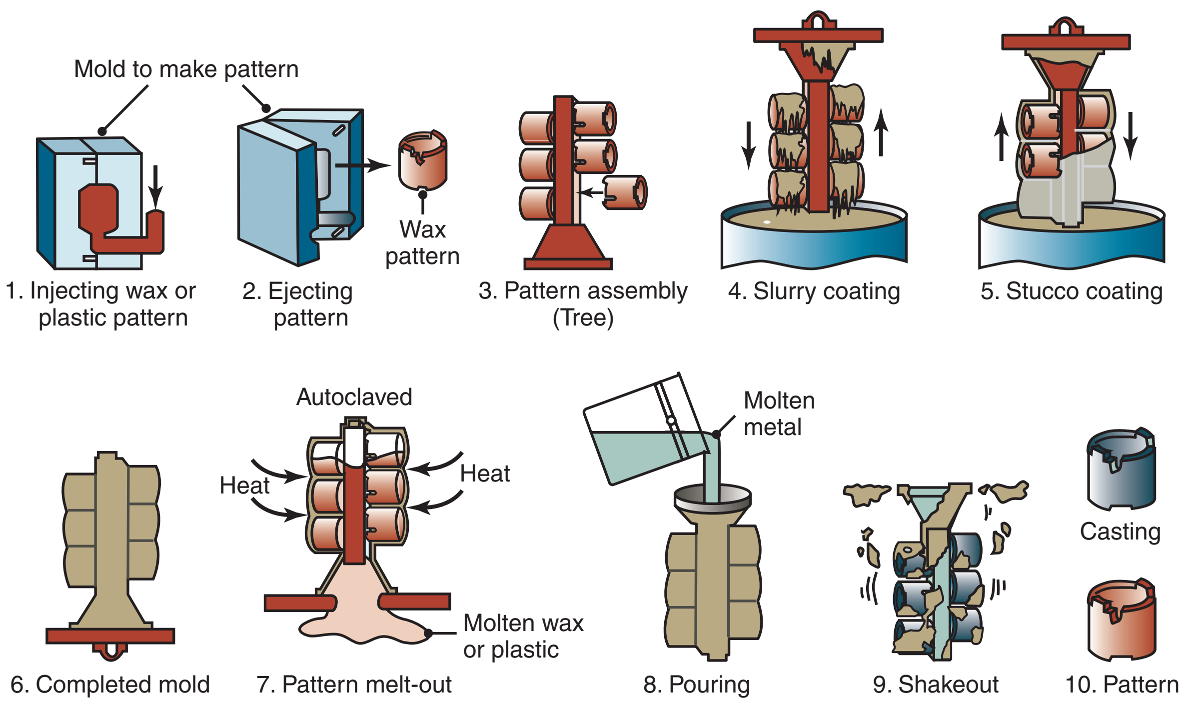


Figure 11.14: Schematic illustration of the investment-casting (lost-wax) process. Castings produced by this method can be made with very fine detail and from a variety of metals. *Source:* Courtesy of Steel Founders' Society of America.

is then dipped into a slurry of refractory material, such as very fine silica and binders, including water, ethyl silicate, and acids. After this initial coating has dried, the pattern is coated repeatedly to increase its thickness, for higher strength. Note that smaller particles can be used for the initial coating to develop a better surface finish in the casting; subsequent layers use larger particles and are intended to increase the coating thickness quickly.

The term *investment* derives from the fact that the pattern is *invested* (surrounded) with the refractory material. Wax patterns require careful handling because they are not sufficiently strong to withstand the forces encountered during mold making; unlike plastic patterns, however, wax can be recovered and reused.

The one-piece mold is then dried in air and heated to a temperature of 90° to 175°C. It is held in an inverted position for a few hours to melt out the wax. The mold is then fired to 650° to 1050°C for about four hours (depending on the metal to be cast), to drive off the water of crystallization (chemically combined water) and to burn off any residual wax. After the metal has been poured and has solidified, the mold is broken up and the casting is removed.

A number of patterns can be joined together to make one mold, called a **tree** (Fig. 11.14), significantly increasing the production rate. For small parts, the tree can be inserted into a permeable flask and filled with a liquid slurry. The investment is then placed into a chamber and evacuated (to remove any air bubbles) until the mold solidifies. The flask is usually placed in a vacuum-casting machine, so that the molten metal is drawn into the permeable mold and onto the part, thus producing fine detail.

Although the mold materials and the labor involved make the lost-wax process costly, it is suitable for casting high-melting-point alloys, with good surface finish and close dimensional tolerances. Few or no finishing operations are required, which otherwise would add significantly to cost of the casting. The process is capable of producing intricate shapes from a wide variety of ferrous and nonferrous metals and alloys, with parts weighing from 1 g to 35 kg. Advances include the casting of titanium aircraft-engine and structural airframe components, with wall thicknesses on the order of 1.5 mm, thus competing with previously used sheet-metal structures (see Chapter 40).

Ceramic-shell Investment Casting. A variation of the investment-casting process is *ceramic-shell casting*. It uses the same type of wax or plastic patterns, which is first dipped in ethyl silicate gel and, subsequently into a fluidized bed (see Section 4.12) of fine-grained fused silica or zircon powder. The pattern is then dipped into coarser grained silica, to build up additional coatings and develop a proper thickness so that the pattern can withstand the thermal shock due to pouring of the hot metal. The rest of the procedure is similar to investment casting. The process is economical and is used extensively for the precision casting of steels and high-temperature alloys.

The sequence of operations involved in making a turbine disk by this method is shown in Fig. 11.25. If the cores are made of ceramics, they are later removed by leaching with caustic solutions under high pressure and temperature. The molten metal may also be poured in a vacuum, to extract evolved gases and reduce oxidation, thus improving the casting quality. To further reduce microporosity, the castings made by this, as well as other processes, are subjected to hot isostatic pressing.

Case Study 11.2 Investment Casting of Total Knee Replacements

With major advances in medical care, life expectancies have increased significantly, so the expectations for the quality of life in the later years of a person’s life remain high. One of the reasons for improvement has been the great success of orthopedic implants. Hip, knee, shoulder, spine, and other implants have resulted in greatly increased activity and reduced pain for millions worldwide.

An example of an orthopedic implant that has greatly improved quality of life is total knee replacement (TKR), as shown in Fig. 11.15a. TKRs are very popular and reliable for the relief of osteoarthritis, a chronic and painful degenerative condition of the knee joint that typically sets in after middle age. TKRs consist of multiple parts, including femoral, tibial, and patellar components. Typical materials used include cobalt alloys, titanium alloys, and ultrahigh-molecular-weight polyethylene (UHMWPE). Each material is chosen for specific properties important in the application of the implant.

This case study describes the investment casting of femoral components of TKRs, which are produced from cobalt–chrome alloy (Section 6.6). The manufacturing process begins with injection molding of the patterns, which are then hand assembled onto *trees*, as shown in Fig. 11.15b. The patterns are spaced properly on a central wax sprue; they are then welded in place by dipping them into molten wax and pressing them against the sprue until the patterns are held in place. The final assembled tree, shown in Fig. 11.16a, contains 12 knee implants arranged in four rows.

The completed trees are then placed in a rack, where they form a queue and are then taken in order by an industrial robot (Section 37.6). The robot follows a set sequence in building up the mold. It first dips the pattern into dilute slurry, then rotates it under a sifting of fine particles. Next, the robot moves the tree beneath a blower to quickly dry the ceramic coating, and then it repeats the cycle. After a few cycles of such exposure to dilute slurry and fine particles, the details of the patterns are well produced, and good surface finish is ensured. The robot then dips the pattern into a thicker slurry which quickly builds up the mold thickness (Fig. 11.16c). The trees are then dried and placed into a furnace to melt out and burn the wax. The trees are placed into another furnace to preheat them in preparation for casting.

Figure 11.16 shows the progression of investment casting, from tree, to investment, to casting. A mold, ready for investment casting, is placed into a casting machine. The mold is placed upside down on the machine, directly over a measured volume of molten cobalt chrome alloy. The machine then rotates so that the metal flows into the mold, as shown in Fig. 11.15d. The tree is allowed to cool and the mold is removed; the cast parts are machined from the tree and are further machined and polished to the required surface finish and dimensional tolerance.

Source: Courtesy of M. Hawkins, Zimmer Biomet, Inc.

11.4 Permanent-mold Casting Processes

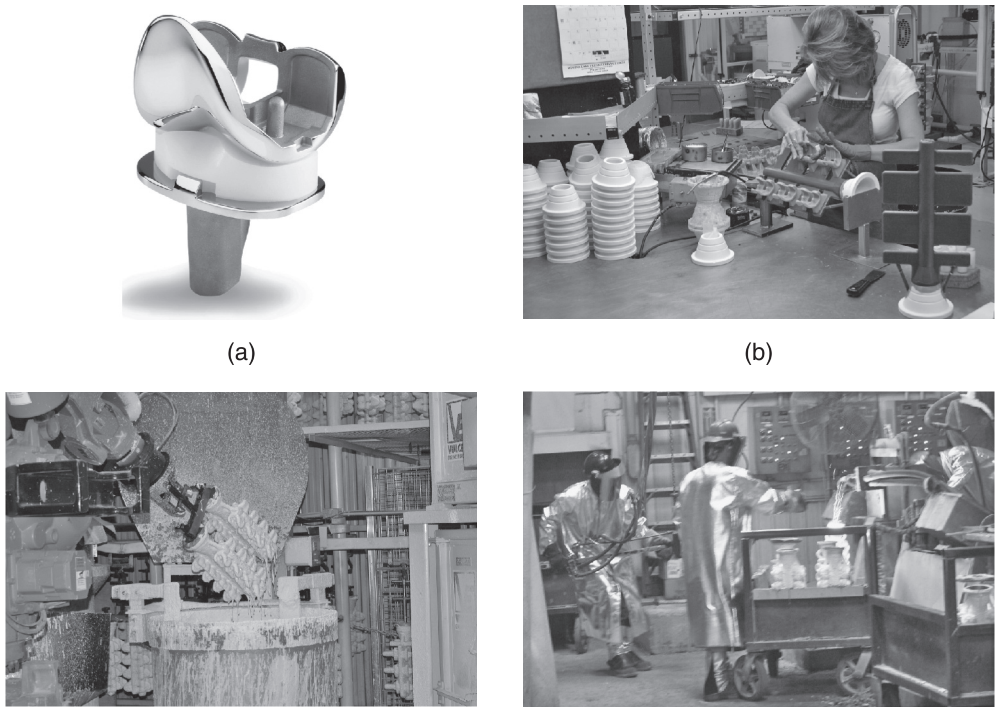
Permanent-mold casting processes have certain advantages over other casting processes, as described below.

11.4.1 Permanent-mold Casting

In *permanent-mold casting* (also called *hard-mold casting*), two halves of a mold are made from such materials as cast iron, steel, bronze, graphite, or refractory metal alloys, with high resistance to erosion and thermal fatigue. Typical parts made are automobile pistons, cylinder heads, connecting rods, gear blanks for appliances, and kitchenware. Parts that can be made economically typically weigh less than 25 kg, although special castings, weighing a few hundred kilograms, have been made using this process. The capabilities of permanent-mold casting are given in Table 11.3.

The mold cavity and the gating system are machined into the mold and thus become an integral part of the mold. To produce castings with internal cavities, cores, made of metal or sand aggregate, are placed in the mold prior to casting. Typical core materials are oil-bonded or resin-bonded sand, plaster, graphite, gray iron, low-carbon steel, and hot-work die steel. Gray iron is used most commonly, particularly for large molds for aluminum and magnesium casting. Inserts also are used in various locations of the mold.

In order to increase the life of permanent molds, the surfaces of the mold cavity are usually coated with a refractory slurry, such as sodium silicate and clay, or are sprayed with graphite every few castings. These coatings also serve as parting agents and as thermal barriers, thus controlling the rate of cooling of the casting. Mechanical ejectors, such as pins located in various parts of the mold, may be required for the removal of complex castings. Ejectors usually leave small round impressions, which generally are not significant.



(a) (b)

(c) (d)

Figure 11.15: Manufacture of total knee replacements. (a) The Zimmer NexGen mobile-bearing knee (MBK); the femoral portion (top component) of the total knee replacement is the subject of Case Study 11.2. (b) Assembly of patterns onto a central tree. (c) Dipping of the tree into slurry to develop a mold from investment. (d) Pouring of metal into a mold. *Source:* Courtesy of M. Hawkins, Zimmer, Inc.

(a) (b) (c)

Figure 11.16: Progression of the tree. (a) After assembly of blanks onto the tree; (b) after coating with investment; (c) after removal from the mold. *Source:* Courtesy of M. Hawkins, Zimmer Biomet, Inc.

Table 11.3: Properties and Typical Applications of Some Common Die-casting Alloys. *Source:* American Die Casting Institute.

Alloy	Ultimate tensile strength (MPa)	Yield strength (MPa)	Elongation in 50 mm (%)	Applications
Aluminum				
380 (3.5 Cu–8.5 Si)	320	160	2.5	Appliances, automotive components, electrical motor frames and housings
13 (12 Si)	300	150	2.5	Complex shapes with thin walls, parts requiring strength at elevated temperatures
Brass 858 (60 Cu)	380	200	15	Plumbing fixtures, lock hardware, bushings, ornamental castings
Magnesium AZ91 B (9 Al–0.7 Zn)	230	160	3	Power tools, automotive parts, sporting goods
Zinc				
No. 3 (4 Al)	280	—	10	Automotive parts, office equipment, household utensils, building hardware, toys
No. 5 (4 Al–1 Cu)	320	—	7	Appliances, automotive parts, building hardware, business equipment

The two molds are clamped together by mechanical means, and heated to about 150° to 200°C to facilitate metal flow and reduce thermal damage to the dies. Molten metal is then poured through the gating system; after solidification, the molds are opened and the casting is removed. The mold often incorporates special cooling features, such as a means for pumping cooling water through the channels located in the mold and the use of cooling fins. Although the permanent-mold casting operation can be performed manually, it is often automated for large production runs.

This process is used mostly for aluminum, magnesium, and copper alloys, as well as for gray iron because of their generally lower melting points; however, steels also can be cast using graphite or heat-resistant metal molds. Permanent-mold casting produces castings with a good surface finish, close dimensional tolerances, uniform and good mechanical properties, and at high production rates.

Although equipment costs can be high because of high die costs, labor costs are kept low through automation. The process is not economical for small production runs and is not suitable for intricate shapes, because of the difficulty in removing the casting from the mold. However, in a process called **semipermanent mold casting**, easily collapsible sand cores can be used, which are then removed from castings, leaving intricate internal cavities.

11.4.2 Vacuum Casting

A schematic illustration of the *vacuum-casting* process, also called *countergravity low-pressure* (CL) *process* (not to be confused with the vacuum molding process described in Section 11.2.1) is shown in Fig. 11.17. Vacuum casting is an alternative to investment, shell-mold, and green-sand casting, and is suitable particularly for thin-walled complex shapes. With uniform properties, typical parts made are superalloy gas-turbine components with walls as thin as 0.5 mm.

In this process, a mixture of fine sand and urethane is molded over metal dies, and cured with amine vapor. The mold is then held with a robot arm and immersed partially into molten metal held in an induction furnace. The metal may be melted in air (*CLA process*) or in a vacuum (*CLV process*). The vacuum reduces the air pressure inside the mold to about two-thirds of the atmospheric pressure, thus drawing the molten metal into the mold cavities through a gate in the bottom of the mold. The metal in the furnace is

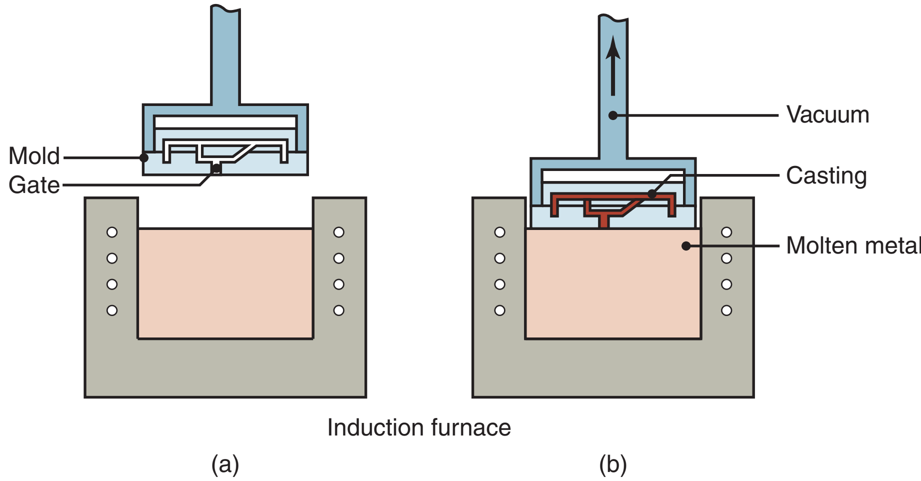


Figure 11.17: Schematic illustration of the vacuum-casting process. Note that the mold has a bottom gate. (a) Before and (b) after immersion of the mold into the molten metal. *Source:* After R. Blackburn.

usually at a temperature of 55°C above the liquidus temperature of the alloy; consequently, it begins to solidify within a very short time.

This process can be automated, with production costs that are similar to those for green-sand casting. Carbon, low- and high-alloy steel, and stainless steel parts, weighing as much as 70 kg, have been vacuum cast by this method. CLA castings are made easily at high volume and relatively low cost; CLV parts usually involve reactive metals, such as aluminum, titanium, zirconium, and hafnium.

11.4.3 Slush Casting

Note in Fig. 10.13 that a solidified skin develops in a casting, which becomes thicker with time. Thin-walled hollow castings can be made by permanent-mold casting using this principle, in a process called *slush casting*. The molten metal is poured into the metal mold; after the desired thickness of solidified skin is obtained, the mold is inverted (or *slung*) and the remaining liquid metal is poured out. The mold halves are then opened and the casting is removed. Note that this operation is similar to making hollow chocolate shapes and other confectionaries. Slush casting is suitable for small production runs, and is generally used for making ornamental and decorative objects, such as lamp bases and stems, and toys from low-melting-point metals, such as zinc, tin, and lead alloys.

11.4.4 Pressure Casting

In the two permanent-mold processes described previously, the molten metal flows into the mold cavity by gravity. In *pressure casting*, also called *pressure pouring* or *low-pressure casting*, the molten metal is forced by gas pressure into a graphite or metal mold. The molten metal is tapped from below the surface, and thus avoiding entrainment of dross and oxides into the mold cavity. The pressure is maintained until the metal has completely solidified in the mold. The molten metal may be forced upward by a vacuum, which also removes dissolved gases and produces a casting with lower porosity. Pressure casting is generally used for high-quality castings, such as steel railroad-car wheels; these wheels also may be cast in sand molds or semipermanent molds made of graphite and sand.

11.4.5 Die Casting

The *die-casting* process, developed in the early 1900s, is a further example of permanent-mold casting. The European term for this process is *pressure die casting*, and should not be confused with pressure casting, described in Section 11.4.4. Typical parts made by die casting are housings for transmissions, business-machine and appliance components, hand-tool components, and toys. The weight of most castings typically ranges from less than 90 g to about 25 kg. Equipment costs, particularly the cost of dies, are somewhat high, but labor costs are generally low when the process is semi- or fully automated. The capabilities of die casting are given in Table 11.3.

In the die-casting process, molten metal is forced into the die cavity at pressures ranging from 0.7 to 700 MPa. There are two basic types of die-casting machines: hot- and cold-chamber.

The **hot-chamber process** (Fig. 11.18) involves the use of a piston, which forces a specific volume of metal into the die cavity through a gooseneck and nozzle; pressures range up to 35 MPa, with an average of about 15 MPa. The metal is held under pressure until it solidifies in the die. To improve die life and to aid in rapid metal cooling (thereby reducing cycle time), dies are usually cooled by circulating water or oil through various passageways in the die block. Low-melting-point alloys, such as zinc, magnesium, tin, and lead, are commonly cast using this process. Cycle times usually range from 200 to 300 shots (individual injections) per hour for zinc, although very small components, such as zipper teeth, can be cast at rates of 18,000 shots per hour.

In the **cold-chamber process** (Fig. 11.19), molten metal is poured into the injection cylinder (*shot chamber*). The chamber is not heated, hence the term *cold chamber*. The metal is forced into the die cavity at pressures usually ranging from 20 to 70 MPa, although they may be as high as 150 MPa.

The machines may be horizontal, as shown in the figure, or vertical, in which case the shot chamber is vertical. High-melting-point alloys of aluminum, magnesium, and copper normally are cast using this method, although ferrous and other metals also can be cast. Molten-metal temperatures start at about 600°C for aluminum and some magnesium alloys, and increase considerably for copper-based and iron-based alloys.

Freeze casting. Die casting can be used for *freeze casting* (see Section 10.2.4) to produce porous metals. In this case, the die is maintained at room temperature, thereby freezing a carrier fluid and separating a suspended powder; the cast part is then sintered (see Section 17.4).

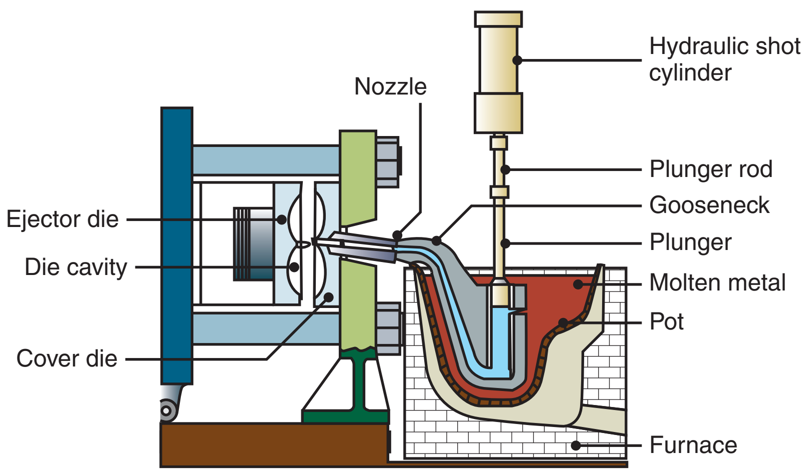


Figure 11.18: Schematic illustration of the hot-chamber die-casting process.

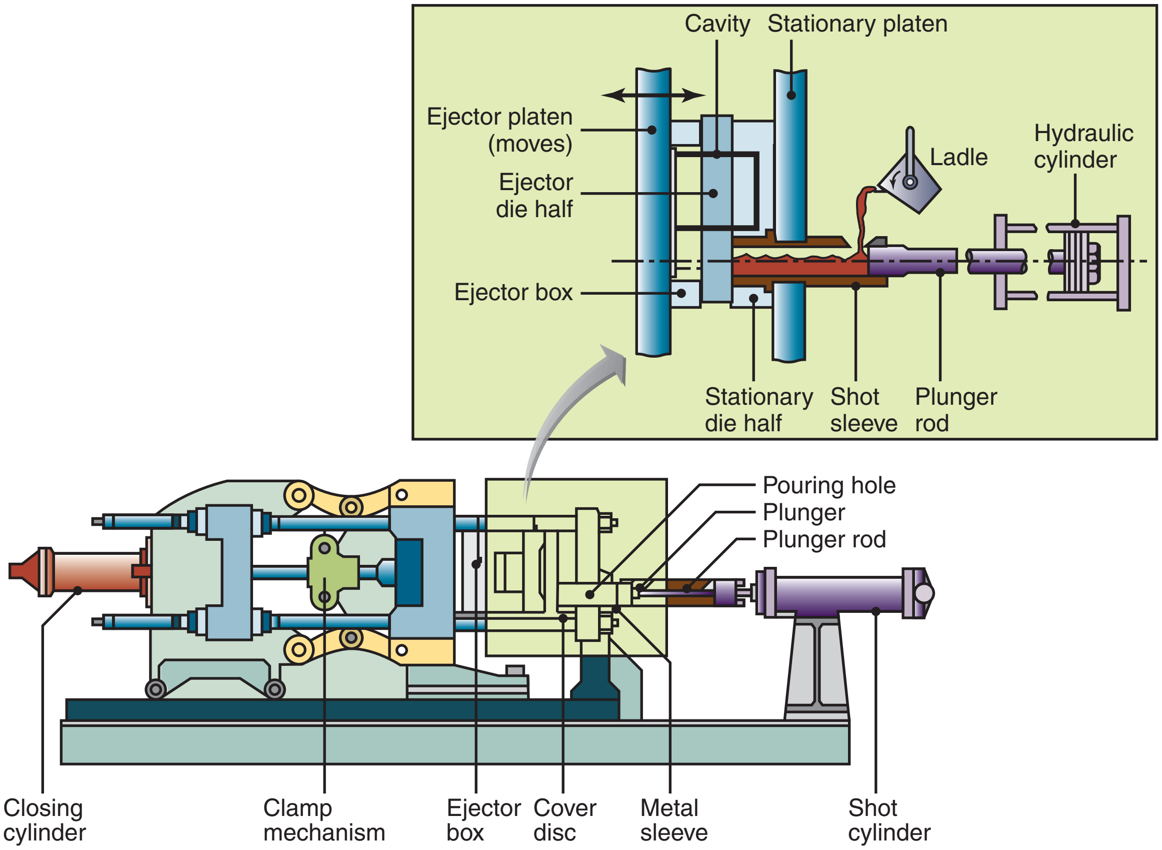


Figure 11.19: Schematic illustration of the cold-chamber die-casting process. These machines are large as compared to the size of the casting, because high forces are required to keep the two halves of the dies closed under pressure.

Process Capabilities and Machine Selection. Die casting has the capability for rapid production of high-quality parts with complex shapes, especially with aluminum, brass, magnesium, and zinc (Table 11.3). It also produces good dimensional accuracy and surface details, so that parts require little or no subsequent finishing operations (net-shape forming; Section 1.6). Because of the high pressures involved, walls as thin as 0.38 mm are produced, which are thinner than those obtained by other casting methods. However, ejector marks remain on part surfaces, as may small amounts of flash (thin material squeezed out between the dies at the die parting line).

Cycle time greatly depends on the ability of a die to extract heat from the molten metal. It is a common practice to incorporate cooling channels in the die, and to pump coolant through the cooling channels; the forced heat transfer keeps the die cool and allows continuous operation. **Conformal** or **contoured cooling** can be performed with dies produced in additive manufacturing; in this case, the cooling channels closely follow the contour of the mold to most efficiently extract heat from the desired location. Conformal cooling is described in greater detail in Section 20.10.

A typical part made by die casting is the aluminum impeller shown in Fig. 11.1d; note the intricate shape and fine surface detail. For certain parts, die casting can compete favorably with other manufacturing methods (such as sheet-metal stamping and forging) or other casting processes. In addition, because the molten metal chills rapidly at the die walls, the casting has a fine-grained, hard skin with high strength; consequently, the strength-to-weight ratio of die-cast parts increases with decreasing wall thickness. With good

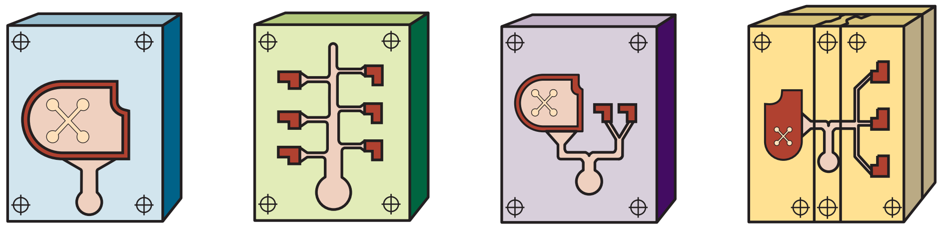


Figure 11.20: Various types of cavities in a die-casting die. *Source:* Courtesy of American Die Casting Institute.

surface finish and dimensional accuracy, die casting can produce smooth surfaces, such as for bearings, that otherwise would normally have to be machined.

Multimaterial components, such as pins, shafts, and certain threaded fasteners, can be die cast integrally. Called **insert casting**, this process is similar to placing wooden sticks in popsicles prior to freezing (see also Section 19.3). For high interfacial strength, insert surfaces may be knurled (see Fig. 23.11), grooved, or splined. Steel, brass, and bronze inserts are commonly used with die-casting alloys. In selecting insert materials, the possibility of *galvanic corrosion* should be taken into account. To avoid this potential problem, the insert can be insulated, plated, or surface treated.

Because of the high pressures involved, dies for die casting have a tendency to separate unless they are clamped together tightly (see Fig. 11.19). Die-casting machines are hence rated according to the clamping force that can be exerted to keep the dies closed during casting. The capacities of commercially available machines range from about 22.5 to 2700 metric tons. Other factors involved in the selection of die-casting machines are die size, piston stroke, shot pressure, various features, and cost.

Die-casting dies (Fig. 11.20) may be *single cavity, multiple cavity* (several identical cavities), *combination cavity* (several different cavities), or *unit dies* (simple, small dies that can be combined in two or more units in a master holding die). Typically, the ratio of die weight to part weight is 1000 to 1. Thus, for example, the die for a casting weighing 2 kg would weigh about 2000 kg. The dies are usually made of hot-work die steels or mold steels (see Section 5.7). **Heat checking** of dies (surface cracking from cyclic heating and cooling of the die, described in Section 3.6) can be a problem. When the die materials are selected and maintained properly, however, dies can last more than a half million shots before any significant die wear takes place.

Case Study 11.3 Die Casting of a Headlight Mount

Figure 11.22 shows a die-cast aluminum component of a daytime running lamp and turn signal for an automobile. Aluminum was preferable to plastic because of its higher heat-sink characteristics and rigidity, and also because tight tolerances were required for mounting and providing wiring access to LED bulbs. The fin size, thickness, and spacing were determined from a heat transfer analysis. The fins were tapered to allow for easy removal from a die, and the corner radii were designed to prevent distortion during ejection. The part was then oriented so that mounting holes and pockets were coplanar to the die parting line to simplify die fabrication. Heating channels were incorporated into the die near the thin sections to slow cooling, while cooling channels were incorporated near the thick sections. The resulting thermal balance led to lower distortion in the final product. The final product was cast from 380 aluminum; it measures 100 $\times$ 75 $\times$ 100 mm for the turn signal and 250 $\times$ 100 $\times$ 50 mm for the daytime running light sub-assembly.

Case Study 11.4 Die Cast Magnesium Liftgate

Figure 11.21 shows a complex high-pressure die-casting produced from a 2017 Chrysler Pacifica, and represents the first high-volume magnesium application of its kind in the automotive industry. The casting forms part of a four-piece assembly, with aluminum sheet upper and lower outer panels and a wiper reinforcement. The AM60B magnesium alloy part reduced the weight of the liftgate by more than 10 kg, representing a 50% reduction over the previous generation design. However, the liftgate affected the designs of the motor, strut, hinge, and other mechanical components, so that the weight savings was actually much greater.

Lightweighting strategies such as design optimization and exploitation of materials with high strength-to-weight ratios are common in automotive and aerospace applications so as to achieve ever increasing fuel economy goals. The original design was a weldment of seven steel stampings; the redesigned liftgate was combined into a single magnesium casting, with 10 spot welds and rivets compared to 84 in the original design. The thin-walled casting takes special care to extract the heat from the magnesium slowly, in order to prevent solidification in the mold and resulting underfills.

11.4.6 Centrifugal Casting

As its name implies, the *centrifugal-casting* process utilizes *inertia* (caused by rotation) to force the molten metal into the mold cavities, a method that was first suggested in the early 1800s. The capabilities of centrifugal casting are given in Table 11.3. There are three types of centrifugal casting: true centrifugal casting, semicentrifugal casting, and centrifuging.

Figure 11.21: A magnesium liftgate for a 2017 Chrysler Pacifica, saving more than 10 kg in weight over the previous generation steel stamping, and with reduction of welds from 84 to 10. Magnesium is increasingly used for vehicle lightweighting, and high-pressure die castings such as this one can be produced with thin walls and intricate part details. *Source:* Courtesy of American Foundry Society.

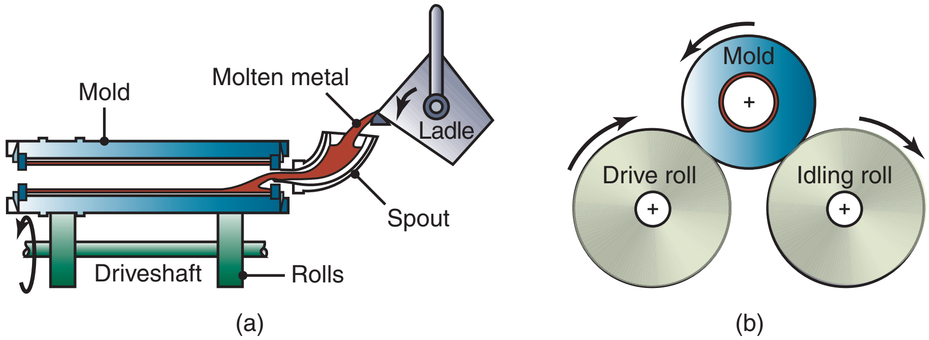


Figure 11.22: (a) Schematic illustration of the centrifugal-casting process. Pipes, cylinder liners, and similarly shaped parts can be cast with this process. (b) Side view of the machine.

True Centrifugal Casting. In *true centrifugal casting*, hollow cylindrical parts (such as pipes, gun barrels, bushings, engine-cylinder liners, bearing rings with or without flanges, and street lampposts) are produced by the technique shown in Fig. 11.22. In this process, molten metal is poured into a rotating mold; the axis of rotation is usually horizontal, but can be vertical for short workpieces. Molds are made of steel, iron, or graphite, and may be coated with a refractory lining to increase mold life. The mold surfaces can be shaped so that pipes with various external designs can be cast. The inner surface of the casting remains cylindrical, because the molten metal is distributed uniformly by the centrifugal forces. However, because of density differences, lighter elements (such as dross, impurities, and pieces of the refractory lining in the mold) tend to collect on the inner surface of the casting. Consequently, the properties of the casting can vary throughout its thickness.

Cylindrical parts ranging from 13 mm to 3 m in diameter and 16 m long can be cast centrifugally, with wall thicknesses ranging from 6 to 125 mm. The pressure generated by the centrifugal force is high (the angular acceleration can be as much as 150 times gravity); such high pressure is necessary for casting thick-walled parts. Castings with good quality, dimensional accuracy, and external surface detail are produced by this process.

Semicentrifugal Casting. An example of semicentrifugal casting is shown in Fig. 11.23. This method is used to cast parts with rotational symmetry, such as a wheel with spokes.

Centrifuging. In *centrifuging*, also called *centrifuge casting*, mold cavities are placed at a certain distance from the axis of rotation. The molten metal is poured from the center, and is forced into the mold by centrifugal forces (Fig. 11.23). The properties of the castings can vary by distance from the axis of rotation, as in true centrifugal casting.

11.4.7 Squeeze Casting and Semisolid-metal Forming

Two casting processes that incorporate the features of both casting and forging (Chapter 14) are squeeze casting and semisolid-metal forming.

Squeeze Casting. The *squeeze-casting* or *liquid-metal forging* process was invented in the 1930s, but developed for industrial applications in the 1960s, and involves the solidification of molten metal under high pressure (Fig. 11.24). Typical products made are automotive components and mortar bodies (a cannon with a short body). The machinery includes a die, punch, and ejector pin. The pressure applied by the punch keeps the entrapped gases in solution, while the contact under high pressure at the die–metal interface

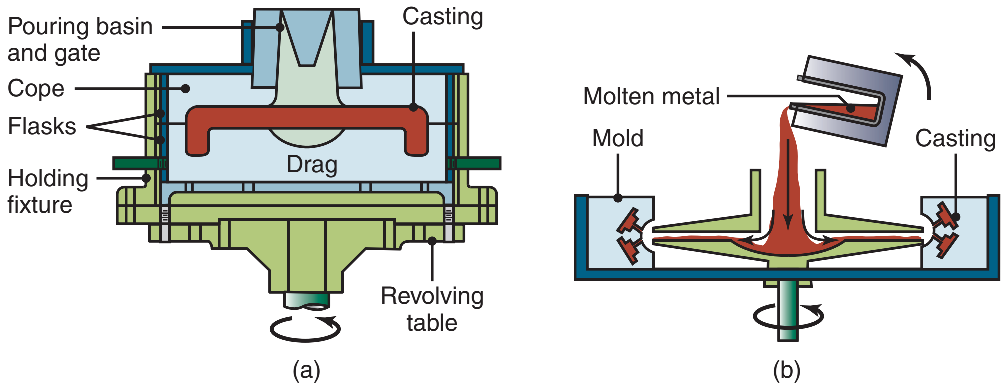


Figure 11.23: (a) Schematic illustration of the semicentrifugal casting process. Wheels with spokes can be cast by this process. (b) Schematic illustration of casting by centrifuging. The molds are placed at the periphery of the machine, and the molten metal is forced into the molds by centrifugal force.

promotes rapid heat transfer, thus resulting in a casting with a fine microstructure and good mechanical properties.

The application of pressure also overcomes hot-metal feeding difficulties that may arise when casting metals with a long freezing range (Section 10.2.2). Complex parts can be made to near-net shape, with fine surface detail from both nonferrous and ferrous alloys.

Semisolid-metal Forming. *Semisolid-metal forming*, also called *mushy-state processing* (see Fig. 10.4) was developed in the 1970s. When it enters the die, the metal (consisting of liquid and solid components) is stirred so that all of the dendrites are broken into fine solids. When cooled in the die, a fine-grained structure is developed. The alloy exhibits *thixotropic behavior*, described in Section 10.2.3, hence the process is also called **thixoforming** or **thixomolding**, meaning its viscosity decreases when agitated. Thus, at rest and above its solidus temperature, the molten alloy has the consistency of butter at room temperature, but when agitated vigorously, its consistency becomes more like that of motor oil.

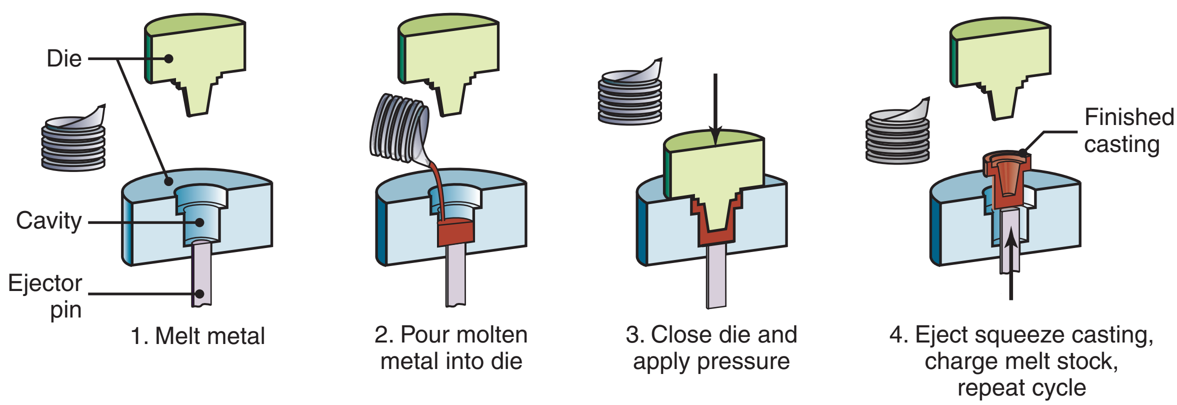


Figure 11.24: Sequence of operations in the squeeze-casting process. This process combines the advantages of casting and forging.

Processing metals in their mushy state also has led to developments in *mushy-state extrusion*, similar to injection molding (described in Section 19.3), *forging*, and *rolling* processes, hence the term *semisolid metalworking*. These processes are also used in making parts with specially designed casting alloys, wrought alloys, and metal-matrix composites (Section 9.5). They also have the capability for blending granules of different alloys, called *thixoblending*, for specific applications.

Thixotropic behavior has also been utilized in developing technologies that combine casting and forging, using cast billets that are then forged when the metal is 30–40% liquid. Parts made include automotive control arms, brackets, and steering components. Processing steels by thixoforming has not yet reached the same stage as with aluminum and magnesium, largely because of the high temperatures involved (which adversely affect die life) and the difficulty in making complex shapes.

The advantages of semisolid metal forming over die casting are: (a) the structures developed are homogeneous, with uniform properties, lower porosity, and high strength; (b) both thin and thick parts can be made; (c) casting alloys as well as wrought alloys can be used; (d) parts can subsequently be heat treated; and (e) the lower superheat results in shorter cycle times. However, material and overall costs are higher than those for die casting.

Rheocasting. This technique, first investigated in the 1960s, is used for forming metals in their semisolid state. The metal is heated to just above its solidus temperature, and poured into a vessel to cool it down to the semisolid state. The slurry is then mixed and delivered to the mold or die. This process is being used successfully with aluminum and magnesium alloys.

11.4.8 Composite-mold Casting Operations

Composite molds are made of two or more different materials and are used in shell molding and various other casting processes. They are generally employed in casting complex shapes, such as impellers for turbines. Composite molds increase the strength of the mold, improve the dimensional accuracy and surface finish of the castings, and can help reduce overall costs and processing time. Molding materials commonly used are shells (made as described in Section 11.2.2), plaster, sand with binder, metal, and graphite. These molds may include cores and chills to control the rate of solidification in critical areas of castings.

11.5 Casting Techniques for Single-crystal Components

This section describes the techniques used to cast single-crystal components, such as gas turbine blades which generally are made of nickel-based superalloys, and used in the hot stages of the engine.

Conventional Casting of Turbine Blades. In the *conventional-casting process*, the molten metal is poured into a ceramic mold, and begins to solidify at the mold walls. The grain structure developed is polycrystalline, similar to that shown in Fig. 10.2c. However, the presence of grain boundaries makes this structure susceptible to creep and cracking along the boundaries under the centrifugal forces and elevated temperatures commonly encountered in an operating gas turbine.

Directionally Solidified Blades. The *directional-solidification process* (Fig. 11.25) was first developed in 1960. The ceramic mold, supported by a water-cooled chill plate, is preheated by radiant heating. After the metal is poured into the mold, the chill-plate assembly is lowered slowly. Crystals begin to grow at the chill-plate surface and on upward, like the *columnar grains* shown in Fig. 10.3. The blade is solidified directionally, with longitudinal but no transverse grain boundaries. The blade is thus stronger in the direction of centrifugal forces developed in the gas turbine.

Single-crystal Blades. In *crystal growing*, developed in the late 1960s, the mold has a constriction in the shape of a corkscrew or helix (Figs. 11.25b and c); its cross section is so small that it allows only one crystal

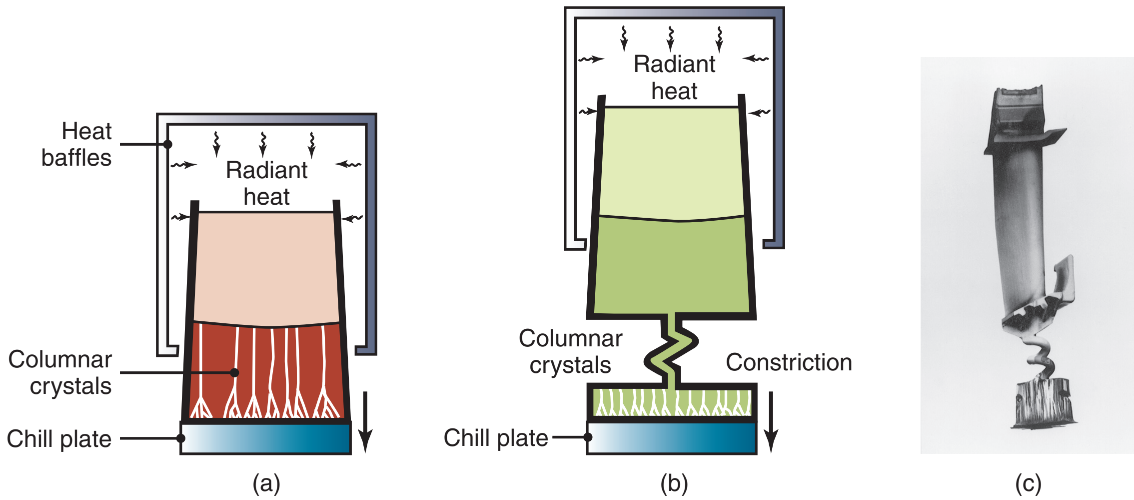


Figure 11.25: Methods of casting turbine blades: (a) directional solidification; (b) method to produce a single-crystal blade; (c) a single-crystal blade with the constriction portion still attached (see also Fig. 1.1). *Source:* (a) and (b) After B.H. Kear, (c) Courtesy of ASM International.

to fit through. The mechanism of crystal growth is such that only the most favorably oriented crystals are able to grow through the helix (a situation similar to that shown in Fig. 10.3), because all other crystals are intercepted by the walls of the helical passage.

As the assembly is slowly lowered, a single crystal grows upward through the constriction and begins to grow in the mold; strict control of the rate of movement is essential. Although single-crystal blades are more expensive than other types, the lack of grain boundaries makes them resistant to creep and thermal shock, hence they have a longer and more reliable service life.

Single-crystal Growing. Single-crystal growing is a major activity in the semiconductor industry in the manufacture of the silicon wafers for microelectronic devices (Chapter 28). There are two basic methods of crystal growing:

- In the **crystal-pulling method**, also known as the **Czochralski** (CZ) **process** (Fig. 11.26), a seed crystal is dipped into the molten metal and then pulled out slowly, at a rate of about 10 μm/s, while being rotated. The liquid metal begins to solidify on the seed, and the crystal structure of the seed continues throughout. *Dopants* (Section 28.3) may be added to the liquid metal to impart specific electrical properties. Single crystals of silicon, germanium, and various other elements are grown using this process. Single-crystal ingots up to 400 mm in diameter and over 2 m in length have been produced by this technique, although 200- and 300-mm ingots are common in the production of silicon wafers for integrated circuit manufacture (Part V).

- The **floating-zone method** (Fig. 11.26b) starts with a rod of polycrystalline silicon resting on a single crystal; an induction coil then heats these two pieces while the coil moves slowly upward. The single crystal grows upward, while maintaining its orientation. Thin wafers are then cut from the rod (see Section 28.4), cleaned, and polished for use in microelectronic device fabrication. This process is suitable for producing diameters under 150 mm, with very low levels of impurities.

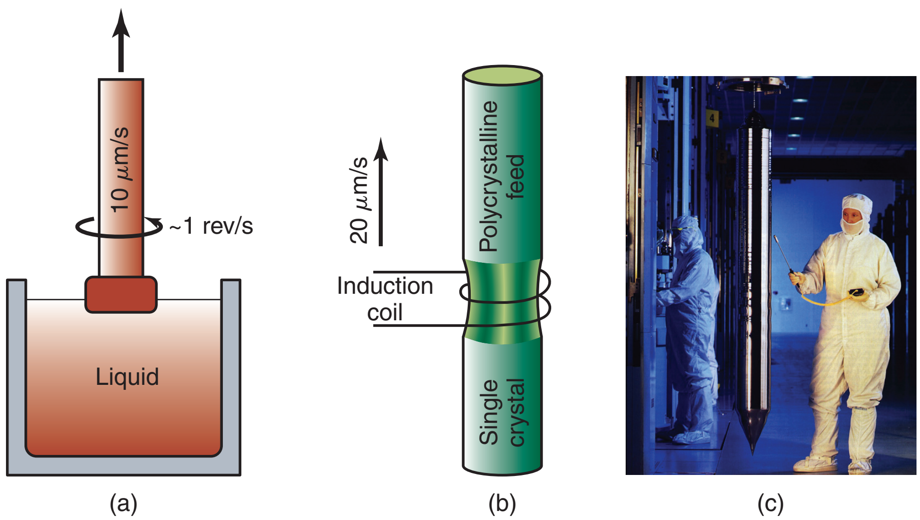


Figure 11.26: Two methods of crystal growing: (a) crystal pulling (Czochralski process) and (b) the floating-zone method. (c) A single-crystal ingot produced by the Czochralski process. *Source:* Courtesy of Intel Corp.

11.6 Rapid Solidification

The properties of *amorphous alloys*, also known as *metallic glasses*, are described in Section 6.15. The technique for making these alloys, called *rapid solidification*, involves cooling the molten metal at rates as high as 10^6 K/s so that it does not have sufficient time to crystallize (see also Fig. 1.11). Rapid solidification results in a significant extension of solid solubility (Section 4.2), grain refinement, and reduced microsegregation (see Section 10.2.3). Metallic glasses have very high strength but limited ductility; this behavior can be thought of as an extension of the Hall–Petch effect (see Section 1.5.1), where the grain size is on the order of one atom.

In a method called **melt spinning** (Fig. 19.6), the alloy is melted by induction in a ceramic crucible. It is then propelled, under high gas pressure, against a rotating copper disk (chill block), which rapidly chills the alloy (**splat cooling**), forming a metallic glass strip. Significant research is taking place to produce bulk forms of metallic glass.

11.7 Inspection of Castings

Several methods can be used to inspect castings to determine their quality and the presence and types of any defects. Castings can be inspected *visually*, or *optically*, for surface defects. Subsurface and internal defects are investigated using various *nondestructive* techniques, described in Section 36.10. In *destructive testing* (Section 36.11), specimens are removed from various locations in a casting, and tested for strength, ductility, and various other mechanical properties, and to determine the presence, location, and distribution of porosity and other defects.

Pressure tightness of cast components, such as valves, pumps, and pipes, is usually determined by sealing the openings in the casting, then pressurizing it with water, oil, or air. For leak tightness requirements in critical applications, pressurized helium or specially scented gases, with detectors (sniffers), are used. The casting is then inspected for leaks while the pressure is maintained; unacceptable or defective castings are remelted for reprocessing.

11.8 Melting Practice and Furnaces

Melting practice is an important aspect of casting operations, because it has a direct bearing on the quality of castings. Furnaces are charged with *melting stock*, consisting of metal, alloying elements, and various other materials, such as **flux** and slag-forming constituents. Fluxes are inorganic compounds that refine the molten metal by removing dissolved gases and various impurities. They may be added manually or can be injected automatically into the molten metal.

Melting Furnaces. The melting furnaces commonly used in foundries are electric-arc furnaces, induction furnaces, crucible furnaces, and cupolas.

1. **Electric arc** furnaces, described in Section 5.2.3 and illustrated in Fig. 5.2, are used extensively in foundries, because of their high rate of melting (thus high-production rate), much less pollution than other types, and their ability to hold the molten metal (keeping it at a constant temperature for a period of time) for alloying purposes.

2. **Induction** furnaces (Fig. 5.2c) are especially useful in smaller foundries, and produce composition-controlled melts. There are two basic types. (a) The *coreless induction furnace* consists of a crucible, surrounded with a water-cooled copper coil through which high-frequency current passes. Because there is a strong electromagnetic stirring action during induction heating, this type of furnace has excellent mixing characteristics and is used for alloying and adding a new charge of metal into the furnace. (b) A core or channel furnace, uses low-frequency current (as low as 60 Hz), and has a coil that surrounds only a small portion of the unit. These furnaces are commonly used in nonferrous foundries, and are particularly suitable for *superheating* (heating above normal casting temperature to improve fluidity), for *holding*, which makes it suitable for die-casting applications, and for *duplexing* (using two furnaces: melting the metal in one furnace and then transferring it to another).

3. **Crucible** furnaces (Fig. 11.27a), which have been used extensively throughout history, are heated using various fuels, such as commercial gases, fuel oil, and fossil fuel, and with electricity. Crucible furnaces may be stationary, tilting, or movable.

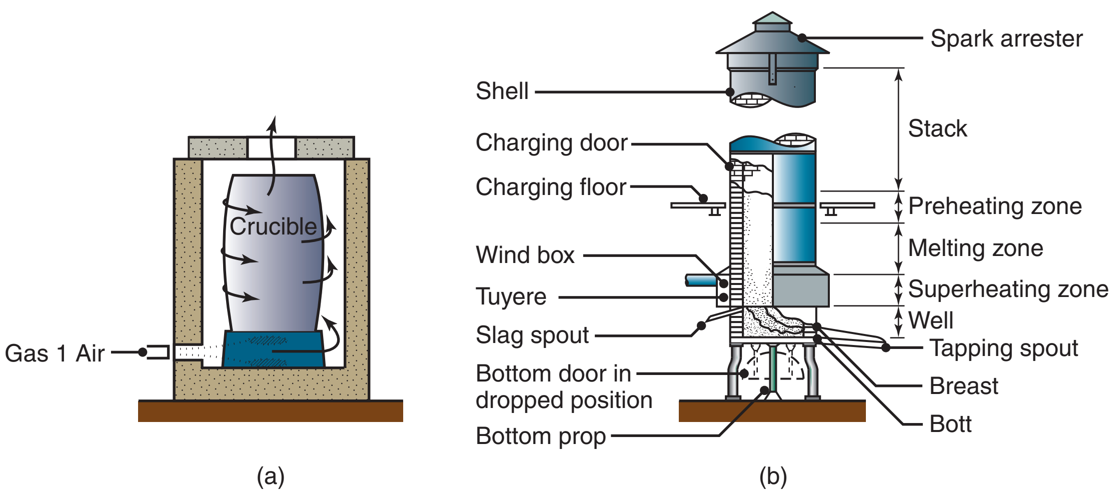


Figure 11.27: Two types of melting furnaces used in foundries: (a) crucible and (b) cupola.

4. **Cupolas** are basically vertical refractory-lined steel vessels, charged with alternating layers of metal, coke, and flux (Fig. 11.27b). Although they require major investments and are increasingly replaced by induction furnaces, cupolas operate continuously, have high melting rates, and produce large amounts of molten metal.

5. **Levitation melting** involves *magnetic suspension* of the molten metal. An induction coil simultaneously heats a solid billet and stirs and confines the melt, thus eliminating the need for a crucible (which could contaminate the molten metal with its oxide inclusions). The molten metal flows downward into an investment-casting mold, placed directly below the coil. Investment castings made by this method are free of refractory inclusions and of gas porosity, and have a uniform fine-grained structure.

11.9 Foundries and Foundry Automation

Casting operations are carried out in **foundries** (from the Latin *fundere*, meaning melting and pouring). Although these operations traditionally have involved much manual labor, modern foundries have efficient automated and computer-integrated facilities for all aspects of their operations.

As outlined in Fig. 11.2, foundry operations initially involve two separate groups of activities. The first group is pattern and mold making using computer-aided design and manufacturing (Chapter 38) and rapid-prototyping techniques (Chapter 20), thus improving efficiency and lowering costs. A variety of automated machinery is used to minimize labor costs, which can be significant in the production of castings. The second group of activities involves melting the metals, controlling their compositions and impurities, and pouring them into molds.

The rest of the operations in a foundry, such as pouring into molds (some carried along conveyors), shakeout, cleaning, heat treatment, and inspection, also are automated. Automation minimizes labor, reduces the possibility of human error, increases the production rate, and attains higher quality levels. Industrial robots (Section 37.6) are used extensively, such as for cleaning, cutting risers, mold venting, mold spraying, pouring, sorting, and inspection. Other operations involve automatic storage and retrieval systems for cores and patterns, using automated guided vehicles (Section 37.5).

Summary

- Expendable-mold, permanent-pattern processes include sand, shell-mold, plaster-mold, and ceramic-mold casting. These processes require the destruction of the mold for each casting produced, but mold production is facilitated by a reusable pattern.

- Expendable-mold, expendable-pattern processes include lost-foam and investment casting. In these processes, a pattern is consumed for each mold produced, and the mold is destroyed after each casting.

- Permanent-mold processes have molds or dies that can be used to produce castings at high production rates. Common permanent-mold processes include slush casting, pressure casting, die casting, and centrifugal casting.

- The molds used in permanent-mold casting are made of metal or graphite, and are used repeatedly to produce a large number of parts. Because metals are good heat conductors but do not allow gases to escape, permanent molds have fundamentally different effects on castings than sand or other aggregate mold materials.

- In permanent-mold casting, die and equipment costs are relatively high, but the processes are economical for large production runs. Scrap loss is low, dimensional accuracy is relatively high, and good surface details can be achieved.

- Casting processes include squeeze casting (a combination of casting and forging), semisolid-metal forming, rapid solidification (for the production of amorphous alloys), and the casting of single-crystal components (such as turbine blades and silicon ingots for making wafers in integrated-circuit manufacture).
- Melting processes and their control are important factors in casting operations. They include proper melting of the metals, preparation for alloying and removal of slag and dross, and pouring the molten metal into the molds. Inspection of castings for possible internal or external defects also is essential.
- Castings are generally subjected to subsequent processing, such as heat treatment and machining operations, to produce the final desired shapes, surface characteristics, and the required surface finish and dimensional accuracy.

Key Terms

Binders
Centrifugal casting
Ceramic-mold casting
Chaplets
Composite mold
Core print
Cores
Crystal growing
Die casting
Evaporative-pattern casting
Expendable mold
Expendable-pattern casting
Flux
Foundry
Green molding sand
Insert casting
Investment casting
Levitation melting
Lost-foam process
Lost-wax process
Parting agent
Patterns
Permanent mold
Permanent-mold casting
Plaster-mold casting
Precision casting
Pressure casting
Rammed-graphite molding
Rapid solidification
Rheocasting
Sand casting
Semisolid-metal forming
Shell-mold casting
Slush casting
Squeeze casting
Thixotropic
Vacuum casting

Bibliography

Analysis of Casting Defects, American Foundrymen's Society, 2002.
ASM Handbook, Vol. 1A: **Cast Iron Science and Technology**, ASM International, 2017.
ASM Handbook, Vol. 15: **Casting**, ASM International, 2008.
Beeley, P.R., **Foundry Technology**, Butterworth-Heinemann, 2002.
Beeley, P.R., and Smart, R.F. (eds.), **Investment Casting**, Maney Publishing, 2008.
Campbell, J., **Complete Casting Handbook: Metal Casting Processes, Techniques and Design**, 2nd ed., Butterworth-Heinemann, 2015.
Hirt, G., and Kopp, R., **Thixoforming**, Wiley-VCH, 2009.

Investment Casting Handbook, Investment Casting Institute, 1997.
Kaufman, J.G., and Rooy, E.L., **Aluminum Alloy Castings: Properties, Processes and Applications**, ASM International, 2004.
Kirkwood, D.H., Suery, M., Kapranos, P., and Atkinson, H.V., **Semi-solid Processing of Alloys**, Springer, 2009.
Nafisi, S., and Ghomashchi, R., **Semi-solid Processing of Aluminum Alloys**, Springer, 2016.
Sias, F.R., **Lost-wax Casting**, Woodsmere Press, 2006.
Steel Castings Handbook, 6th ed., ASM International, 1995.
Vinarcik, E.J., **High Integrity Die Casting**, Wiley, 2002.
Young, K.P., **Semi-solid Processing**, Chapman & Hall, 1997.

Review Questions

11.1. Describe the differences between expendable and permanent molds.

11.2. Name the important factors in selecting sand for molds.

11.3. What are the major types of sand molds? What are their characteristics?

11.4. List important considerations when selecting pattern materials.

11.5. What is the function of a core?

11.6. What is the difference between sand-mold and shell-mold casting?

11.7. What are composite molds? Why are they used?

11.8. Describe the features of plaster-mold casting.

11.9. Name the type of materials typically used for permanent-mold casting processes.

11.10. What are the advantages of pressure casting over other processes?

11.11. List the advantages and limitations of die casting.

11.12. What is the purpose of a riser? What is a blind riser?

11.13. Explain the purpose of a vent and a runner in a casting mold.

11.14. How are shell molds produced?

11.15. What keeps the mold together in vacuum casting?

11.16. What is squeeze casting? What are its advantages?

11.17. What are the advantages of the lost-foam casting process?

11.18. How are single-crystal turbine blades produced?

Qualitative Problems

11.19. What are the reasons for the large variety of casting processes that have been developed over the years? Explain with specific examples.

11.20. Why are risers not as useful in die casting as they are in sand casting?

11.21. Describe the drawbacks to having a riser that is (a) too large and (b) too small.

11.22. Why can blind risers be smaller than open-top risers?

11.23. Why does die casting produce the smallest cast parts?

11.24. Why is the investment-casting process capable of producing fine surface detail on castings?

11.25. What differences, if any, would you expect in the properties of castings made by permanent-mold versus sand-casting processes?

11.26. Recently, cores for sand casting have been produced from salt. What advantages and disadvantages would you expect from using salt cores?

11.27. Would you recommend preheating the molds used in permanent-mold casting? Would you remove the casting soon after it has solidified? Explain your reasons.

11.28. Give reasons for, and examples of, using die inserts.

11.29. Referring to Fig. 11.3, do you think it is necessary to weigh down or clamp the two halves of the mold? Explain your reasons. Do you think that the kind of metal cast, such as gray cast iron versus aluminum, should make a difference in the clamping force? Explain.

11.30. Explain why squeeze casting produces parts with better mechanical properties, dimensional accuracy, and surface finish than do expendable-mold processes.

11.31. How are the individual wax patterns attached on a "tree" in investment casting?

11.32. Describe the measures that you would take to reduce core shifting in sand casting.

11.33. You have seen that, even though die casting produces thin parts, there is a limit to how thin they can be. Why can't even thinner parts be made by this process?

11.34. How are hollow parts with various cavities made by die casting? Are cores used? If so, how? Explain.

11.35. It was stated that the strength-to-weight ratio of die-cast parts increases with decreasing wall thickness. Explain why.

11.36. How are risers and sprues placed in sand molds? Explain, with appropriate sketches.

11.37. In shell-mold casting, the curing process is critical to the quality of the finished mold. In this stage of the process, the shell-mold assembly and cores are placed in an oven for a short period of time to complete the curing of the resin binder. List probable causes of unevenly cured cores or of uneven core thicknesses.

11.38. Why does the die-casting machine shown in Fig. 11.19 have such a large mechanism to close the dies? Explain.

11.39. Chocolate forms are available in hollow shapes. What process should be used to make these chocolates?

11.40. What are the benefits to heating the mold in investment casting before pouring in the molten metal? Are there any drawbacks? Explain.

11.41. The "slushy" state of alloys refers to that state between the solidus and liquidus temperatures, as described in Section 10.2.2. Pure metals do not have such a slushy state. Does this mean that pure metals cannot be slush cast? Explain.

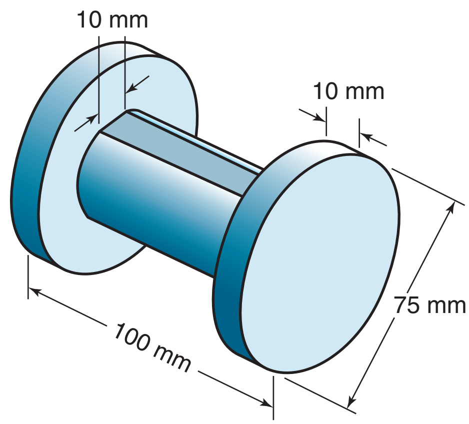
11.42. Can a chaplet also act as a chill? Explain.

11.43. Rank the casting processes described in this chapter in terms of their solidification rate. (That is, which processes extract heat the fastest from a given volume of metal?)

Quantitative Problems

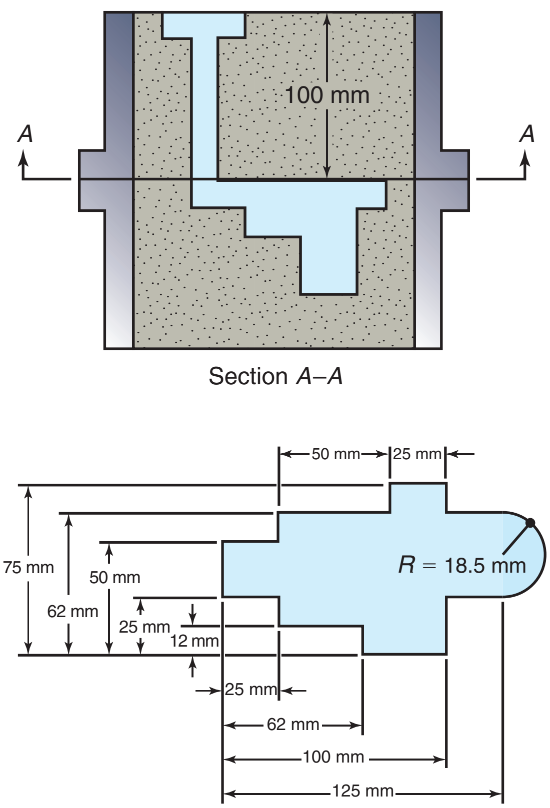
11.44. Estimate the clamping force for a die-casting machine in which the casting is rectangular with projected dimensions of 125 mm $\times$ 175 mm. Would your answer depend on whether it is a hot-chamber or cold-chamber process? Explain.

11.45. The blank for the spool shown below is to be sand cast out of A-319, an aluminum casting alloy. Make a sketch of the wooden pattern for this part, and include all necessary allowances for shrinkage and machining.



11.46. Repeat Problem 11.45, but assume that the aluminum spool is to be cast by expendable-pattern casting. Explain the important differences between the two patterns.

11.47. In sand casting, it is important that the cope-mold half be weighted down with sufficient force to keep it from floating when the molten metal is poured in. For the casting shown below, calculate the minimum amount of weight necessary to keep the cope from floating up as the molten metal is poured in. (*Hint:* The buoyancy force exerted by the molten metal on the cope is dependent on the effective height of the metal head above the cope.)



Material: Low-carbon steel
Density: 7196 kg/m^3
All dimensions in inches

11.48. If an acceleration of 125 g is necessary to produce a part in true centrifugal casting and the part has an inner diameter of 300 mm, a mean outer diameter of 400 mm, and a length of 8 m, what rotational speed is needed?

11.49. A jeweler wishes to produce 20 gold rings in one investment-casting operation. The wax parts are attached to a wax central sprue with a 20-mm diameter. The rings are located in four rows, each 15 mm from the other on the sprue. The rings require a 3-mm diameter, 12-mm long runner to the sprue. Estimate the weight of gold needed to completely fill the rings, runners, and sprues. Assume a typical ring has a 25-mm outer diameter, 19-mm inner diameter, and 5-mm width. The specific gravity of gold is 19.3.

11.50. Assume that you are an instructor covering the topics described in this chapter, and you are giving a quiz on the numerical aspects of casting processes to test the understanding of the students. Prepare two quantitative problems and supply the answers.

Synthesis, Design, and Projects

11.51. Describe the procedures that would be involved in making a large outdoor bronze statue. Which casting process(es) would be suitable? Why?

11.52. The optimum shape of a riser is spherical to ensure that it cools more slowly than the casting it feeds. However, spherically shaped risers are difficult to cast. (a) Sketch the shape of a blind riser that is easy to mold, but also has the smallest possible surface-area-to-volume ratio. (b) Compare the solidification time of the riser in part (a) with that of a riser shaped like a right circular cylinder. Assume that the volume of each riser is the same and the height of each is equal to the diameter. (See Example 10.1.)

11.53. Sketch and describe a casting line consisting of machinery, conveyors, robots, sensors, etc., that automatically could perform the expendable-pattern casting process.

11.54. Outline the casting processes that would be most suitable for making small toys. Explain your choices.

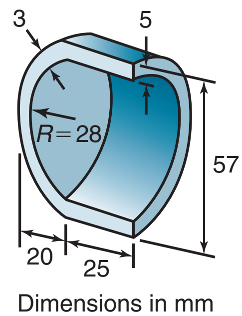
11.55. Make a list of the mold and die materials used in the casting processes described in this chapter. Under each type of material, list the casting processes that are employed and explain why these processes are suitable for that particular mold or die material.

11.56. Write a brief report on the permeability of molds and the techniques that are used to determine permeability.

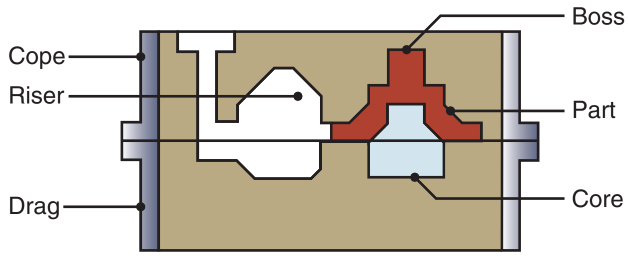
11.57. Light metals commonly are cast in vulcanized rubber molds. Conduct a literature search and describe the mechanics of this process.

11.58. It sometimes is desirable to cool metals more slowly than they would be if the molds were maintained at room temperature. List and explain the methods you would use to slow down the cooling process.

11.59. The part shown below is a hemispherical shell used as an acetabular (mushroom-shaped) cup in a total hip replacement. Select a casting process for making this part, and provide a sketch of all the patterns or tooling needed if it is to be produced from a cobalt–chrome alloy.



11.60. Porosity that has developed in the boss of a casting is illustrated below. Show that the porosity can be eliminated simply by repositioning the parting line of this casting.



11.61. Review Fig. II.1b, and note that the gemstones have been cast in place. Design a ring with a means of securing a gemstone in the wax pattern, such that it will remain in the mold as the wax is being melted. Could such an approach be used in lost foam casting?

Chapter 12

Metal Casting: Design, Materials, and Economics

- This final chapter on metal casting serves as a general guide to the interrelationships among product design, material and process selection, and economic considerations in casting.
- The chapter describes in detail the design considerations for casting operations, and discusses the general guidelines for successful casting practices.
- The characteristics and applications of the most common ferrous and nonferrous alloys are then described.
- The chapter ends with a brief review of casting economics.

12.1 Introduction

In the preceding two chapters, it was noted that successful casting practice requires the proper control of a large number of variables. These variables pertain to the particular characteristics of the metals and alloys cast, method of casting, mold and die materials, mold design, and processing parameters. Factors such as the flow of the molten metal in the mold cavities, the gating systems, the rate of cooling, and the gases evolved all influence the quality of a casting.

This chapter describes general design considerations for metal casting, and presents guidelines for avoiding defects. It then describes the characteristics of the metals and alloys that are commonly cast, together with their typical applications. Because the economics of casting operations are just as important as their technical aspects, the chapter also outlines the basic economic factors relevant to all casting operations.

12.2 Design Considerations in Casting

As in all manufacturing operations, certain **design principles** pertaining to casting have been developed over the years. Although these principles have been established primarily through experience, analytical methods, process simulation and modeling, and computer-aided design and manufacturing techniques have now all come into wide use as well, improving the quality of castings and productivity, resulting in significant cost savings.

All casting processes share several basic characteristics; consequently, a number of design considerations apply equally to, for example, sand casting and die casting. However, each process still has its own particular design considerations. Sand casting will require consideration of mold erosion and associated sand inclusions in the casting process, whereas die casting will not have this concern, but it has others, such as heat checking of dies, which significantly reduces die life.

Troubleshooting the causes of defects in cast products can be complicated; the considerations presented in this chapter are to serve only as guidelines. Moreover, defects frequently are random and can be difficult to reproduce, thus complicating the implementation of corrective measures. In most situations, a given mold design will produce mostly good parts as well as some defective parts. For these reasons, strict quality control procedures have to be implemented, especially for critical applications.

It should also be noted that many of the design rules developed over the years are now put in doubt or somewhat relaxed through the application of additive manufacturing (Chapter 20). For example, a sand mold produced through additive manufacturing need not be designed with certain accommodations, such as draft and corner radii, to allow for pattern removal. Additive manufacturing has its own design concerns, as described in Section 20.12, as well as its economic usefulness for short production runs.

12.2.1 General Design Considerations for Castings

There are two types of design issues in casting: (a) geometric features, tolerances, etc., that should be incorporated into the part and (b) mold features that are needed to produce the desired casting. Robust design of castings usually involves the following steps:

1. Design the part so that the shape is cast as easily as possible. If secondary operations, such as machining, are required, include a machining allowance (that is, a slight oversize of the part), surfaces for fixturing, and reinforcement where necessary (to support the machining forces). Several design considerations are given in this chapter to assist in such efforts.

2. Select a casting process and a material suitable for the part, its size, the required production quantity, and mechanical properties. Often, the shape, the material, and the process(s) need to be specified simultaneously, which can be a demanding design challenge.

3. Locate the parting line of the mold in the part.

4. Design and locate the gates to allow uniform feeding of the mold cavity with molten metal.

5. Select an appropriate runner geometry for the system.

6. Locate mold features, such as sprues, screens, and risers, as appropriate.

7. Check that proper controls and good practices are in place.

Design of Parts to Be Cast. The following considerations are important in designing castings, as outlined in Fig. 12.1:

1. **Corners, angles, and section thickness.** Sharp corners, angles, and fillets should be avoided as much as possible, because they act as stress raisers and may cause cracking and tearing of the metal (as

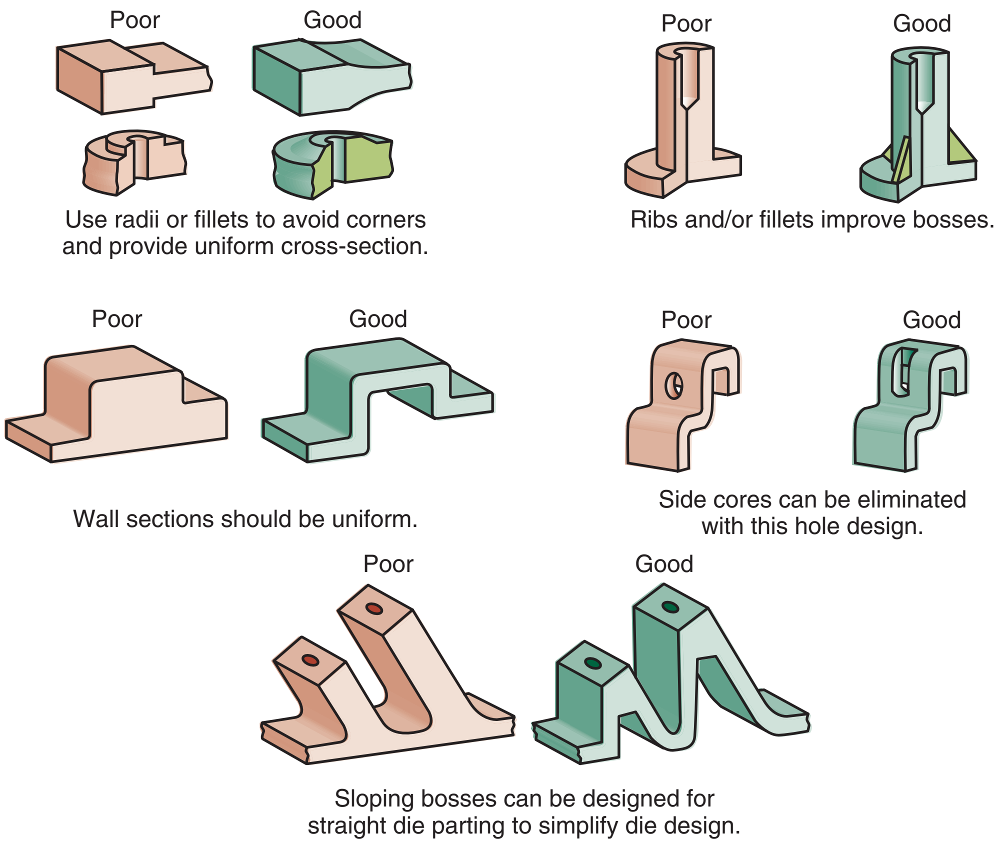


Figure 12.1: Suggested design modifications to avoid defects in castings. *Source:* Courtesy of the American Die Casting Institute.

well as of the dies) during solidification. Fillet radii should be selected so as to minimize stress concentrations and to ensure proper molten-metal flow during pouring. Fillet radii usually range from 3 to 25 mm, although smaller radii may be permissible in small castings and for specific applications. On the other hand, if the fillet radii are too large, the volume of the material in those regions also is large, and hence the cooling rate is lower.

Section changes should be blended smoothly into each other. The location of the largest circle that can be inscribed in a particular region (Figs. 12.2a and b) is critical so far as shrinkage cavities are concerned. Because the cooling rate in regions with larger circles is lower, these regions are called **hot spots**, and can cause **shrinkage cavities** and **porosity** (Figs. 12.2c and d). Cavities at hot spots can be eliminated by using small cores; although they produce cored holes in the casting (Fig. 12.2e), these holes do not significantly affect its strength. It is also important to maintain uniform cross sections and wall thicknesses throughout the casting, in order to avoid or minimize shrinkage cavities. Although they increase the production cost, *metal paddings* or *chills* in the mold can eliminate or minimize hot spots (see Fig. 10.14).

2. **Flat areas.** Large flat areas (plane surfaces) should be avoided, since (a) they may warp during cooling because of temperature gradients or (b) result in poor surface finish because of uneven flow of the metal during pouring. One of the common techniques for avoiding these problems is to break up flat surfaces with staggered ribs and serrations, as described below.

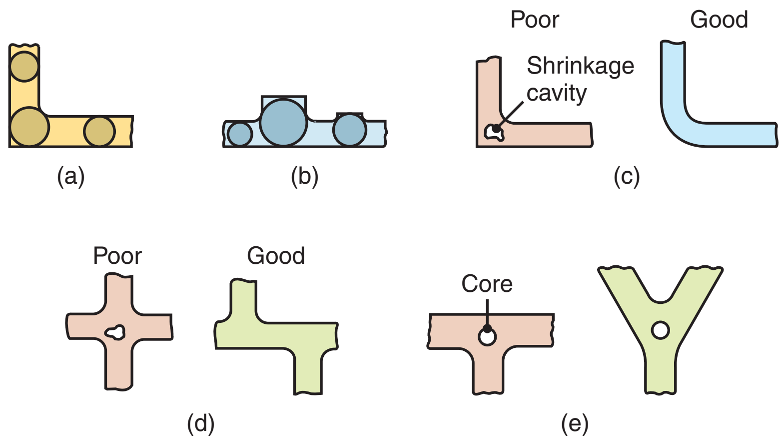


Figure 12.2: Examples of designs showing the importance of maintaining uniform cross sections in castings to avoid hot spots and shrinkage cavities. (a) Illustration of the method of inscribing the largest possible circle in a cross section; locations where abrupt changes in circle size occurs are concerns for hot spots and shrinkage pores. (b)–(e) Common geometries and strategies for reducing or eliminating pores.

3. **Ribs.** One method of producing parts with uniform thickness is to eliminate large, bulky volumes in the casting, as shown in Fig. 12.1; however, this can result in a loss in stiffness and, especially with flat regions, can lead to warping. One solution to these difficulties is to use ribs or a support structure on the casting, as shown in Fig. 12.3. They are usually placed on the side that is less visible. Ribs should, in general, have a thickness around 80% of the adjoining member thickness, and should be deeper than their strut thickness. It is beneficial to have the ribs solidify before the members they adjoin. Ribs should not be placed on both sides of a casting, and should not meet at acute angles, because of complications to molding.

4. **Shrinkage.** To avoid cracking of the casting during cooling, allowance should be provided for shrinkage during solidification and/or cooling to room temperature. In castings with intersecting ribs, the tensile stresses developed can be reduced by staggering the ribs or by modifying the intersection

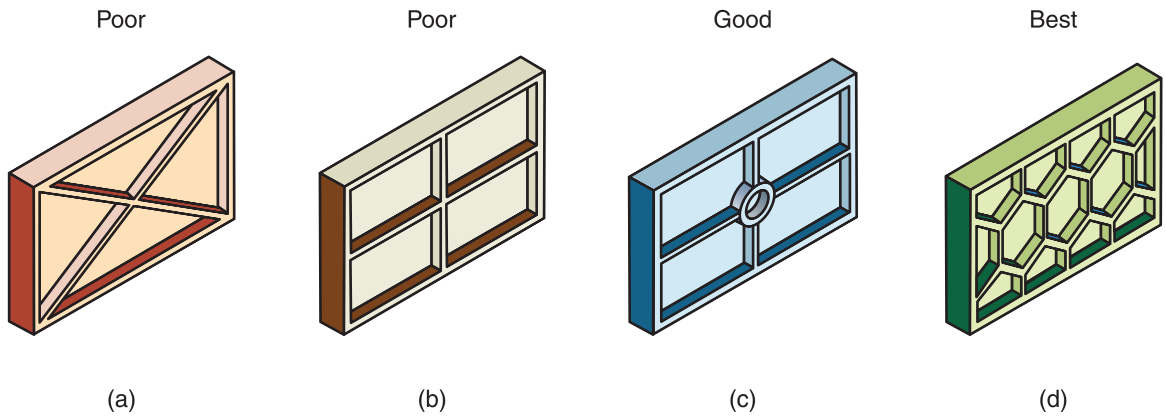


Figure 12.3: Rib designs for use on thin sections or flat surfaces to control or eliminate warping. Note the progression of designs: from left to right, the rib designs have improved castability and reliability.

Table 12.1: Normal Shrinkage Allowance for Some Metals Cast in Sand Molds.

Metal	Shrinkage allowance (%)
Cast Irons	
Gray cast iron	0.83–1.3
White cast iron	2.1
Malleable cast iron	0.78–1.0
Aluminum alloys	1.3
Magnesium alloys	1.3
Copper alloys	
Yellow brass	1.3–1.6
Phosphor bronze	1.0–1.6
Aluminum bronze	2.1
High-manganese steel	2.6

geometry. Pattern dimensions also should allow for shrinkage of the metal during solidification and cooling. Allowances for shrinkage, known as **patternmaker's shrinkage allowances**, usually range from about 10 to 20 mm/m. Table 12.1 gives the normal shrinkage allowance for metals commonly sand cast.

5. **Draft.** A small draft (taper) is typically provided in sand-mold patterns to enable the removal of the pattern without damaging the mold (see Fig. 11.5). Drafts generally range from 5 to 15 mm/m. Depending on the quality of the pattern, draft angles usually range from $0.5°$ to $2°$. The angles on inside surfaces typically are twice this range; they have to be higher than those for outer surfaces because the casting shrinks inward toward the core.

6. **Dimensional tolerances.** Dimensional tolerances depend on the particular casting process employed, size of the casting, and type of pattern used. Tolerances should be as wide as possible, within the limits of good part performance, as otherwise the cost of the casting increases. In commercial practice, tolerances are typically in the range of ± 0.8 mm for small castings, and increase with the size of the castings. For large castings, for instance, they may be as much as ± 6 mm.

7. **Lettering and markings.** It is common practice to include some form of part identification, such as lettering, numbers, or company logos. These features can be sunk into the casting or can protrude from the casting surface; the more desirable one depends on the method of producing the molds. For example, in sand casting, a pattern plate is produced by machining on a computer numerically controlled milling machine (Section 24.2), because it is simpler to machine letters into the pattern plate, they well be recessed in the part. In die casting, it is simpler to machine letters into the mold, leading to letters that protrude.

8. **Finishing operations.** In casting design, it is important to consider the subsequent machining and finishing operations that may be required. For example, if a hole is to be drilled in a casting, it is better to locate it on a flat surface rather than on a curved surface, in order to prevent the drill from wandering. An even better design would incorporate a small dimple on the curved surface as a starting point for the drilling operation. Castings should also include features that allow them to be clamped easily on to machine tools, if secondary machining operations are necessary.

9. **Integrated Computational Materials Engineering (ICME).** The use of modern computational tools allows identification of design and manufacturing issues and allows the prediction of material properties and microstructure that results from a particular mold cavity design. The time spent in process simulation actually saves time that normally would be expended in tooling rework and cost associated with defects.

Selecting a Casting Process. Casting processes cannot be selected separately from economic considerations, as described in Section 12.4. Table 11.2 lists some of the advantages and limitations of casting processes that have an impact on casting design.

Locating the Parting Line. A casting should be oriented in a mold so that the large portion of the casting is relatively low and the height of the casting is minimized. Part orientation also determines the distribution of porosity; for example, in casting aluminum, hydrogen is soluble in liquid metal but is not soluble as the aluminum solidifies (see Fig. 10.17). Hydrogen bubbles can form during the casting of aluminum, which float upwards due to buoyancy and causing a higher porosity in the top regions of castings. Thus, critical surfaces should be oriented so that they face downwards.

A properly oriented casting then can have the parting line determined; this is the line or plane separating the upper (cope) and lower (drag) halves of molds (see Fig. 11.3). In general, the parting line should be along a flat plane rather than be contoured. Whenever possible, the line should be at the corners or edges of castings rather than on flat surfaces in the middle of the casting, so that the **flash** (material squeezing out between the two halves of the mold) at the parting line will not be as visible. The location of the line is also important because it influences mold design, ease of molding, the number and shape of cores required, method of their support, and the gating system.

The parting line should be placed as low as possible (relative to the casting) for metals with lower density (such as aluminum alloys) and be located at around mid-height for denser metals (such as steels). However, the molten metal should not be allowed to flow vertically, especially when unconstrained by a sprue. The placement of the parting line has a large effect on the remainder of the mold design. For example, in sand casting, it is common practice that the runners, gates, and sprue well are all placed in the drag on the parting line. Also, the placement of the parting line and orientation of the part determine the number of cores required, especially when it is preferable to avoid the use of cores whenever practical.

Locating and Designing Gates. Gates are the connections between the runners and the part to be cast. Important considerations in gating system design are:

- Multiple gates often are preferable, and are necessary for large parts; they have the benefits of allowing lower pouring temperature and reducing temperature gradients in the casting.
- Gates should feed into thicker sections of castings.
- A fillet should be used where a gate meets a casting; this feature produces less turbulence than abrupt junctions.
- The gate closest to the sprue should be placed sufficiently away from the sprue, so that the gate can be easily removed. This distance may be as small as a few millimeters for small castings, and up to 500 mm for large ones.
- The minimum gate length should be three to five times the gate diameter, depending on the metal being cast. The gate cross section should be large enough to allow the filling of the mold cavity and should be smaller than the runner cross section.
- Curved gates should be avoided; when necessary, a straight section in the gate should be located immediately adjacent to the casting.

Runner Design. The runner is a horizontal distribution channel that receives molten metal from the sprue and delivers it to the gates. They are used to trap *dross* (a mixture of oxide and metal that forms on the surface of metals) and keep it from entering the gates and mold cavity. Commonly, dross traps are placed at the ends of runners, and the runner projects above the gates to ensure that the metal in the gates is tapped from below the surface. A single runner is used for simple parts, but two-runner systems may be necessary for more complicated castings.

Designing Various Mold Features. The main goal in designing a *sprue* (described in Section 10.3) is to achieve the required molten-metal flow rates, while preventing *aspiration* (entrainment of air) or excessive dross formation. Flow rates are determined such that turbulence is avoided, but also that the mold is filled quickly as compared to the solidification time required. A *pouring basin* can be used to ensure that the metal flow into the sprue is uninterrupted; also, if molten metal is maintained in the pouring basin during pouring, the dross will float and will not enter the mold cavity. *Filters* are used to trap large contaminants, also serving to reduce the metal velocity and make the flow more laminar. *Chills* can be used to speed solidification of the metal in a particular region of a casting.

Establishing Good Practices. It has been widely observed that a given mold design can produce acceptable castings as well as defective ones, and it rarely will produce only good or only defective castings. To check for defective ones, quality control procedures are necessary. Some common concerns are the following:

- Starting with a high-quality molten metal is essential for producing superior castings. Pouring temperature, metal chemistry, gas entrainment, and handling procedures all can affect the quality of metal being poured into a mold.
- The pouring of the molten metal should not be interrupted, because it can lead to dross entrainment and turbulence. The meniscus of the molten metal in the mold cavity should experience a continuous, uninterrupted, and upward advance.
- The different cooling rates within the body of a casting can cause residual stresses. Thus, stress relieving (Section 4.11) may be necessary to avoid distortions of castings in critical applications.

12.2.2 Design for Expendable-mold Casting

Expendable-mold processes have certain specific design requirements, mainly involving the mold material, part sizes, and the manufacturing method. Recall that a casting in an expendable-mold process, such investment casting, will cool much more slowly than it would in, say, die casting; this has important implications in the layout of molds. Important design considerations for expendable-mold casting are as follows.

Mold Layout. The various features in a mold must be placed logically and compactly, with gates as necessary. One of the most important goals in mold layout is to have solidification initiate at one end of the mold and progress across the casting in a uniform front, with the risers solidifying last. Traditionally, mold layout has been based on experience and on considerations of fluid flow and heat transfer. Commercial computer programs have now become widely available assisting in the analysis of fluid flow and heat transfer. These programs simulate mold filling and allow the rapid evaluation and design of mold layouts.

Riser Design. A major concern in the design of castings is the size of risers and their placement. Risers are very useful in affecting the solidification-front progression across a casting, and are an essential feature in mold layout described previously. Blind risers are good design features and maintain heat longer than open risers do.

Risers are designed according to the following basic rules:

1. The riser must not solidify before the casting does. This rule usually is satisfied by avoiding the use of small risers and by using cylindrical risers with small aspect ratios (i.e., small ratios of height to cross section). Spherical risers are the most efficient shape, but are difficult to work with.
2. The riser volume must be sufficiently large to provide enough molten metal to compensate for shrinkage in the casting.
3. Junctions between the casting and the riser should not develop hot spots, where shrinkage porosity can occur.
4. Risers must be placed such that the molten metal can reach locations where it is most needed.

5. There must be sufficient pressure to drive the molten metal into locations in the mold where it is needed. Risers are not as useful for metals with low density (such as aluminum alloys) as they are for those with higher density (such as steel and cast irons).

6. The pressure head from the riser should suppress cavity formation and encourage complete filling of the mold cavity.

Machining Allowance. Most expendable-mold castings require some additional finishing operations, such as machining and grinding; allowances have to be included in casting design for these operations. Machining allowances, which are included in pattern dimensions, depend on the type of casting operation, noting also that they increase with the size and section thickness of the casting. Allowances usually range from about 2 to 5 mm for small castings to more than 25 mm for large castings.

12.2.3 Design for Permanent-mold Casting

General design guidelines for permanent-mold casting are described in Example 12.1. Although designs may be modified to eliminate the draft, for better dimensional accuracy, a draft angle of 0.5° or even 0.25° is usually required; otherwise, galling (localized seizure or sticking of two surfaces, Section 33.5) may occur between the part and the dies, causing distortion of the casting. Die-cast parts are nearly net shaped, typically requiring only the removal of gates and minor trimming to remove flashing and other minor defects. The surface finish and dimensional accuracy of die-cast parts are very good (see Table 11.3) and, in general, they do not require a machining allowance.

Case Study 12.1 Illustrations of Poor and Good Casting Designs

Several examples of poor and good designs in permanent-mold and die casting are illustrated in Fig. 12.4. The significant differences in design are outlined here for each example:

1. The lower portion of the design on the left has a thin wall, with no apparent specific function; at this location, the part may fracture if subjected to high forces or to impact. The good design eliminates this possibility, and also may simplify die and mold making.

2. Large flat surfaces always present difficulties, as they tend to warp and develop uneven surfaces. A common practice to avoid this situation is to break up the surface with *ribs* (see Fig. 12.3) and serrations on the reverse side of the casting. This approach greatly reduces part distortion, while not adversely affecting the appearance and function of the flat surface. In addition to ribs, it is beneficial to use a *textured* surface, as shown in Fig. 12.4b, since very smooth surfaces are difficult to cast without objectionable aesthetic features.

3. This example of poor and good design is relevant not only to castings, but also to parts that are subsequently machined or ground. It is difficult to produce sharp internal radii or corners that may be required for functional purposes, such as inserts designed to reach the bottom of the part cavity. Also, in the case of lubricated cavities, the lubricant can accumulate at the bottom and, because it is incompressible, prevent full insertion of an insert. The placement of a small radius at the corners or periphery at the bottom of the part eliminates this problem.

4. A cast part could function, for instance, as a knob to be gripped and rotated, hence the outer features along its periphery. Note in the design on the left that the inner periphery of the knob also has features which are not functional but help save material; the die for the good design is easier to manufacture.

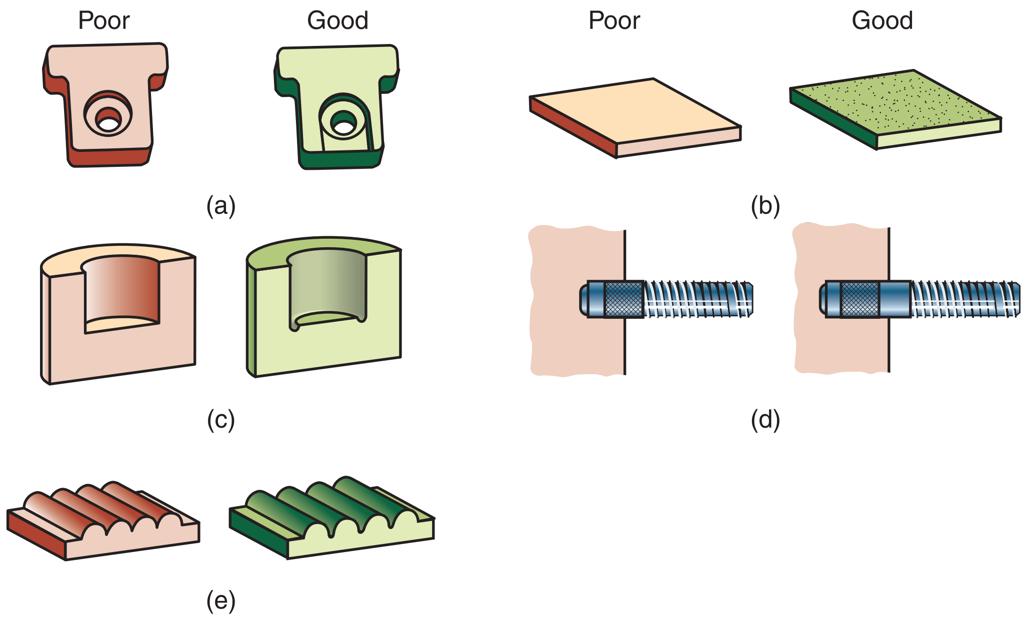


Figure 12.4: Examples of undesirable (poor) and desirable (good) casting designs.

5. Note that the poor design has sharp fillets at the base of the longitudinal grooves, indicating that the die has sharp (knife-edge) protrusions. It is thus possible that, with overextended use of the die, these edges may chip off.

12.2.4 Computer Modeling of Casting Processes

Because casting involves complex interactions among several material and process variables, a quantitative study of these interactions is essential to the proper design and production of high-quality castings. Rapid advances in modeling techniques have led to important innovations in modeling casting processes. These include fluid flow, heat transfer, and the microstructures developed during solidification under various casting conditions, as described in Section 10.3.

Simulations are capable of predicting, for example, the width of the mushy zone (see Fig. 10.4) during solidification and the grain size in castings. Similarly, the capability to calculate *isotherms* (lines of equal temperature) give insight into possible hot spots and the subsequent development of shrinkage cavities. With the availability of user-friendly software and advances in computer-aided design and manufacturing (Chapter 38), modeling techniques have become easier to implement. The benefits of this approach are improved quality, easier planning and cost estimating, increased productivity, and faster response to design changes.

12.3 Casting Alloys

The general properties and typical applications of ferrous and nonferrous metals and alloys were presented in Chapters 5 and 6, respectively. This section describes the properties and applications of cast metals and alloys; their properties and casting and manufacturing characteristics are summarized in Fig. 12.5 and Tables 12.2 through 12.5. In addition to their casting characteristics, other important considerations in casting alloys include their machinability and weldability, since they are assembled with other components to produce the entire assembly.

The most commonly used casting alloy (in tonnage) is gray iron, followed by ductile iron, malleable iron, steel, copper, aluminum, magnesium, and zinc. Shipments of castings in the United States alone are around 9.07 million metric tons per year.

12.3.1 Nonferrous Casting Alloys

Common nonferrous casting alloys are as follows:

Aluminum-based Alloys. Aluminum alloys have a wide range of mechanical properties, mainly because of various hardening mechanisms and heat treatments that can be used (Section 4.9). Parts made of aluminum and magnesium alloys are known as **light-metal** castings. They have high electrical conductivity and generally good atmospheric corrosion resistance; however, their resistance to all alkalines and some acids is poor, and care must be taken to prevent galvanic corrosion.

Aluminum alloys are lightweight, nontoxic, and have good machinability. Except for alloys containing silicon, they generally have low resistance to wear and abrasion. They have numerous applications, including architectural and decorative purposes. An increasing trend is their use in automobiles, for components such as engine blocks, cylinder heads, intake manifolds, transmission cases, suspension components, wheels and brakes.

Magnesium-based Alloys. These alloys have the lowest density of all commercial casting alloys. They have good corrosion resistance and moderate strength, depending on the particular heat treatment used. Typical applications include automotive wheels, housings, and air-cooled engine blocks. Because of their light weight, magnesium castings are being increasingly used in automobiles to increase fuel economy.

Copper-based Alloys. These alloys have the advantages of good electrical and thermal conductivity, corrosion resistance, and nontoxicity, as well as wear properties suitable as bearing materials. A wide variety of copper-based alloys is available, including brasses, aluminum bronzes, phosphor bronzes, and tin bronzes.

Zinc-based Alloys. A low-melting-point alloy group, zinc-based alloys have good corrosion resistance, good fluidity, and sufficient strength for structural applications. These alloys are commonly used in die casting, particularly for parts with thin walls and complex shapes.

Tin-based Alloys. Although low in strength, these alloys have good corrosion resistance and are typically used for linings or bearing surfaces.

High-temperature Alloys. These alloys have a wide range of properties, and typically require temperatures of up to 1650°C for casting titanium and superalloys, and even higher for refractory alloys (Mo, Nb, W, and Ta). Special techniques are used to cast these alloys for nozzles and various jet- and rocket-engine components. Some high-temperature alloys are more suitable and economical for casting than for shaping by other manufacturing methods, such as forging and powder metallurgy techniques.

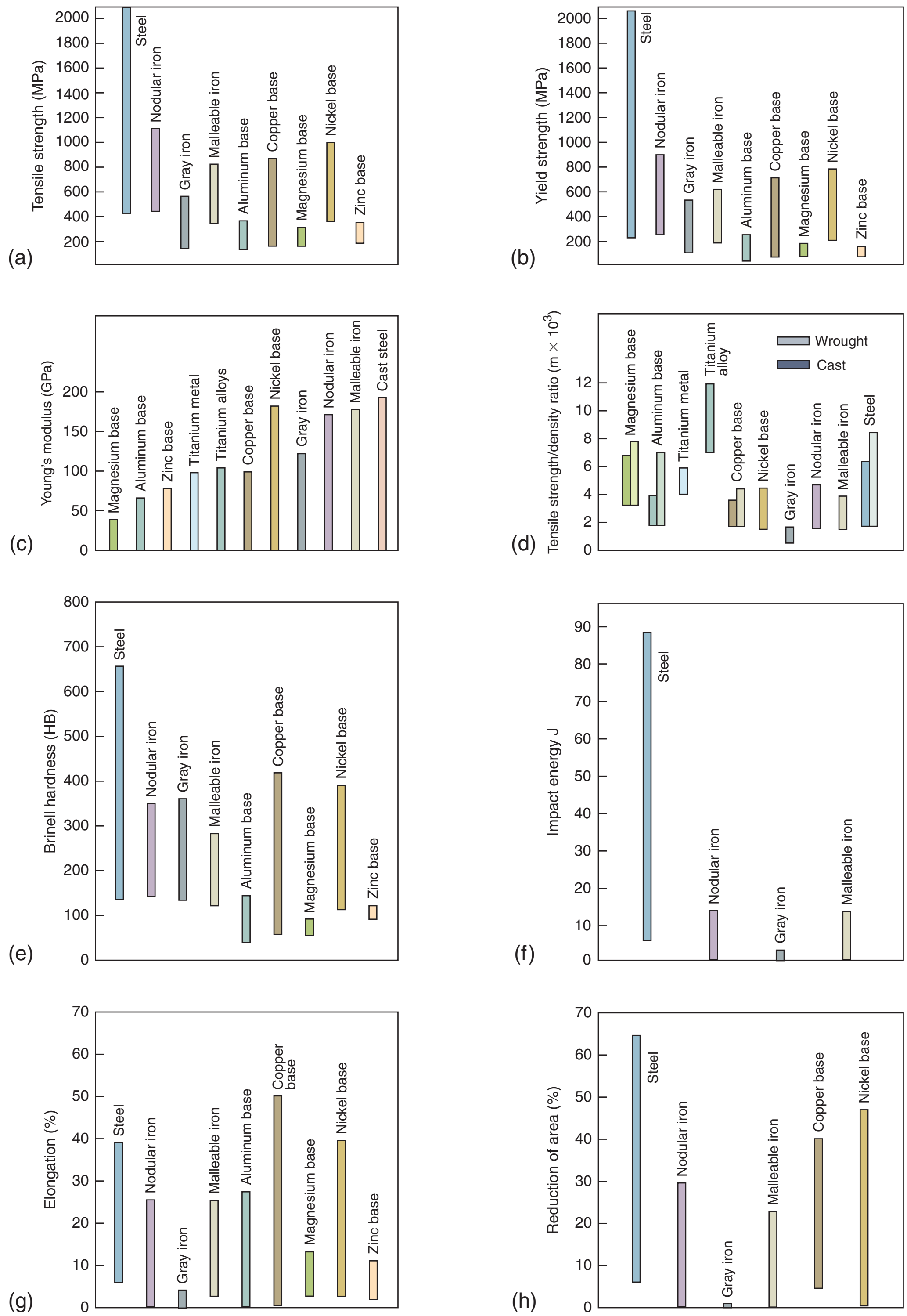


Figure 12.5: Mechanical properties for various groups of cast alloys. Note that even within the same group, the properties vary over a wide range, particularly for cast steels. *Source:* Courtesy of Steel Founders' Society of America.

Table 12.2: Typical Applications for Castings and Casting Characteristics.

Type of alloy	Castability*	Weldability*	Machinability*	Typical applications
Aluminum	E	F	G–E	Pistons, clutch housings, intake manifolds
Copper	F–G	F	F–G	Pumps, valves, gear blanks, marine propellers
Iron				
Ductile	G	D	G	Crankshafts, heavy-duty gears
Gray	E	D	G	Engine blocks, gears, brake disks and drums, machine bases
Malleable iron	G	D	G	Farm and construction machinery, heavy-duty bearings, railroad rolling stock
White iron	G	VP	VP	Mill liners, shot-blasting nozzles, railroad brake shoes, crushers, and pulverizers
Magnesium	G–E	G	E	Crankcase, transmission housings
Nickel	F	F	F	Gas turbine blades, pump and valve components for chemical plants
Steel				
Carbon and low-alloy	F	E	F	Die blocks, heavy-duty gear blanks, aircraft undercarriage members, railroad wheels
High-alloy	F	E	F	Gas-turbine housings, pump and valve components, rock-crusher jaws
Zinc	E	D	E	Door handles, radiator grills

Note: * E = excellent; G = good; F = fair; VP = very poor; D = difficult

Table 12.3: Properties and Typical Applications of Cast Irons.

Cast iron	Type	Ultimate tensile strength (MPa)	Yield strength (MPa)	Elongation in 50 mm (%)	Typical applications
Gray	Ferritic	170	140	0.4	Pipe, sanitary ware
	Pearlitic	275	240	0.4	Engine blocks, machine tools
	Martensitic	550	550	0	Wear surfaces
Ductile (Nodular)	Ferritic	415	275	18	Pipe, general service
	Pearlitic	550	380	6	Crankshafts, highly stressed parts
	Tempered martensite	825	620	2	High-strength machine parts, wear-resistant parts
Malleable	Ferritic	365	240	18	Hardware, pipe fittings, general engineering service
	Pearlitic	450	310	10	Railroad equipment, couplings
	Tempered martensite	700	550	2	Railroad equipment, gears, connecting rods
White	Pearlitic	275	275	0	Wear-resistant parts, mill rolls

Table 12.4: Mechanical Properties of Gray Cast Irons.

ASTM class	Ultimate tensile strength (MPa)	Compressive strength (MPa)	Elastic modulus (GPa)	Hardness (HB)
20	152	572	66–97	156
25	179	669	79–102	174
30	214	752	90–113	210
35	252	855	100–119	212
40	293	965	110–138	235
50	362	1130	130–157	262
60	431	1293	141–162	302

Table 12.5: Properties and Typical Applications of Nonferrous Cast Alloys.

Alloys (UNS)	Condition	Ultimate tensile strength (MPa)	Yield strength (MPa)	Elongation in 50 mm (%)	Typical applications
Aluminum alloys					
195 (AO1950)	Heat treated	220–280	110–220	8.5–2	Sand castings
319 (AO3190)	Heat treated	185–250	125–180	2–1.5	Sand castings
356 (AO3560)	Heat treated	260	185	5	Permanent mold castings
Copper alloys					
Red brass (C83600)	Annealed	235	115	25	Pipe fittings, gears
Yellow brass (C86400)	Annealed	275	95	25	Hardware, ornamental
Manganese bronze (C86100)	Annealed	480	195	30	Propeller hubs, blades
Sulfur tin bronze (C83470)	As cast	190	95	15	Water supply piping and fittings, valves
Copper Bismuth (C89836)	As cast	230	95	20	Antimicrobial; water supply and fittings
Gun metal (C90500)	Annealed	275	105	30	Pump parts, fittings
Nickel silver (C97600)	Annealed	275	175	15	Marine parts, valves
Magnesium alloys					
AZ91A	F	230	150	3	Die castings
AZ63A	T4	275	95	12	Sand and permanent mold castings
AZ91C	T6	275	130	5	High-strength parts
EZ33A	T5	160	110	3	Elevated-temperature parts
HK31A	T6	210	105	8	Elevated-temperature parts
QE22A	T6	275	205	4	Highest-strength parts

12.3.2 Ferrous Casting Alloys

Commonly cast ferrous alloys are as follows:

Cast Irons. Cast irons represent the largest quantity of all metals cast. They can easily be cast into intricate shapes, and generally possess several desirable properties, such as high hardness, wear resistance, and good machinability. The term *cast iron* refers to a family of alloys, and as described in Section 4.6, they are classified as gray cast iron (gray iron), ductile (nodular or spheroidal) iron, white cast iron, malleable iron, and compacted-graphite iron. Their general properties and typical applications are given in Tables 12.3 and 12.4.

1. **Gray cast iron.** Gray iron castings have relatively few shrinkage cavities and low porosity. Various forms of gray cast iron are *ferritic, pearlitic,* and *martensitic*, and because of differences in their structures, each type has different properties (Table 12.4). Gray cast irons are specified by a two-digit ASTM designation; thus, for example, class 20 specifies that the material must have a minimum tensile strength of 140 MPa. Typical uses of gray cast iron are in engine blocks, electric-motor housings, pipes, and wear surfaces for machines. Also, because of its high damping capacity, gray iron is used widely for machine-tool bases (Section 25.3).

2. **Ductile (nodular) iron.** Typically used for machine parts, housings, gears, pipe, rolls for rolling mills, and automotive crankshafts, ductile irons are specified by a set of two-digit numbers. For example, class or grade 80-55-06 indicates that it has a minimum tensile strength of 550 MPa, a minimum yield strength of 380 MPa, and 6% elongation in 50 mm.

3. **White cast iron.** Because of its very high hardness and wear resistance, white cast iron is used typically for rolls for rolling mills, railroad-car brake shoes, and liners in machinery for processing abrasive materials.

4. **Malleable iron.** The principal use of malleable iron is for railroad equipment and various types of hardware, fittings, and components for electrical applications. Malleable irons are specified by a five-digit designation. For example, 35018 indicates that the yield strength is 240 MPa and its elongation is 18% in 50 mm.

5. **Compacted-graphite iron.** First produced commercially in 1976, compacted-graphite iron (CGI) has properties that are between those of gray irons and ductile irons. Gray iron has good damping and thermal conductivity, but low ductility, whereas ductile iron has poor damping and thermal conductivity, but high tensile strength and fatigue resistance. Compacted-graphite iron has damping and thermal properties similar to gray iron and strength and stiffness that are comparable to those of ductile iron. Because of its strength, castings made of CGI can be smaller, thus lighter. This iron is easy to cast and has properties that are consistent throughout the casting. Moreover, its machinability is better than that of ductile iron (an important consideration since compacted-graphite iron is used for automotive engine blocks and cylinder heads, which require extensive machining).

Cast Steels. Because of the high temperatures required to melt steels (up to about 1650°C), casting steels requires special considerations. The high temperatures involved present difficulties in the selection of mold materials, particularly in view of the high reactivity of steels with oxygen during the melting and pouring of the metal. Steel castings possess properties that are more uniform (isotropic) than those made by mechanical working processes (Part III). Although they can be welded, welding alters the cast microstructure in the heat-affected zone (see Fig. 30.15), thus influencing the strength, ductility, and toughness of the base metal. Subsequent heat treatment would be required to restore the mechanical properties of the casting. Cast weldments have gained importance for assembling large machines and structures. Cast steels have important applications in mining, chemical plants, oil fields, heavy construction, and equipment for railroads.

Cast Stainless Steels. Casting of stainless steels involves considerations similar to those for steels. Stainless steels generally have long freezing ranges (see Section 10.2.2) and high melting temperatures. They can develop several structures, depending on their composition and processing parameters. Cast stainless steels are available in various compositions, and they can be heat treated and welded. Cast stainless-steel parts have high heat and corrosion resistance, especially useful in the chemical and food industries. Nickel-based casting alloys are used for very corrosive environments and for very high temperature service.

Table 12.6: General Cost Characteristics of Casting Processes.

Casting process	Cost* Die	Equipment	Labor	Production rate (pieces/hr)
Sand	L	L	L–M	< 20
Shell mold	L–M	M–H	L–M	< 10
Plaster	L–M	M	M–H	< 10
Investment	M–H	L–M	H	< 1000
Permanent mold	M	M	L–M	< 60
Die	H	H	L–M	< 200
Centrifugal	M	H	L–M	< 50

* L = low; M = medium; H = high.

12.4 Economics of Casting

As in all manufacturing processes, the cost of each cast part (**unit cost**) depends on several factors, including materials, equipment, and labor. Recall that among various casting processes described in Chapter 11, some require more labor than others, some require expensive dies and machinery, and some require a long production times to produce the castings (Table 12.6). Each of these individual factors affects the overall cost of a casting operation and to varying degrees.

As can be noted in Table 12.6, relatively little cost is involved in making molds for sand casting, whereas molds for other casting processes and especially dies for die-casting require expensive materials and manufacturing operations. There are also major costs involved in making patterns for casting, although much progress continues to be made in utilizing additive manufacturing techniques (Section 20.10) to reduce costs and production time.

Costs are also incurred in melting and pouring the molten metal into molds, and in heat treating, cleaning, and inspecting the castings. Heat treating is an important part of the production of many alloy groups (especially ferrous castings), and may be necessary for improving the mechanical properties. However, heat treating may also introduce another set of production problems, such as scale formation on casting surfaces and warpage of the part, that can be a significant aspect of production costs.

The labor and the skills required can vary considerably, depending on the particular casting operation and level of automation in the foundry. Investment casting, for example, requires much labor because of the several steps involved in the operation, although some automation in a plant can be implemented, such as using robots (Fig. 11.13a). On the other hand, operations such as in highly automated die-casting maintain high production rates, with little labor involved.

Note also that the equipment cost per casting decreases as the number of parts cast increases. Sustained high production rates can justify the high cost of dies and machinery. However, if demand is relatively small, the cost per casting increases rapidly. It then becomes more economical to manufacture the parts either by other casting processes described in this chapter or by considering other manufacturing processes, described in detail in Parts III and IV, singly or in combination.

Summary

- General guidelines have been established to aid in the production of castings without defects, and to meet dimensional tolerances, surface finish, service requirements, and various specifications and standards. The guidelines concern the shape of the casting and the various techniques to minimize hot spots that could lead to shrinkage cavities. Because of the large number of variables involved, close control of all parameters is essential, particularly those related to the nature of liquid-metal flow into molds and dies, and the rate of cooling in different regions of the mold.

- Numerous nonferrous and ferrous casting alloys are available, with a wide range of properties, casting characteristics, and applications. Because many castings are designed and produced to be assembled with other mechanical components and structures (subassemblies), several other considerations, such as weldability, machinability, and surface characteristics, also are important.
- Within the limits of good performance, the economics of casting is just as important as the technical considerations. Factors affecting the overall cost are the cost of materials, molds, dies, equipment, and labor, each of which varies with the particular casting operation.

Key Terms

Cast iron
Compacted-graphite iron
Design principles
Draft
Flash
Hot spots
Machining allowance
Parting line
Patternmaker's shrinkage allowance
Porosity
Shrinkage cavities
Unit cost

Bibliography

Abrasion-Resistant Cast Iron Handbook, American Foundry Society, 2000.
ASM Handbook, Vol. 15: **Casting**, ASM International, 2008.
ASM Specialty Handbook: **Cast Irons**, ASM International, 1996.
Campbell, J., **Complete Casting Handbook: Metal Casting Processes, Techniques and Design**, Butterworth-Heinemann, 2011.
Casting Design and Performance, ASM International, 2009.
Casting Design Handbook, ASM International, 2012.
Powell, G.W., Cheng, S.-H., and Mobley, C.E., Jr., **A Fractography Atlas of Casting Alloys**, Battelle Press, 1992.
Product Design for Die Casting, Diecasting Development Council, 1988.
Steel Castings Handbook, 6th ed., ASM International, 1995.
Tuttle, R.B., **Foundry Engineering: The Metallurgy and Design of Castings**, CreateSpace Independent Publishing Platform, 2012.

Review Questions

12.1. Why are steels more difficult to cast than cast irons?

12.2. What is the significance of hot spots in metal casting?

12.3. What is shrinkage allowance? Machining allowance?

12.4. Explain the reason for drafts in molds.

12.5. Why are ribs useful for flat surfaces?

12.6. What are light castings and where are they used most commonly?

12.7. Name the types of cast irons generally available, and list their major characteristics and applications.

12.8. Comment on your observations regarding Fig. 12.5.

12.9. Describe the difference between a runner and a gate.

12.10. What is the difference between machining allowance and dimensional tolerance?

12.11. What is dross? Can it be eliminated?

Qualitative Problems

12.12. Describe your observation concerning the design changes shown in Fig. 12.1.

12.13. If you need only a few castings of the same design, which three processes would be the most expensive per piece cast?

12.14. Do you generally agree with the cost ratings in Table 12.6? If so, why?

12.15. Describe the nature of the design differences shown in Fig. 12.4. What general principles do you observe in this figure?

12.16. Note in Fig. 12.5 that the ductility of some cast alloys is very low. Do you think that this should be a significant concern in engineering applications of castings? Explain.

12.17. Do you think that there will be fewer defects in a casting made by gravity pouring versus one made by pouring under pressure? Explain.

12.18. Explain the difference in the importance of drafts in green-sand casting versus permanent-mold casting.

12.19. What type of cast iron would be suitable for heavy-machine bases, such as presses and machine tools? Why?

12.20. Explain the advantages and limitations of sharp and rounded fillets, respectively, in casting design.

12.21. Explain why the elastic modulus, E, of gray cast iron varies so widely, as shown in Table 12.4.

12.22. If you were to incorporate lettering or numbers on a sand-cast part, would you make them protrude from the surface or recess them into the surface? What if the part were to be made by investment casting? Explain your answer.

12.23. The general design recommendations for a well in sand casting (see Fig. 11.3) are that (a) its diameter should be at least twice the exit diameter of the sprue and (b) its depth should be approximately twice the depth of the runner. Explain the consequences of deviating from these guidelines.

12.24. The heavy regions of parts typically are placed in the drag in sand casting and not in the cope. Explain why.

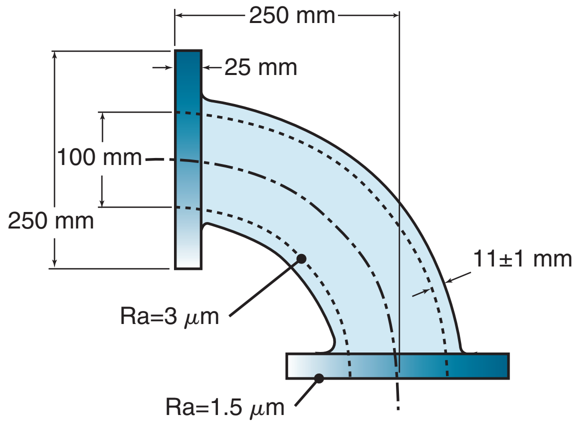
12.25. What are the benefits and drawbacks to having a pouring temperature that is much higher than the metal's melting temperature? What are the advantages and disadvantages in having the pouring temperature remain close to the melting temperature?

Quantitative Problems

12.26. When designing patterns for casting, patternmakers use special rulers that automatically incorporate solid shrinkage allowances into their designs. For example, a 300 mm patternmaker's ruler is longer than 300 mm. How long should a patternmaker's ruler be for making patterns for (a) aluminum castings and (b) high-manganese steel?

12.27. Using the data given in Table 12.2, develop approximate plots of (a) castability versus weldability and (b) castability versus machinability, for at least five of the materials listed in the table.

12.28. The part in the figure below is to be cast of 10% Sn bronze at the rate of 100 parts per month. To find an appropriate casting process, consider all casting processes, then reject those that are (a) technically inadmissible, (b) technically feasible but too expensive for the purpose, and (c) identify the most economical one. Write a rationale using common-sense assumptions about cost.



Synthesis, Design, and Projects

12.29. Describe the general design considerations pertaining to metal casting.

12.30. Add more examples to those shown in Fig. 12.2.

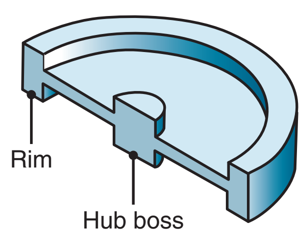
12.31. Explain how ribs and serrations are helpful in casting flat surfaces that otherwise may warp. Give a specific illustration.

12.32. List casting processes that are suitable for making hollow parts with (a) complex external features, (b) complex internal features, and (c) both complex external and complex internal features. Explain your choices.

12.33. Small amounts of slag and dross often persist after skimming and are introduced into the molten metal flow in casting. Recognizing that slag and dross are less dense than the molten metal, design mold features that will remove small amounts of slag before the metal reaches the mold cavity.

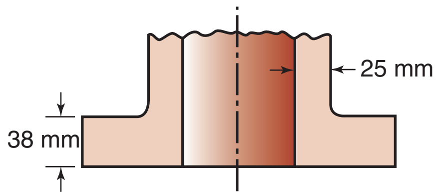
12.34. If you need only a few units of a particular casting, which process(es) would you use? Why?

12.35. For the cast metal wheel illustrated below, show how (a) riser placement, (b) core placement, (c) padding, and (d) chills may be used to help feed molten metal and eliminate porosity in the isolated hub boss.

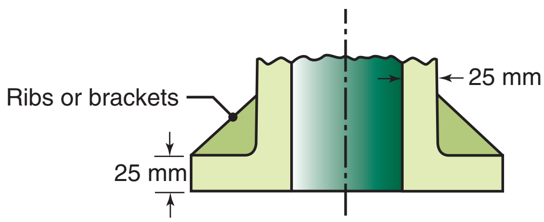


12.36. Assume that the introduction to this chapter is missing. Write a brief introduction to highlight the importance of the topics covered in it.

12.37. In the figure below, the original casting design shown in (a) was resized and modified to incorporate ribs in the design shown in (b). The casting is round and has a vertical axis of symmetry. What advantages do you think the new design has as a functional part over the old one?

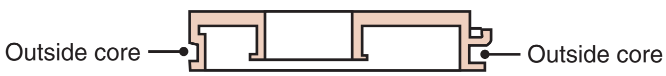


(a)

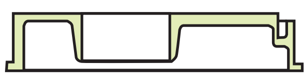


(b)

12.38. An incorrect and a correct design for casting are shown below. Review the changes made and comment on their advantages.

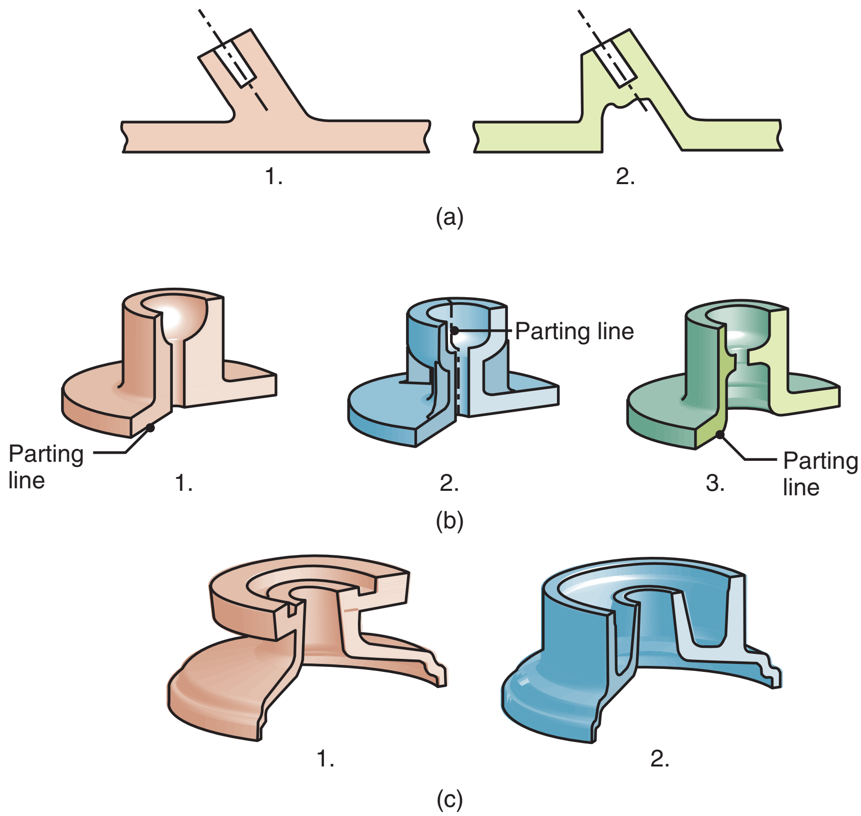


(a) Incorrect



(b) Correct

12.39. Three sets of designs for die casting are shown below. Note the changes made to design 1 and comment on the reasons for then or them.



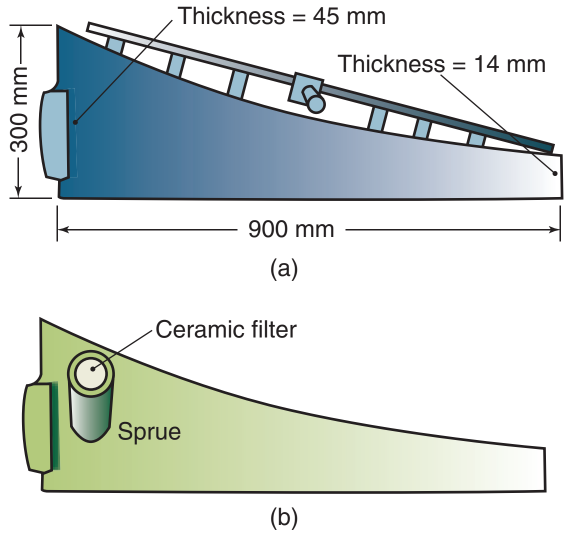
12.40. Using the method of inscribed circles shown in Fig. 12.2, justify the trend shown in Fig. 12.3.

12.41. A growing trend is the production of patterns and molds through rapid prototyping approaches described in Chapter 20. Consider the case of an injection molding operation, where the patterns are produced by rapid prototyping, and then hand assembled onto trees and processed in traditional fashion. What design rules discussed in this chapter would still be valid, and which would not be as important in this case?

12.42. Repeat Problem 12.41 for the case where (a) a pattern for sand casting is produced by rapid prototyping; (b) a sand mold for sand casting is produced.

12.43. It is sometimes desirable to cool metals more slowly than they would be if the molds were maintained at room temperature. List and explain the methods you would use to slow down the cooling process.

12.44. The two illustrations shown are proposed designs of a gating system for an aluminum low-power water turbine blade. The first uses a conventional sprue-runner-gate system, while the second uses a ceramic filter underneath a pouring cup, but without gates (direct pour method). Evaluate the two designs, and list their advantages and disadvantages. Based on your analysis, select a preferred approach.



12.45. Note that in cast jewelry, gemstones are usually cast in place; that is, they are not attached after the ring is cast, but are incorporated into the ring. Design a ring with a means of securing a gemstone in the wax pattern, such that it will remain in the mold as the wax is being melted. Could such an approach be used in lost foam casting?

PART III
Forming and Shaping Processes and Equipment

Examination of various products soon leads to the realization that a wide variety of materials and processes have been used in making them, as can also be seen from the example of the automobile shown in Fig. III.1. It will be noted that some products consist of only one part (screws, bolts, washers) or a few parts (pens, eyeglasses, microscopes), while others consist of hundreds or thousands of parts (automobiles, computers) or millions of parts (airplanes, ships). Some products are used for routine applications (paper clips, forks, door keys) while others are used in critical applications (elevator cables, stents, turbine blades). Some are very thin (aluminum foil, plastic film) whereas others are thick (boiler plates, submarine hulls).

Note that the words **forming** and **shaping** are both used in the title of Part III of this book. Although there are not always clear distinctions between the two terms, *forming* generally indicates changing the shape of an existing solid body. Thus, in **forming processes**, the starting material (workpiece, stock, or blank) may be in the shape of a plate, sheet, bar, rod, wire, or tubing. For example, a common wire coat hanger is made by taking a straight piece of wire and bending and twisting it into the shape of a hanger. Likewise, the sheet-metal body for a washer or dryer is generally made of flat, cold-rolled steel (occasionally aluminum) sheet, which is then formed into various shapes.

Shaping processes typically involve molding or casting, producing a part that generally is at or near the final desired shape. A plastic coat hanger, for example, is made by forcing molten plastic into a two-piece mold, with a cavity in the shape of the hanger. Computer mouse or video game console housings, refrigerator-door liners, some auto-body parts, and countless other plastic products are likewise shaped by forcing molten polymer into a mold, and removing it after it solidifies.

Some forming and shaping operations produce long *continuous* products, such as plates, sheets, tubing, wire, and rod and bars, which then are shaped into specific products. Rolling, extrusion, and drawing processes (Chapters 13 and 15) are capable of making such long products, which then are cut into desired lengths. On the other hand, processes such as forging (Chapter 14), sheet metal forming and stamping (Chapter 16), powder metallurgy (Chapter 17), ceramic slip casting and glass pressing (Chapter 18),

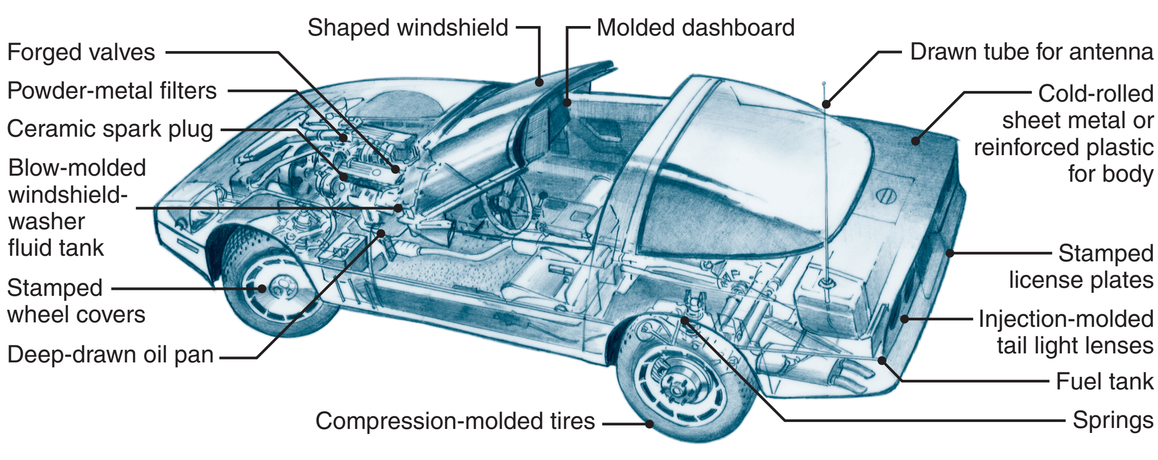


Figure III.1: Formed and shaped parts in a typical automobile.

Table III.1: General Characteristics of Forming and Shaping Processes.

Process	Characteristics
Rolling	
Flat	Production of flat plate, sheet, and foil at high speeds; good surface finish, especially in cold rolling; very high capital investment; low-to-moderate labor cost
Shape	Production of various structural shapes (such as I-beams and rails) at high speeds; includes thread rolling; requires shaped rolls and expensive equipment; low-to-moderate labor cost; requires moderate operator skill
Forging	Production of discrete parts with a set of dies; some finishing operations usually required; usually performed at elevated temperatures, but also cold for smaller parts; die and equipment costs are high; moderate-to-high labor cost; requires moderate-to-high operator skill
Extrusion	Production of long lengths of solid or hollow shapes with constant cross section; product is then cut into desired lengths; usually performed at elevated temperatures; cold extrusion has similarities to forging and is used to make discrete products; moderate-to-high die and equipment cost; low-to-moderate labor cost; requires low-to-moderate operator skill
Drawing	Production of long rod and wire with various cross sections; good surface finish; low-to-moderate die, equipment, and labor costs; requires low-to-moderate operator skill
Sheet-metal forming	Production of a wide variety of shapes with thin walls and simple or complex geometries; generally low-to-moderate die, equipment, and labor costs; requires low-to-moderate operator skill
Powder metallurgy	Production of simple or complex shapes by compacting and sintering metal powders; moderate die and equipment cost; low labor cost and skill
Processing of plastics and composite materials	Production of a wide variety of continuous or discrete products by extrusion, molding, casting, and fabricating processes; moderate die and equipment costs; requires high operator skill in processing of composite materials
Forming and shaping of ceramics	Production of discrete products by various shaping, drying, and firing processes; low-to-moderate die and equipment cost; requires moderate-to-high operator skill
Additive manufacturing	Production of discrete parts by various computer-controlled methods; no dies involved; low-to-high equipment cost; requires high operator skill

processes involving plastics and reinforced plastics (Chapter 19), and additive manufacturing (Chapter 20), typically produce *discrete* products.

The initial raw material used in forming and shaping metals is usually molten metal, which is *cast* into individual *ingots* or *continuously cast* into slabs, rods, or pipes. Cast structures are converted to *wrought structures* by plastic-deformation processes. The raw material used also may consist of *metal powders*, which then are pressed and sintered (heated without melting) into individual parts. For plastics, the starting material is usually pellets, flakes, or powder; for ceramics, it is clay, powder, and oxides, obtained from ores or produced synthetically.

In this part of the text, the important factors involved in each forming and shaping process are described, along with how material properties and processes affect the quality and integrity of the product made (Table III.1). Detailed mathematical models of processes are now available and can be found in the Bibliographies at the end of the chapters. This book provides only general models for the various forming and shaping processes considered. It will also become clear that some materials can be processed only by certain specific manufacturing methods, and why parts with particular shapes can only be processed by certain specific techniques and not by others. Also included are descriptions of the characteristics of the machinery and equipment used, as they can significantly affect product quality, production rate, and the economics of a particular manufacturing operation.

Chapter 13

Metal-rolling Processes and Equipment

- This chapter describes the rolling of metals, which is the most important metal-forming operation based on the total volume of metals rolled.
- The chapter begins with a description of the flat-rolling process and analysis of the force, torque, and power required, in terms of relevant material and process parameters; it also includes a review of defects and their causes in rolled products.
- Shape-rolling processes are then described, where workpieces pass through a series of shaped rolls.
- Special rolling processes such as cross rolling, ring rolling, thread rolling, tube rolling, and tube piercing are also described.
- The chapter ends with a description of the characteristics of rolling mills and roll arrangements for making specific products.

Typical products made by various rolling processes: Plates for ships, bridges, structures, large machines; sheet metal for car bodies, aircraft fuselages, appliances, containers; foil for packaging; I-beams, railroad rails, architectural shapes, large rings, seamless pipe and tubing; bolts, screws, and threaded components.

Alternative processes: Continuous casting, extrusion, drawing, machining of threaded components.

13.1 Introduction

Rolling is the process of reducing the thickness or changing the cross section of a long workpiece by compressive forces applied through a set of **rolls** (Fig. 13.1). Rolling, which accounts for about 90% of all metals

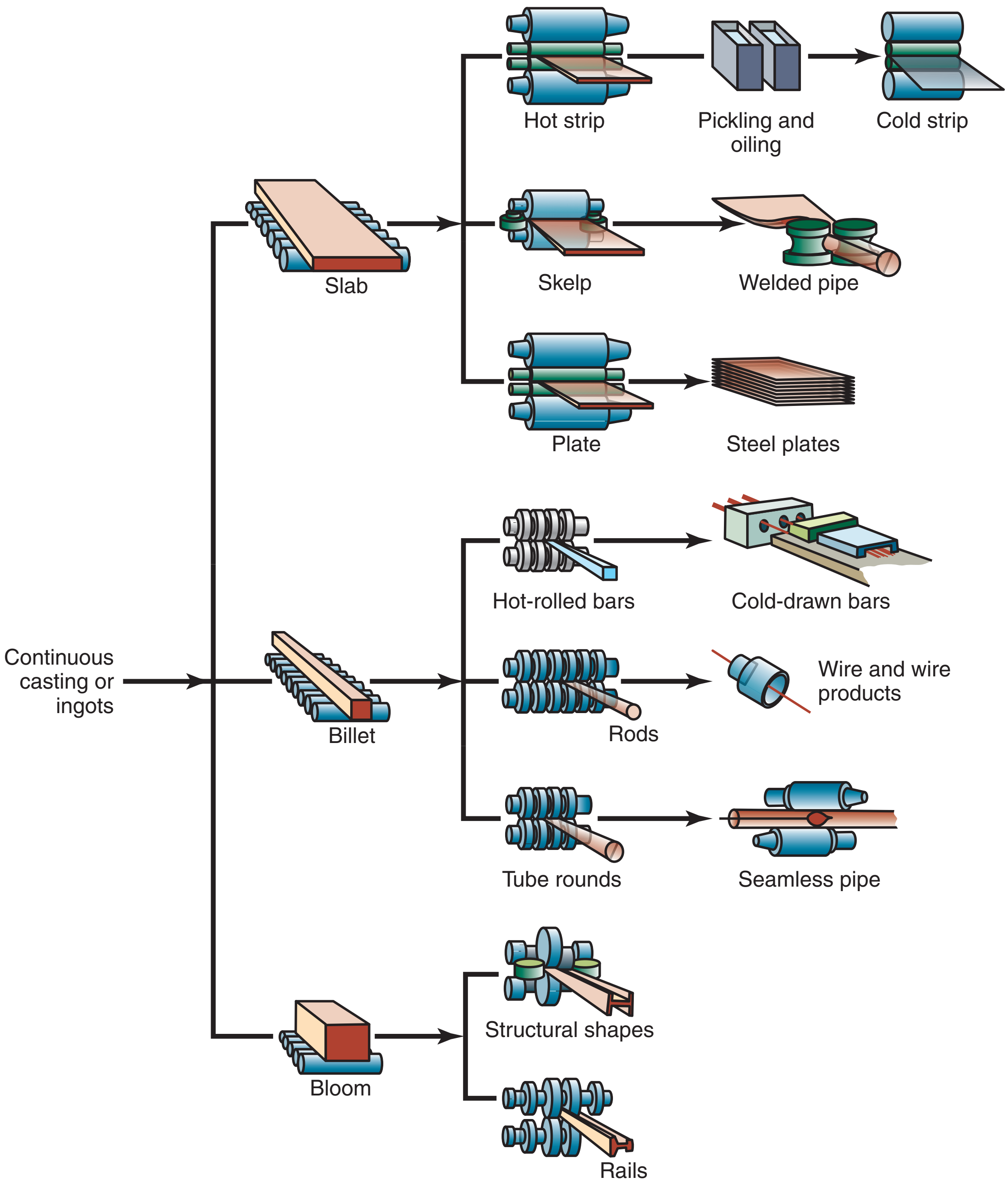


Figure 13.1: Schematic outline of various flat-rolling and shape-rolling processes. *Source:* After American Iron and Steel Institute.

produced by metalworking processes, was first developed in the late 1500s. Modern steelmaking practices and the production of various ferrous and nonferrous metals and alloys now generally integrate continuous casting with rolling processes. This method greatly improves productivity and lowers production costs, as described in Section 5.4. Nonmetallic materials also are rolled to reduce their thickness and enhance their properties.

Rolling is first carried out at elevated temperatures (*hot rolling*). During this stage, the coarse-grained, brittle, and porous structure of the ingot (or the continuously cast metal) is broken down into a *wrought structure*, having a finer grain size and enhanced properties, such as increased strength and hardness. Subsequent rolling is generally carried out at room temperature (*cold rolling*), whereby the rolled sheet has higher strength and hardness, and better surface finish. However, cold rolling will result in a product with *anisotropic* properties, due to preferred orientation or mechanical fibering, described in Section 1.6.

Plates generally have a thickness of greater than 6 mm, and are used for structural applications, such as ship hulls, boilers, bridges, and heavy machinery. Plates can be as thick as 300 mm for large structural supports, 150 mm for reactor vessels, and 100 to 125 mm for machinery frames and warships.

Sheets are generally less than 6 mm thick, and are typically provided to manufacturing facilities as coils, weighing as much as 30,000 kg, or as flat sheets for further processing into a wide variety of sheet-metal products. Sheets are typically used for aircraft bodies, appliances, food and beverage containers, and kitchen and office equipment. Commercial aircraft fuselages and trailer bodies are usually made of a minimum of 1-mm thick aluminum-alloy sheets. The skin thickness of a Boeing 747 fuselage, for example, is 1.8 mm; for a Lockheed L1011 it is 1.9 mm. Steel sheets used for appliances bodies are typically about 0.7 mm thick. Aluminum beverage cans are made from sheets 0.28 mm thick, which becomes a cylindrical body with a wall thickness of 0.1 mm after processing into a can (Section 16.7). Aluminum **foil** typically has a thickness of 0.008 mm, although thinner foils, down to 0.003 mm, also can be produced.

13.2 The Flat-rolling Process

A schematic illustration of the *flat-rolling* process is shown in Fig. 13.2a. A metal strip of thickness h_o enters the **roll gap** and is reduced to thickness h_f by a pair of rotating rolls, each powered individually by electric motors. The surface speed of the rolls is V_r. The velocity of the strip increases from its entry value of V_o as it moves through the roll gap, and is highest at the exit from the roll gap, and is denoted as V_f. The metal accelerates in the roll gap, in the same manner as an incompressible fluid flows through a converging channel.

Because the surface speed of the rigid roll is constant, there is *relative sliding* between the roll and the strip along the contact length, L. At one point, called the **neutral point** or **no-slip point**, the velocity of the strip is the same as that of the roll. To the left of this point, the roll moves faster than the strip; to the

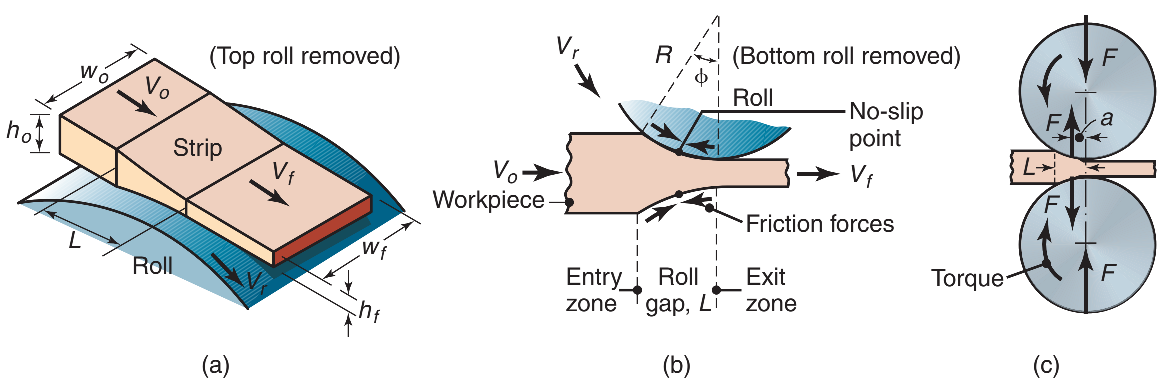


Figure 13.2: (a) Schematic illustration of the flat-rolling process. (b) Friction forces acting on strip surfaces. (c) Roll force, F, and torque, T, acting on the rolls. The width of the strip, w, usually increases during rolling, as shown later in Fig. 13.5.

right, the strip moves faster than the roll. Consequently, the frictional forces act on the strip as shown in Fig. 13.2b.

Forward slip in rolling is defined in terms of the exit velocity of the strip, V_f, and the surface speed of the roll, V_r, as

$$\text{Forward slip} = \frac{V_f - V_r}{V_r}, \tag{13.1}$$

and is a measure of the relative velocities in the roll gap. Forward slip can easily be calculated by measuring the roll and workpiece velocities on a rolling mill, and gives a real-time indication of the neutral point location. Forward slip also correlates with the surface finish of the rolled strip, with low values being preferable to high values.

The rolls pull the material into the roll gap through a net *frictional force* on the material; thus, a net frictional force must exist and be to the right in Fig. 13.2b. This also means that the frictional force to the left of the neutral point must be higher than the friction force to the right. Although friction is essential to enable rolling (just as it is in driving a car on a road), energy is dissipated in overcoming friction. Note that increasing friction also increases rolling forces and power requirements. Furthermore, high friction could damage the surface of the rolled product or cause sticking. A compromise is therefore made in practice through lubricant selection, leading to low and controlled levels of friction.

The maximum possible **draft** is defined as the difference between the initial and final strip thicknesses, or $(h_o - h_f)$. A large draft could cause the rolls to slip. It can be shown that the maximum draft is a function of the roll radius, R, and the coefficient of friction, μ, given by

$$h_o - h_f = \mu^2 R. \tag{13.2}$$

Thus, as expected, the higher the friction and the larger the roll radius, the greater the maximum possible draft. This is a situation similar to the use of large tires (hence high R) and rough treads (hence high μ) on farm tractors and off-road equipment, which allows the vehicles to travel over rough terrain without skidding.

13.2.1 Roll Force, Torque, and Power Requirements

The rolls apply pressure on the flat strip, indicating the presence of a *roll force*, F, as shown in Fig. 13.2c. Note in the figure that this force appears to be perpendicular to the plane of the strip, rather than being at an angle. This is because, in practice, the arc of contact is very small compared with the roll radius, thus it can be assumed that the roll force is approximately perpendicular to the strip.

The roll force in flat rolling can be estimated from the expression

$$F = Lw\sigma_{\text{avg}}, \tag{13.3}$$

where w is the width of the strip, and σ_{avg} is the average true stress (see Section 2.2.3) of the strip in the roll gap. L is the length of contact and can be approximated as the projected length; thus

$$L = \sqrt{R\Delta h}, \tag{13.4}$$

where R is the roll radius and Δh is the difference between the original and final thicknesses of the strip (called **draft**). Equation (13.3) is for a *frictionless* condition; however, an estimate of the *actual roll force*, including friction, may be made by increasing this calculated force by about 20%.

The *torque* on the roll is the product of F and a for frictionless rolling (see Fig. 13.2c). The power required per roll can then be estimated by assuming that F acts in the middle of the arc of contact, or $a \approx L/2$; thus, in Fig. 13.2c, the *total power* (for two rolls), in S.I. units, is

$$\text{Power (kW)} = \frac{2\pi FLN}{60,000}, \tag{13.5}$$

where F is in newtons, L is in meters, and N is the revolutions per minute of the roll.

Example 13.1 Calculation of Roll Force and Torque in Flat-rolling

Given: An annealed copper strip 225 mm wide and 25 mm thick is being rolled to a thickness of 20 mm in one pass. The roll radius is 300 mm, and the rolls rotate at 100 rpm.

Find: Calculate the roll force and the power required in this operation.

Solution: The roll force is determined from Eq. (13.3), in which L is the roll-strip contact length. From Eq. (13.4),

$$L = \sqrt{R(h_o - h_f)} = \sqrt{(300)(25-20)} = 38.7 \text{ mm} = 0.0387 \text{ m}.$$

The average true stress, σ_{avg}, for annealed copper is determined as follows: First note that the absolute value of the true strain that the strip undergoes in this operation is

$$\epsilon = \ln\left(\frac{25}{20}\right) = 0.223.$$

Referring to Fig. 2.5, annealed copper has a true stress of about 80 MPa in the unstrained condition and at a true strain of 0.223, the true stress is around 275 MPa. Hence the average true stress in the roll gap is $(80 + 275)/2 = 178$ MPa. Thus, the roll force is

$$F = Lw\sigma_{\text{avg}} = (0.0387)(0.225)(178 \times 10^6) = 1.55 \text{ MN}.$$

The total power is calculated from Eq. (13.5), with $N = 100$ rpm. Thus,

$$\text{Power} = \frac{2\pi FLN}{60,000} = \frac{2\pi(1.55 \times 10^6)(0.0387)(100)}{60,000} = 628 \text{ kW}.$$

Exact calculation of the force and the power requirements in rolling can be difficult, because of the uncertainties involved in (a) determining the exact contact geometry between the roll and the strip and (b) accurately estimating both the coefficient of friction and the strength of the material in the roll gap. The calculation can be difficult, particularly for hot rolling because of the sensitivity of the strength of the material to temperature and strain rate (see Section 2.2.7).

Reducing Roll Force. Roll forces can cause significant deflection and flattening of the rolls, as it does in a rubber tire. Such changes will, in turn, affect the rolling process and its ability to produce a uniform thickness in the rolled sheet (known as *gage control*, see Section 13.3.2). Also, the columns of the **roll stand** (including the housing, chocks, and bearings, as shown in Fig. 13.3) would deflect under high roll forces to such an extent that the roll gap may open up significantly. Consequently, the rolls have to be set closer than originally calculated in order to compensate for this deflection and to ensure the desired final thickness.

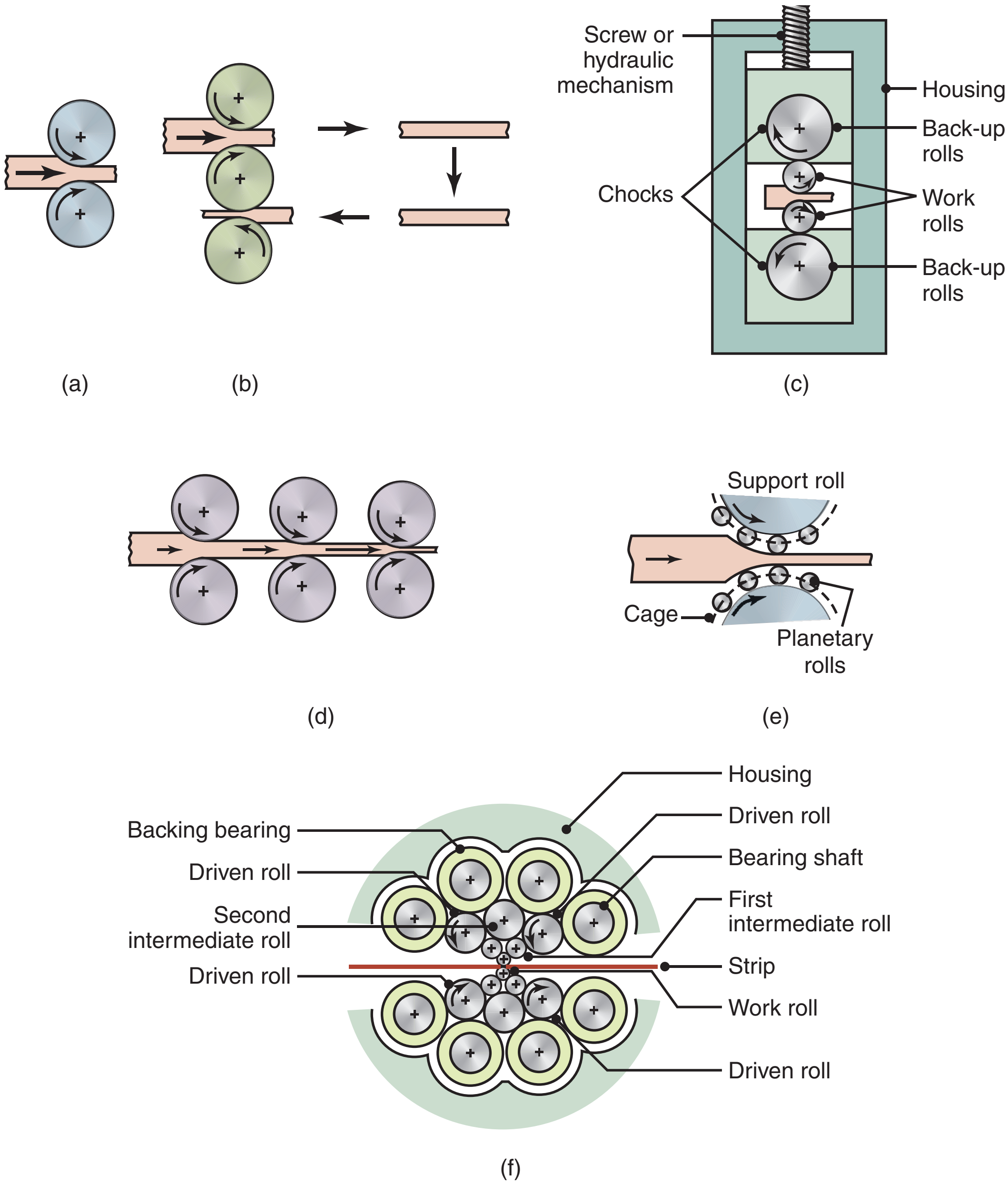


Figure 13.3: Schematic illustration of various roll arrangements: (a) two-high mill; if a two-high mill is used for thick but short workpieces, it will commonly roll a billet back-and-forth in multiple passes, known as a reversing mill; (b) three-high mill with elevator for multiple passes; (c) four-high rolling mill showing various features. The stiffness of the housing, the rolls, and the roll bearings are all important in controlling and maintaining the thickness of the rolled strip; (d) tandem rolling, with three stands; (e) planetary mill; and (f) cluster mill, also known as a *Sendzimir* or Z-mill.

Roll forces can be reduced by the following means:

- Reducing friction at the roll-workpiece interface
- Using smaller diameter rolls, to reduce the contact area
- Taking smaller reductions per pass, to reduce the contact area
- Rolling at elevated temperatures, to lower the strength of the material
- Applying *tensions* to the strip to reduce the roll pressure, as a result of which the compressive stresses required to plastically deform the material become smaller. Because they require high roll forces, tensions are particularly important in rolling of high-strength metals.

Tensions can be applied to the strip at either the entry zone (**back tension**), the exit zone (**front tension**), or both. Back tension is applied to the sheet by a braking action to the reel that supplies the sheet into the roll (*pay-off reel*). Front tension is applied by increasing the rotational speed of the *take-up reel*. Although it has limited and specialized applications, rolling also can be carried out by front tension only, with no power supplied to the rolls, known as **Steckel rolling**.

13.2.2 Geometric Considerations

Just as a straight beam deflects under a transverse load, roll forces tend to *elastically* bend the rolls during rolling, as shown in Fig. 13.4a. The higher the elastic modulus of the roll material, the smaller is the roll deflection. As a result of roll deflections, the rolled strip will now be thicker at its center than at its edges, known as **crown**. A common method of avoiding this problem is to grind the rolls in such a way that their diameter at the center is slightly larger than at their edges (called **camber**). Thus, when the rolls bend, the strip being rolled will have a constant thickness along its width (Fig. 13.4b).

For rolling sheet metals, the radius of the maximum camber is generally 0.25 mm greater than the radius at the ends of the roll. However, a particular camber is correct only for a specific load and strip width. To reduce deflection, the rolls can also be subjected to external bending by applying moments at their bearings.

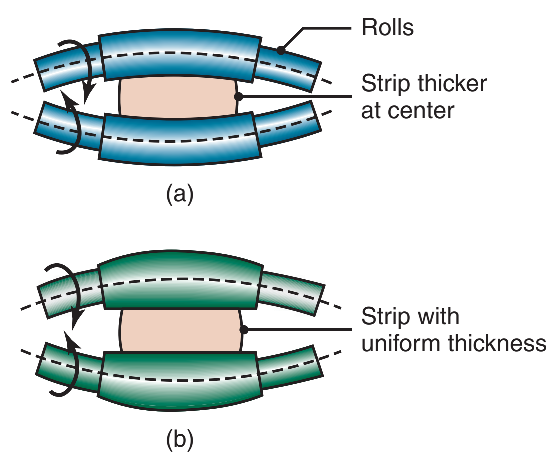


Figure 13.4: (a) Bending of straight cylindrical rolls caused by roll forces. (b) Bending of rolls ground with camber, producing a strip with uniform thickness through the strip width. Deflections have been exaggerated for clarity.

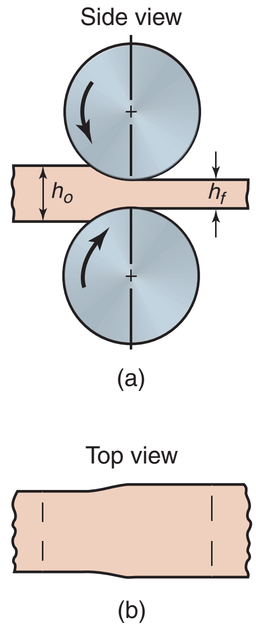


Figure 13.5: Spreading in flat rolling; note that similar spreading can be observed when dough is rolled with a rolling pin.

Because of the heat generated due to the work of plastic deformation during rolling, rolls can become slightly *barrel shaped*, known as **thermal camber**. Unless compensated for by some means, this condition will produce strips that are thinner at the center than at their edges. Thermal camber can be controlled by adjusting the location of coolants and their flow rate along the *length* of the rolls.

Roll forces also tend to *flatten* the rolls *elastically*, producing an effect much like the flattening of automobile tires. Flattening is undesirable because it increases contact area and roll forces. Thin sheets and foils are especially difficult to roll because of roll flattening.

Spreading. The increase in the width of strip during rolling is called *spreading* (Fig. 13.5). In rolling plates and sheets with high width-to-thickness ratios, the width of the strip remains effectively constant during rolling. However, with smaller ratios, such as a bar with a square cross section, its width increases significantly as it passes through the rolls.

It can be shown that spreading increases with (a) decreasing width-to-thickness ratio of the entering strip, (b) increasing friction, and (c) decreasing ratio of roll radius to strip thickness. The last two effects are due to the increased longitudinal constraining force that the material experiences in the roll gap. Spreading can be prevented by using additional rolls, with vertical axes in contact with the edges of the strip; known as *edger mills*, the vertical rolls provide a physical barrier to spreading.

13.2.3 Vibration and Chatter

Chatter is a complex phenomenon (see also Section 25.4), resulting from interactions between the structural dynamics of the mill stand and the dynamics of the rolling operation. Generally defined as *self-excited vibration*, chatter in rolling leads to periodic variations in the thickness of the rolled sheet and in its surface finish, and may lead to excessive scrap (see Table 40.4). Chatter in rolling is found predominantly in tandem mills (Fig. 13.3d). It has been estimated, for example, that modern rolling mills could operate at up to 50% higher speeds were it not for chatter.

Rolling speed and lubrication are found to be the two most significant parameters affecting chatter. Although not always practical to implement, chatter may be reduced by (a) increasing the distance between the stands of the rolling mill, (b) increasing the strip width, (c) decreasing the reduction per pass (draft), (d) increasing the roll radius, (e) increasing the strip-roll friction, and (f) incorporating external dampers in the roll supports.

13.3 Flat-rolling Practice

The initial rolling steps (*breaking down*) of the material is usually done by **hot rolling**, above the recrystallization temperature of the metal (Section 1.7). As described in Section 10.2 and illustrated in Fig. 10.2, a **cast structure** typically is dendritic, consisting of coarse and nonuniform grains, a structure that is usually brittle, and may also be porous. Hot rolling converts the cast structure to a **wrought structure** (Fig. 13.6), with finer grains and enhanced ductility, both of which result from the breaking up of brittle grain boundaries and the closing up of internal defects, including porosity, during rolling. Typical temperature ranges for hot rolling are about 450°C for aluminum alloys, up to 1250°C for alloy steels, and up to 1650°C for refractory alloys (see also Table 14.3).

The rolled product of the first hot-rolling operation is called **bloom**, **slab**, or **billet** (see Fig. 13.1). A bloom typically has a square cross section, at least 150 mm on the side, whereas a slab is usually rectangular in cross section. Blooms are further processed by *shape rolling* into structural shapes, such as I-beams and railroad rails (Section 13.5). Slabs are rolled into plates and sheets. Billets usually are square (with a cross-sectional area smaller than that for blooms), and are later rolled into various shapes, such as round rods and bars, using shaped rolls. Hot-rolled round rods, called **wire rods**, are commonly used as the starting material for rod- and wire-drawing operations (Chapter 15).

In hot rolling of blooms, billets, and slabs, the surface of the material is usually **conditioned** (prepared for a subsequent operation) prior to rolling them. Conditioning is often done by means of a torch (*scarfing*), which removes heavy scale or by rough grinding, which smoothens surfaces. Prior to cold rolling, the scale developed during hot rolling may be removed by *pickling* with acids (acid etching), by such mechanical means as blasting with water or by grinding.

Cold rolling is carried out near room temperature and, compared with hot rolling, it produces sheets and strips with a much better surface finish (because of lack of scale), better dimensional tolerances, and enhanced mechanical properties (because of strain hardening).

Pack rolling is a flat-rolling operation in which two or more layers of sheet are rolled together, thus increasing productivity. *Aluminum foil*, for example, is pack rolled in two layers, where only the top and

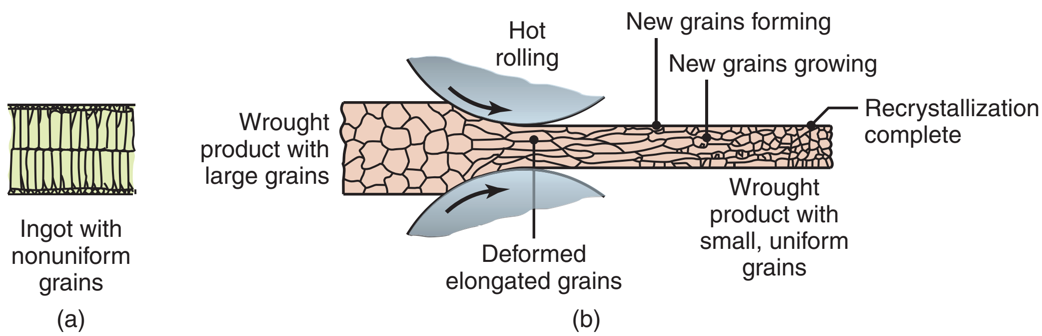


Figure 13.6: Changes in the grain structure of cast or of large-grain wrought metals during hot rolling. Hot rolling is an effective way of reducing grain size in metals for improved strength and ductility. The cast structures of ingots or of continuous castings are converted to a wrought structure by hot working.

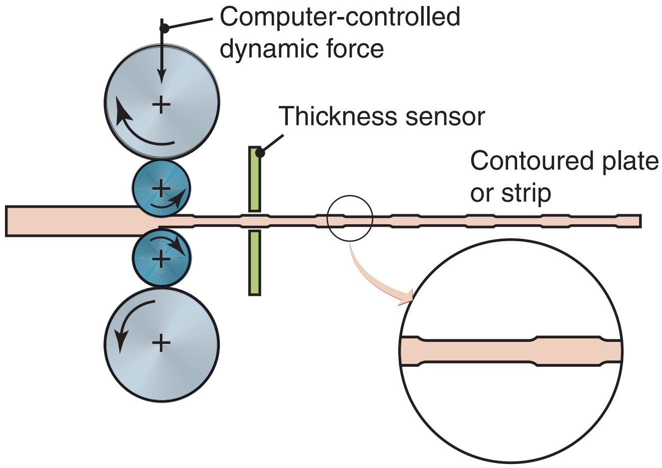


Figure 13.7: Production of tailor rolled blanks, with varying thickness in the rolling direction. A desired contour can be rolled into the workpiece, allowing for optimum placement of material.

bottom outer layers are in contact with the rolls, and hence is smoother. Note that one side of aluminum foil is matte, while the other side is shiny. The foil-to-foil side has a matte and satiny finish, whereas the foil-to-roll side is shiny and bright; this is because it has been in contact with the polished rolls during rolling.

Rolled mild steel, when subsequently stretched during sheet-forming operations, undergoes *yield-point elongation* (Section 16.3), a phenomenon that causes surface irregularities, called *stretcher strains* or *Lüder's bands*. To prevent this situation, the sheet metal is subjected to a final light pass of 0.5–1.5% reduction (known as **temper rolling** or **skin pass**) shortly before stretching it in a subsequent forming operation.

A rolled sheet may not be sufficiently flat as it exits the roll gap, due to factors such as variations in the incoming material or in the processing parameters during rolling. To improve flatness, the rolled strip typically passes through a series of **leveling rolls**. Several roll arrangements can be used, as shown in Fig. 13.8, in which the sheet is basically flexed in opposite directions as it passes through the sets of rolls.

Tailor Rolled Blanks. The thickness of a rolled sheet can be varied by changing the roll forces during the rolling process (Fig. 13.7). Tailor rolled blanks have been used to produce sheet stock as well as tubes. The thickness can be periodic or complex, depending on the dynamic force applied during rolling. Tailor rolled blanks can place material where it is needed for subsequent manufacture or for design purposes. For example, a sheet can be made thicker in locations where sheet-metal forming strains are higher or where stresses are high in service of the part made. **Tailored Coils** are two or more continuously welded coils, intended to provide different materials or thicknesses at different locations across the coil width.

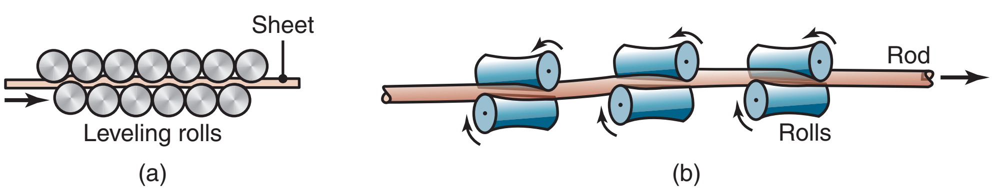


Figure 13.8: (a) A method of roller leveling to flatten rolled sheets. (b) Roller leveling to straighten drawn bars.

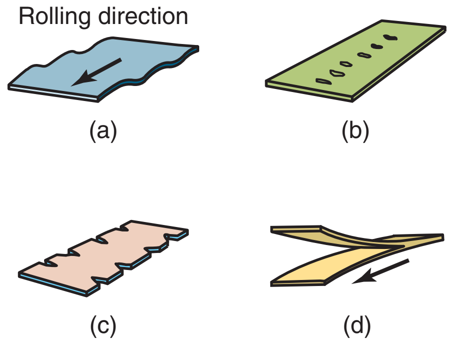


Figure 13.9: Schematic illustration of typical defects in flat rolling: (a) wavy edges; (b) zipper cracks in the center of the strip; (c) edge cracks; and (d) alligatoring.

13.3.1 Defects in Rolled Plates and Sheets

Defects may be present on the surfaces of rolled plates and sheets, or there may be internal structural defects. Defects are undesirable not only because they adversely affect surface appearance, but also because they may affect strength, formability, and other manufacturing characteristics of the rolled sheets. Surface defects such as scale, rust, scratches, gouges, pits, and cracks, may be caused by inclusions and impurities in the original cast material or by various other conditions related to material preparation and to the particular rolling operation itself.

Wavy edges on sheets (Fig. 13.9a) are due to roll bending, whereby the strip becomes thinner along its edges than at its center (see Fig. 13.4a). Because of volume constancy in plastic deformation, the edges then have to elongate more than the material at the center; consequently, the edges buckle because they are constrained by the central region from expanding freely in the longitudinal (rolling) direction.

The **cracks** shown in Fig. 13.9b and c are usually the result of low material ductility at the rolling temperature. Because the quality of the edges of the sheet is important in subsequent forming operations, edge defects in rolled sheets may have to be removed by shearing and slitting operations (Section 16.2). **Alligatoring** (Fig. 13.9d) is typically caused by nonuniform bulk deformation of the billet during rolling or by the presence of defects in the original cast material.

13.3.2 Other Characteristics of Rolled Metals

Residual Stresses. Because of nonuniform deformation of the material within the roll gap, residual stresses can develop in rolled plates and sheets, especially in cold rolling. Small-diameter rolls or small thickness reductions per pass tend to plastically deform the metal to a higher degree at its *surfaces* than in its bulk (Fig. 13.10a). This situation then results in the development of compressive residual stresses on the surfaces and tensile stresses in the bulk. Conversely, large-diameter rolls or high reductions per pass tend to deform the bulk more than its surfaces (Fig. 13.10b). This is due to the higher frictional constraint at the surfaces along the arc of contact.

Dimensional Tolerances. Thickness tolerances for cold-rolled sheets typically range from ± 0.1 to 0.35 mm. Tolerances are much higher for hot-rolled plates, because of thermal effects. *Flatness tolerances* are usually within ± 15 mm/m for cold rolling, and ± 55 mm/m for hot rolling.

Surface Roughness. The ranges of surface roughness in cold and hot rolling are given in Fig. 33.5 which also includes other manufacturing processes for comparison. Note that cold rolling can produce a very fine

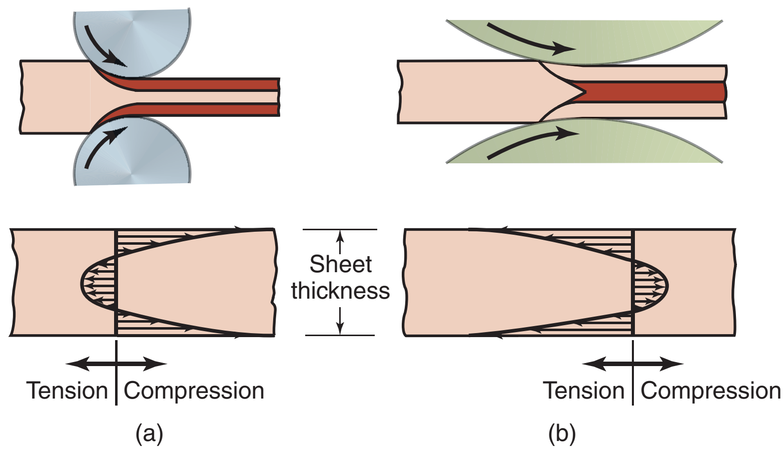


Figure 13.10: (a) Residual stresses developed in rolling with small-diameter rolls or at small reductions in thickness per pass. (b) Residual stresses developed in rolling with large-diameter rolls or at high reductions per pass. Note the reversal of the residual stress patterns.

surface finish; thus, products made of cold-rolled sheets may not require additional finishing operations. Note also in the figure that hot rolling and sand casting produce the same range of surface roughness.

Gage Numbers. The thickness of a sheet is identified by a *gage number*: the smaller the number, the thicker is the sheet. Several numbering systems are used in industry, depending on the type of sheet metal. Rolled sheets of copper and brass are generally identified by thickness changes during rolling, such as 1/4 hard, 1/2 hard, and so on.

13.4 Rolling Mills

Several types of *rolling mills* and equipment are available, with a range of sizes and a variety of roll arrangements. Although the designs of equipment for hot and cold rolling are essentially the same, there are important differences in the roll materials, processing parameters, lubricants, and cooling systems. The design, construction, and operation of rolling mills (Fig. 13.11) require major investments. Highly automated mills now produce close-tolerance, high-quality plates and sheets, at high production rates and at low cost per unit weight, particularly when integrated with continuous casting (Section 5.4). The width of rolled products may range up to 5 m, and rolling speeds are up to 40 m/s.

Two-high rolling mills (Fig. 13.3b) are used for hot rolling in initial breakdown passes (*primary roughing* or **cogging mills**) on cast ingots or in continuous casting, with roll diameters ranging from 0.6 to 1.4 m. In the **three-high mill** (*reversing mill*, Fig. 13.3c) the direction of material movement through the rolls is reversed after each pass, using an elevator mechanism and various manipulators.

Four-high mills (Fig. 13.3a) and **cluster mills** (**Sendzimir** or **Z mill**, Fig. 13.3d) are based on the principle that small-diameter rolls involve lower roll forces because of smaller roll-strip contact area, and thus lower power requirements and reduced spreading. Moreover, when worn or broken, small rolls can easily be replaced and at much lower cost than can large ones. On the other hand, small rolls will deflect more under roll forces, and thus have to be supported by other large-diameter rolls, as is done in four-high and cluster mills. Although the cost of a Sendzimir mill facility is very high, the system is particularly suitable for cold rolling thin sheets of high-strength metals and alloys. Common rolled widths in this mill are 0.66 m, with a maximum of 1.5 m.

In **tandem rolling**, the strip is rolled continuously, through a number of **stands**, to thinner gages with each pass (Fig. 13.12). Each stand consists of a set of rolls, with its own housing and controls. A group of

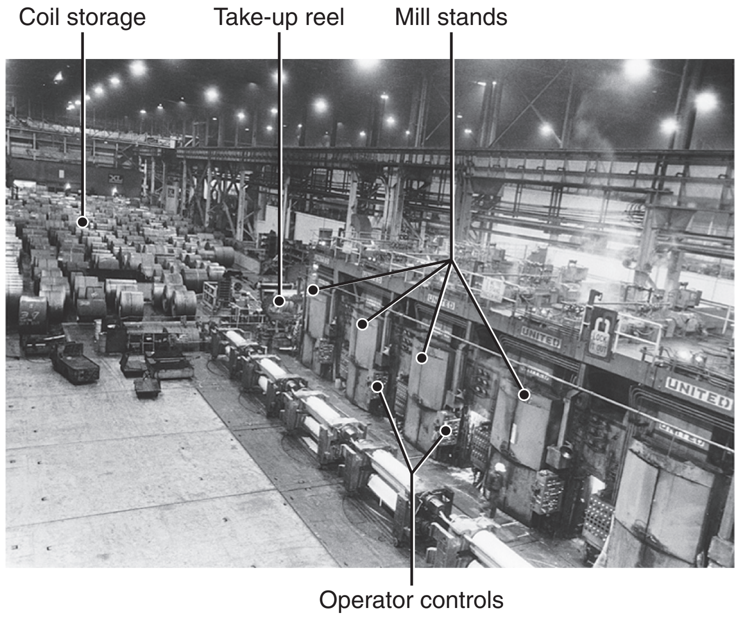


Figure 13.11: View of a rolling mill. *Source: Courtesy of Ispat Inland, an ArcelorMittal legacy company.*

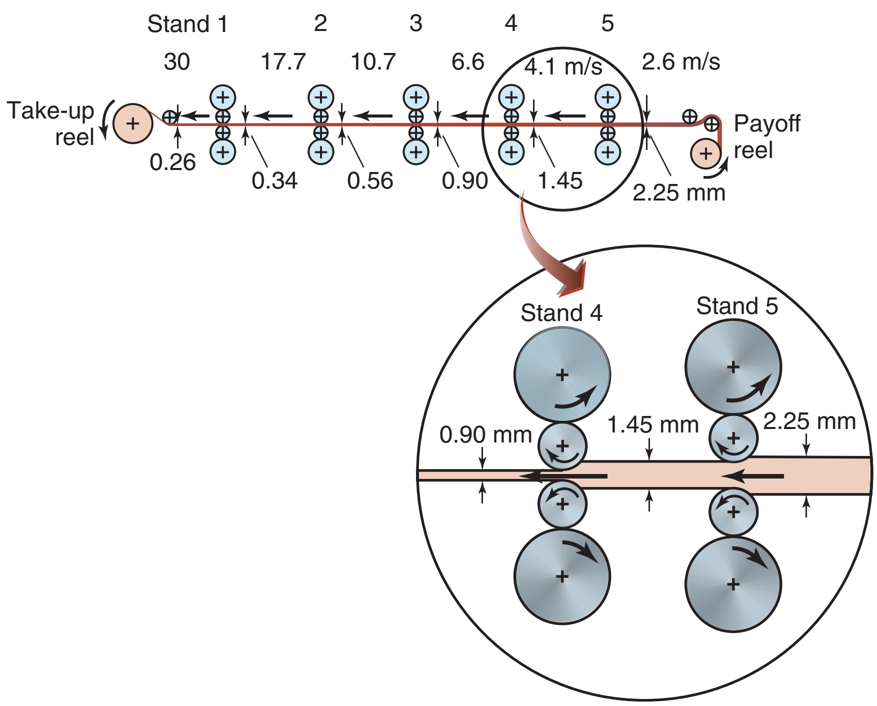


Figure 13.12: An example of a tandem-rolling operation.

stands is called a *train*. The control of the strip thickness and the speed at which the strip travels through each roll gap is critical. Extensive computer controls are used in these operations, particularly in precision rolling at high speeds.

Roll Materials. The basic requirements for roll materials are strength and resistance to wear. Common roll materials are cast iron, cast steel, and forged steel; tungsten carbide is also used for small-diameter rolls, such as the *working roll* in the cluster mill (Fig. 13.3d). Forged-steel rolls, although more costly than cast rolls, have higher strength, stiffness, and toughness than cast-iron rolls. Rolls for cold rolling are ground to a fine finish; for special applications, they are also polished. Rolls made for cold rolling should not be used for hot rolling, because they may crack due to thermal cycling (*heat checking*) or *spall* (cracking or flaking of surface layers).

Lubricants. Hot rolling of ferrous alloys is usually carried out without lubricants, although graphite may be used to reduce friction. Water-based solutions may be used to cool the rolls and to break up the scale on the rolled material. Nonferrous alloys are hot rolled using a variety of compounded oils, emulsions, and fatty acids. Cold rolling is carried out with water-soluble oils or low-viscosity lubricants, such as mineral oils, emulsions, paraffin, and fatty oils (see also Chapter 33).

13.5 Various Rolling Processes and Mills

Several rolling processes and mills have been developed over the years to produce a specific family of product shapes.

Shape Rolling. Straight and long structural shapes, such as channels, I-beams, railroad rails, and solid bars, are formed by *shape rolling* (*profile rolling*), in which the heated stock passes through a set of specially designed rolls (Fig. 13.13; see also Fig. 13.1). *Cold shape rolling* can be done for making rod or wire with various cross sections. Because the entering material's cross section is reduced nonuniformly, the design of

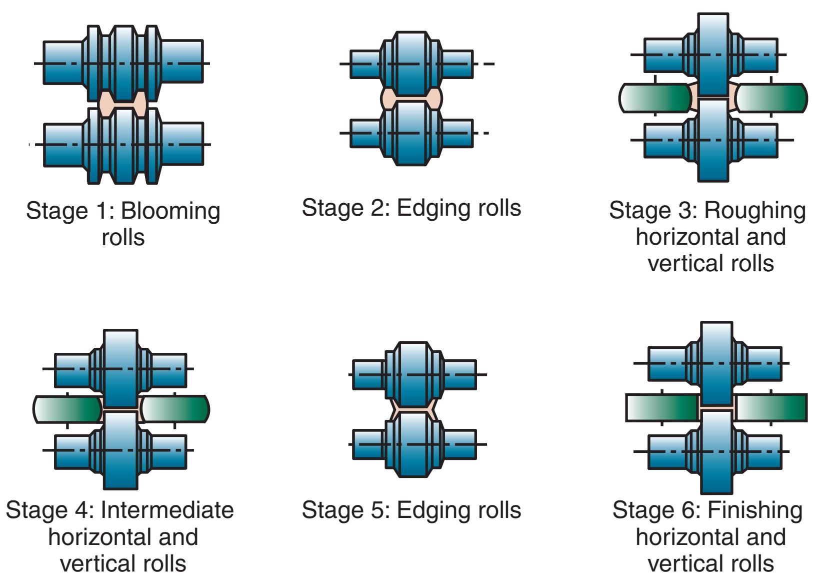


Figure 13.13: Steps in the shape rolling of an I-beam. Various other structural sections, such as channels and rails, also are rolled by this kind of process.

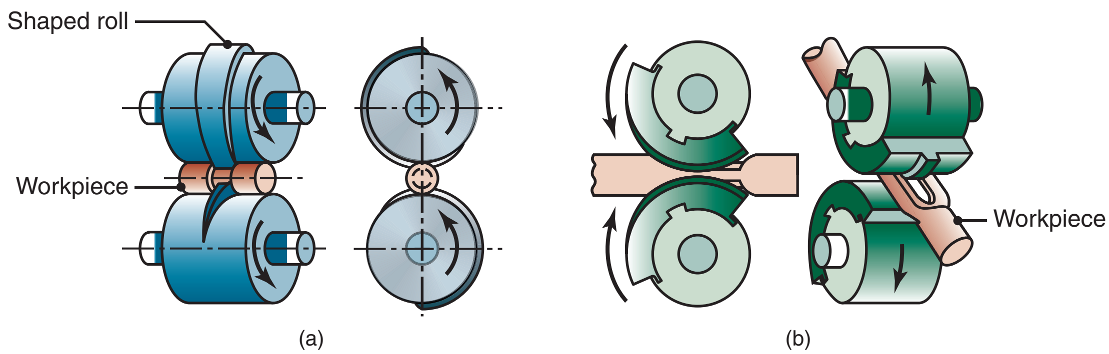


Figure 13.14: Two examples of the roll-forging operation, also known as *cross rolling*. Tapered leaf springs and knives can be made by this process. *Source:* After J. Holub.

a sequence of rolls, called **roll-pass design**, is critical in order to prevent formation of external and internal defects, hold dimensional tolerances, and to reduce roll wear. Also, a coil can use varying thickness or material across the width by using a tailored coil (see Section 13.3).

Roll Forging. In this operation, also called *cross rolling*, the cross section of a round bar is shaped by passing it through a pair of rolls with specially profiled grooves (Fig. 13.14). This process is typically used to produce tapered shafts and leaf springs, table knives, and hand tools. Roll forging may also be used as a preliminary forming operation, to be followed by other forging processes described in Chapter 14.

Skew Rolling. This is a process similar to roll forging and is typically used for making ball bearings (Fig. 13.15a). Round wire or rod is fed into the roll gap, and spherical blanks are formed continuously by the action of the rotating rolls. Another method is illustrated in Fig. 13.15b, which is basically a combined forging and heading operation, described in Fig. 14.12. The balls, which require further finishing, are subsequently ground and polished in special machinery (see Fig. 26.17).

Ring Rolling. In *ring rolling*, a thick ring is expanded into a larger diameter and thinner ring. The ring-shaped blank is placed between two rolls, one of which is driven while the other is idle (Fig. 13.16a). The thickness of the ring is reduced by bringing the rolls closer together as they rotate. Since the volume of the

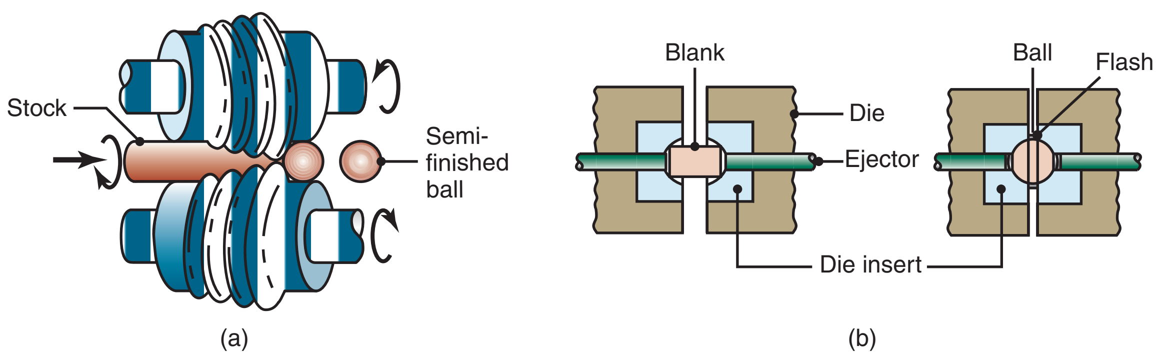


Figure 13.15: (a) Producing steel balls by the skew-rolling process. (b) Producing steel balls by upsetting a cylindrical blank. Note the formation of flash. The balls made by these processes are subsequently ground and polished for use in ball bearings.

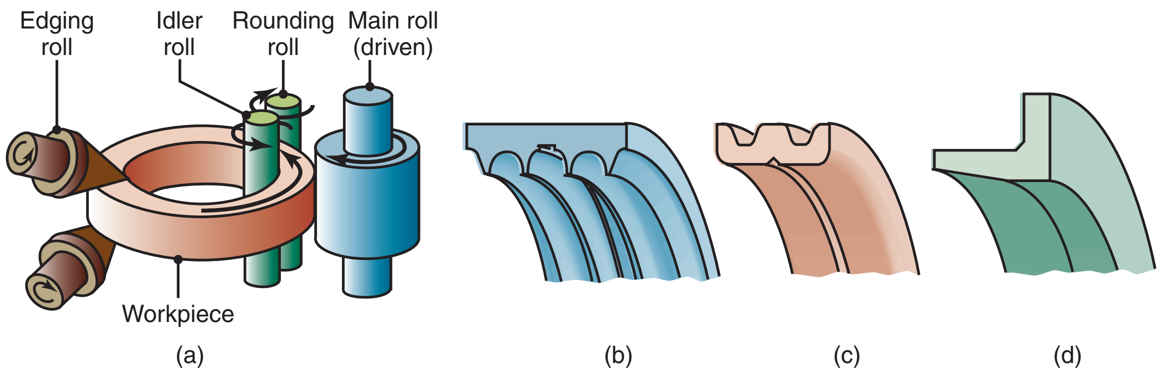


Figure 13.16: (a) Schematic illustration of a ring-rolling operation. Thickness reduction results in an increase in the part diameter. (b) through (d) Examples of cross sections that can be formed by ring rolling.

ring remains constant during deformation (volume constancy), the reduction in ring thickness results in its increase in diameter. The process can be carried out either at room or elevated temperature. The ring size can be up to 3 m in diameter.

The blank may be produced by such means as cutting from a plate, piercing, shearing a thick-walled pipe. Various cross sections can be ring rolled using shaped rolls (Fig. 13.16). The thickness of rings also can be reduced by an open-die forging process, as illustrated in Fig. 14.4c; however, dimensional control and surface finish will not be as good as in ring rolling.

Typical applications of ring rolling are large rings for rockets and turbines, jet-engine cases, ball-bearing and roller-bearing races, flanges, and reinforcing rings for pipes. Compared with other manufacturing processes that are capable of producing the same part, the advantages of ring rolling are short production times, material savings, close dimensional tolerances, and favorable grain flow in the product, thus enhancing its strength in the desired direction.

Thread Rolling. *Thread rolling* is a cold-forming process by which straight or tapered threads are formed on round rods. The threads are formed with each stroke of a pair of flat reciprocating dies (Fig. 13.17a). In another method, threads are formed by using two rolls (Fig. 13.17b) or *rotary* or *planetary dies* (Fig. 13.17c), at production rates as high as 80 pieces per second. Typical parts made are screws, bolts, and threaded parts. Depending on die design, the major diameter of a rolled thread may or may not be larger than a machined thread (Fig. 13.18a), that is, the same as the blank diameter.

The thread-rolling process has the advantages of generating threads with good strength (due to cold working) and without any scrap. The surface finish produced is very smooth, and the process induces compressive residual stresses on the surfaces, thus improving fatigue life. The process is superior to other methods of thread manufacturing, notably thread cutting, as illustrated in Fig. 23.1k. Machining the threads cuts through the grain-flow lines of the material, whereas rolling the threads results in a grain-flow pattern that improves thread strength (Fig. 13.18).

Spur and helical gears can be produced by a cold-rolling process similar to thread rolling (see also Section 24.7). The operation may be carried out on solid cylindrical blanks or on precut gears. Cold rolling of gears has extensive applications in automatic transmissions and in power tools. **Internal thread rolling** can be carried out with a fluteless **forming tap** (Section 23.7), an operation that is similar to external thread rolling; it produces accurate internal threads with good strength.

Lubrication is important in thread-rolling operations, in order to obtain a good surface finish and surface integrity, and minimize defects. Lubrication affects the manner in which the material deforms during

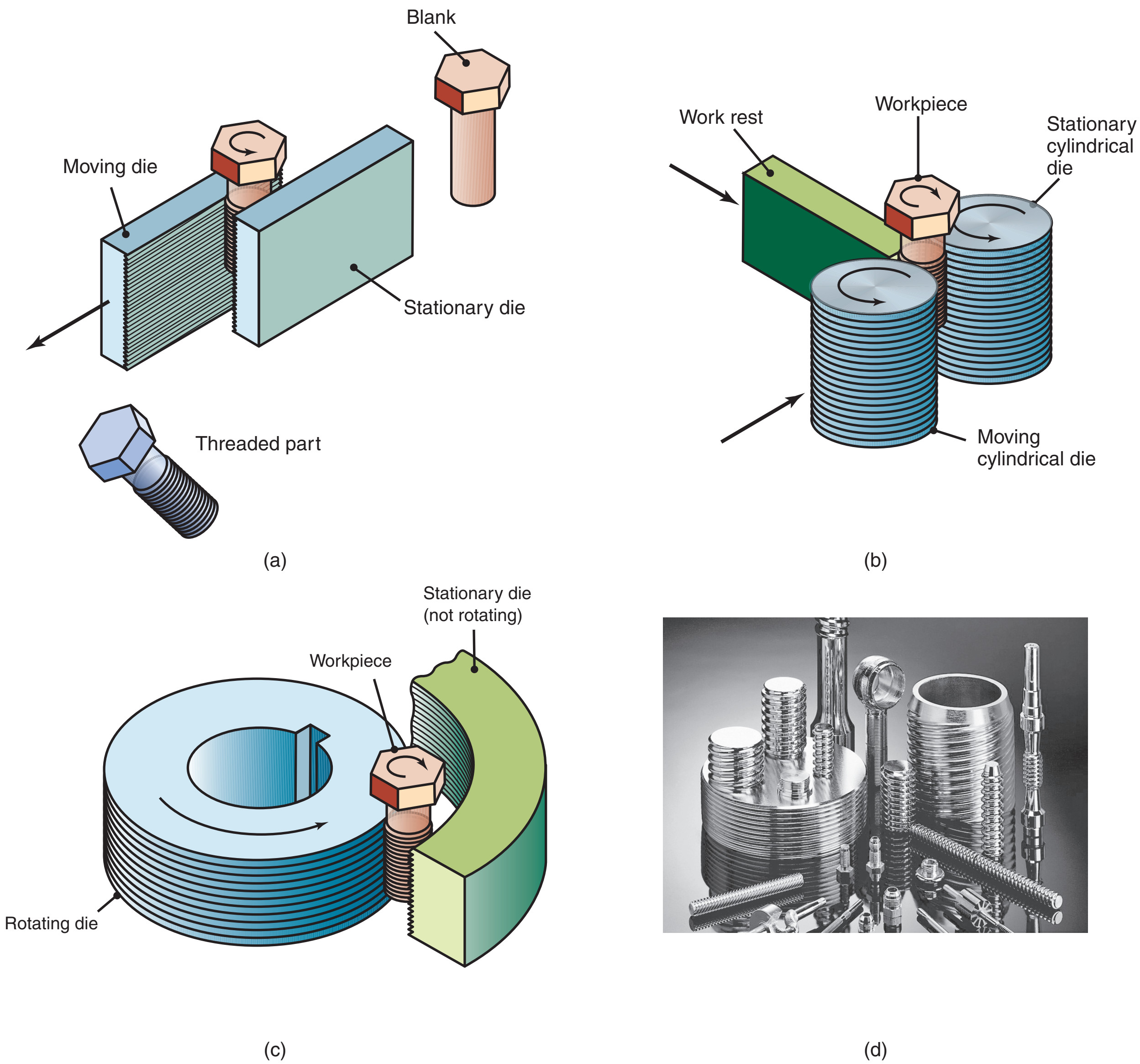


Figure 13.17: Thread-rolling processes: (a) reciprocating flat dies used to produce a threaded fastener; (b) two-roll dies; (c) rotary or planetary die set; (d) A collection of thread-rolled parts made economically at high production rates. *Source:* Courtesy of Tesker Manufacturing Corp.

processing, an important consideration because of the possibility of internal defects being developed (see, for example, Fig. 14.17). Typically made of hardened steel, rolling dies are expensive because of their complex shape, and usually cannot be reground after they are worn. With proper selection of die materials and preparation, die life may range up to millions of pieces.

Rotary Tube Piercing. Also known as the **Mannesmann process**, this is a hot-working operation for producing long, thick-walled *seamless pipe and tubing* (Fig. 13.19). Developed in the 1880s, this process is based

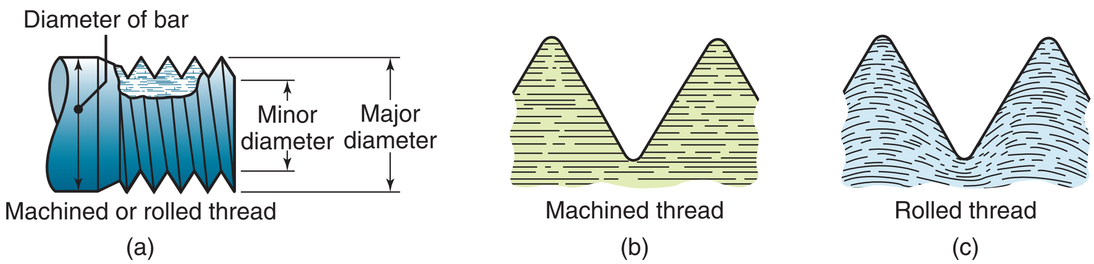


Figure 13.18: (a) Features of a machined or rolled thread. Grain flow in (b) machined and (c) rolled threads. Unlike machining, which cuts through the grains of the metal, the rolling of threads imparts improved strength because of cold working and favorable grain flow.

on the principle that when a round bar is subjected to radial compressive forces, tensile stresses develop at its center (see Fig. 2.9). When continuously subjected to these cyclic compressive stresses (Fig. 13.19b), the bar begins to first develop a small cavity at its center, which then begins to grow. This phenomenon can be demonstrated with a short piece of round eraser, by rolling it back and forth on a hard flat surface, as shown in Fig. 13.19b.

Rotary tube piercing is carried out using an arrangement of rotating rolls (Fig. 13.19c). The axes of the rolls are *skewed* in order to pull the round bar through the rolls by the axial component of the rotary motion. An internal mandrel assists the operation by expanding the hole and *sizing* the inside diameter of the tube. The mandrel may be held in place by a long rod or it may be a floating mandrel, without a support (see Fig. 15.21c for a similar floating mandrel used in drawing). Because of the severe deformation that the bar undergoes, the blank must be of high quality and free of inclusions.

Tube Rolling. The diameter and thickness of pipes and tubing can be reduced by *tube rolling*, which utilizes shaped rolls arranged in various configurations (Fig. 13.20). These operations can be carried out with or without an internal mandrel. In the *pilger mill*, the tube and an internal mandrel undergo a reciprocating motion; the rolls are rotated continuously. During the gap cycle on the roll, the tube is advanced and rotated, starting another cycle of tube reduction, whereby the tube undergoes a reduction in both its diameter and its wall thickness. Steel tubing 265 mm in diameter has been produced by this process. Other operations for tube manufacturing are described in Chapter 15.

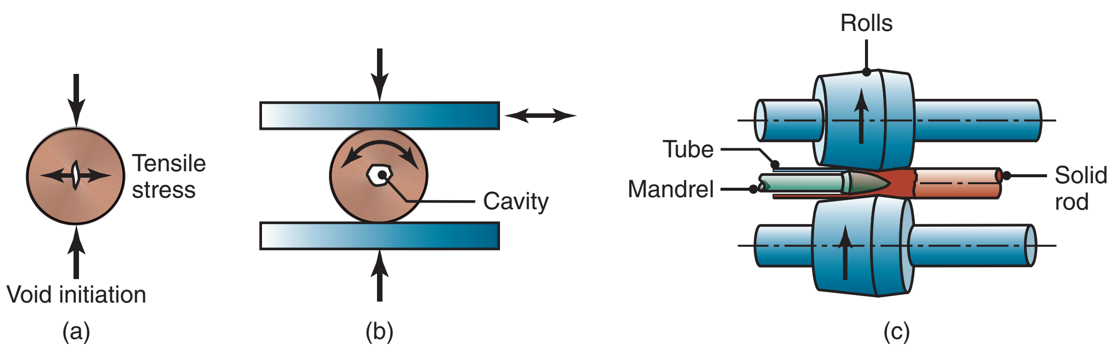


Figure 13.19: Cavity formation in a solid, round bar and its utilization in the rotary tube-piercing process for making seamless pipe and tubing (see also Fig. 2.9).

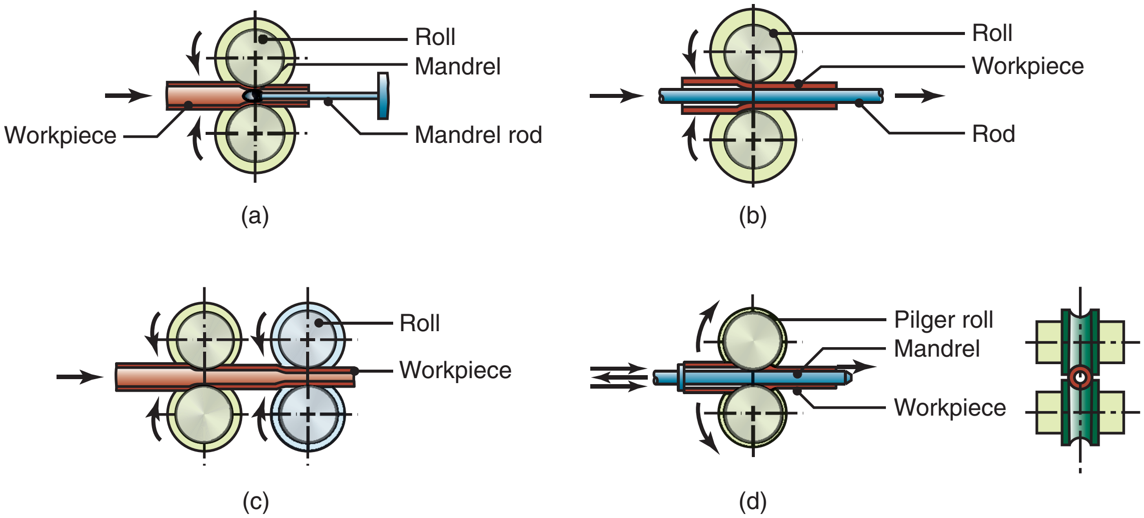


Figure 13.20: Schematic illustration of various tube-rolling processes: (a) with a fixed mandrel; (b) with a floating mandrel; (c) without a mandrel; and (d) pilger rolling over a mandrel and a pair of shaped rolls. Tube diameters and thicknesses also can be changed by other processes, such as drawing, extrusion, and spinning.

13.5.1 Integrated Mills and Minimills

Integrated Mills. These mills are large facilities that involve complete integration of all activities, from the production of hot metal in a blast furnace to the casting and rolling of finished products, ready to be shipped to the customer.

Minimills. In *minimills*, scrap metal is (a) melted in electric-arc furnaces, (b) cast continuously, and (c) rolled directly into specific lines of products. Each minimill produces essentially one type of rolled product (rod, bar, or structural sections such as angle iron), from basically one type of metal or alloy. The scrap metal, obtained locally to reduce transportation costs, is typically old machinery, cars, and farm equipment. Minimills have the economic advantage of lower capital equipment costs for each type of metal and product line, with low labor and energy costs. The products typically are aimed at markets in the mill's particular geographic location.

Summary

- Rolling is the process of reducing the thickness or changing the cross section of a long strip by compressive forces applied through a set of rolls. Shape rolling is used to make products with various cross sections. Other rolling operations include ring rolling and thread rolling.

- The process may be carried out at room temperature (cold rolling) or at elevated temperatures (hot rolling). Rolling involves several material and process variables, including roll diameter (relative to material thickness), reduction per pass, speed, lubrication, and temperature. Spreading, bending, and flattening are important considerations for controlling the dimensional accuracy of the rolled stock.

- Rolling mills have a variety of roll configurations, such as two-high, three-high, four-high, cluster (Sendzimir), and tandem. Front and/or back tension may be applied to the material to reduce roll forces.

- Continuous casting and rolling of ferrous and of nonferrous metals into semi-finished products is a common practice because of the economic benefits.

- Integrated mills are large facilities involving the total sequence of activities, from the production of hot metal in a blast furnace to the casting and the rolling of finished products and ready to be shipped to the customer. On a much smaller scale, minimills utilize scrap metal that is melted in electric-arc furnaces, cast, and continuously rolled into specific lines of products.

Key Terms

Alligatoring
Back tension
Billet
Bloom
Camber
Cast structure
Chatter
Cogging mill
Cold rolling
Crown
Draft
Flat rolling
Foil
Front tension
Gage number
Hot rolling
Mannesmann process
Minimill
Neutral point
Pack rolling
Pilger mill
Plate
Ring rolling
Roll
Roll forging
Roll stand
Rolling
Rolling mill
Rotary tube piercing
Sendzimir mill
Shape rolling
Sheet
Skew rolling
Slab
Spreading
Stand
Steckel rolling
Tandem rolling
Temper rolling
Tailor welded coil
Thread rolling
Tube rolling
Wrought structure

Bibliography

Ginzburg, V.B., and Ballas, R., **Flat Rolling Fundamentals**, CRC Press, 2000.
Hosford, W.F., and Caddell, R.M., **Metal Forming: Mechanics and Metallurgy**, 4th ed., Cambridge University Press, 2014.
Lee, Y., **Rod and Bar Rolling: Theory and Applications**, CRC Press, 2004.
Lenard, J.G., **Primer on Flat Rolling**, Elsevier, 2007.
Pittner, J., and Simaan, M.A., **Tandem Cold Metal Rolling Mill Control: Using Practical Advanced Methods**, Springer, 2010.

Review Questions

13.1. What is the difference between a plate and a sheet?

13.2. Define (a) *roll gap*, (b) *neutral point*, and (c) *draft*.

13.3. What factors contribute to spreading in flat rolling?

13.4. What is forward slip? Why is it important?

13.5. Explain the types of deflections that rolls undergo.

13.6. Describe the difference between a bloom, a slab, and a billet.

13.7. Why may roller leveling be a necessary operation?

13.8. List the defects commonly observed in flat rolling.

13.9. What are the advantages of tandem rolling? Pack rolling?

13.10. How are seamless tubes produced?

13.11. Why is the surface finish of a rolled product better in cold rolling than in hot rolling?

13.12. What is a Sendzimir mill? What are its important features?

13.13. What is the Mannesmann process? How is it different from tube rolling?

13.14. Describe ring rolling. Is there a neutral plane in ring rolling?

13.15. How is back tension generated?

Qualitative Problems

13.16. Explain why the rolling process was invented and developed.

13.17. Flat rolling reduces the thickness of plates and sheets. It is possible, instead, to reduce their thickness simply by stretching the material? Would this be a feasible process? Explain.

13.18. Explain how the residual stress patterns shown in Fig. 13.10 become reversed when the roll radius or reduction-per-pass is changed.

13.19. Explain whether it would be practical to apply the roller-leveling technique shown in Fig. 13.8a to thick plates.

13.20. Describe the factors that influence the magnitude of the roll force, F, in Fig. 13.2c.

13.21. Explain how you would go about applying front and back tensions to sheet metals during rolling. How would you go about controlling these tensions?

13.22. What typically is done to make sure that the product in flat rolling is not crowned?

13.23. Make a list of some parts that can be made by (a) shape rolling and (b) thread rolling.

13.24. Describe the methods by which roll flattening can be reduced. Which property or properties of the roll material can be increased to reduce roll flattening?

13.25. In the chapter, it was stated that spreading in flat rolling increases with (a) a decreasing width-to-thickness ratio of the entering material, (b) decreasing friction, and (c) a decreasing ratio of the roll radius to the strip thickness. Explain why.

13.26. As stated in this chapter, flat rolling can be carried out by front tension only, using idling rolls (Steckel rolling). Since the torque on the rolls is now zero, where, then, is the energy coming from to supply the work of deformation in rolling?

13.27. Explain the consequence of applying too high a back tension in rolling.

13.28. Note in Fig. 13.3d that the driven rolls (powered rolls) are the third set from the work roll. Why isn't power supplied through the work roll itself? Is it even possible? Explain.

13.29. Describe the importance of controlling roll speeds, roll gaps, temperature, and other process variables in a tandem-rolling operation, as shown in Fig. 13.12. Explain how you would go about determining the distance between the stands.

13.30. In Fig. 13.10a, if you remove the top compressive layer by, say, grinding, will the strip remain flat? If not, which way will it curve and why?

13.31. Name several products that can be made by each of the operations shown in Fig. 13.1.

13.32. List the possible consequences of rolling at (a) too high of a speed and (b) too low of a speed.

13.33. It is known that in thread rolling as illustrated in Fig. 13.17, a workpiece must make roughly six revolutions to form the thread. Under what conditions (process parameters, thread geometry or workpiece properties) can deviation from this rule take place?

13.34. If a rolling mill encounters chatter, what process parameters would you change, and in what order? Explain your answer.

13.35. Can the forward slip ever become negative? Why or why not?

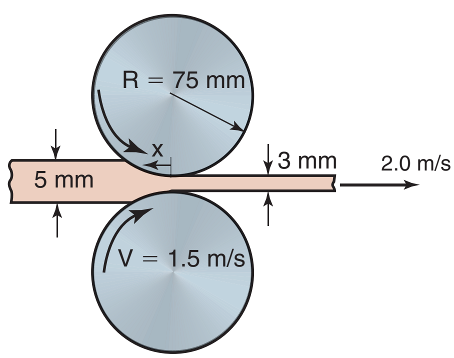
Quantitative Problems

13.36. In Example 13.1, calculate the roll force and the power for the case in which the workpiece material is 1100-O aluminum and the roll radius, R, is 500 mm.

13.37. Calculate the individual drafts in each of the stands in the tandem-rolling operation shown in Fig. 13.12.

13.38. Estimate the roll force, F, and the torque for an AISI 1020 carbon-steel strip that is 200 mm wide, 12 mm thick, and rolled to a thickness of 6 mm. The roll radius is 200 mm, and it rotates at 200 rpm.

13.39. A rolling operation takes place under the conditions shown in the accompanying figure. What is the position, x_n, of the neutral point? Note that there are a front and back tension that have not been specified. Additional data are as follows: Material is 5052-O aluminum; hardened steel rolls; surface roughness of the rolls = 0.025 μm; rolling temperature = 210°C.



Sketch P13.39

13.40. Estimate the roll force and power for annealed low carbon steel strip 200 mm wide and 10 mm thick, rolled to a thickness of 6 mm. The roll radius is 200 mm, and the roll rotates at 200 rpm. Use $\mu = 0.2$.

13.41. A flat-rolling operation is being carried out where $h_o = 6$ mm, $h_f = 5$ mm, $w_o = 300$ mm, $R = 250$ mm, $\mu = 0.25$, and the average flow stress of the material is 275 MPa. Estimate the roll force and the torque.

13.42. It can be shown that it is possible to determine μ in flat rolling without measuring torque or forces. By inspecting equations for rolling, describe an experimental procedure to do so. Note that you are allowed to measure any quantity other than torque or forces.

13.43. A U-channel of 85-15 brass will be shape formed, but first must be flat rolled to a thickness of 0.9 mm. The strip has a width of 25 mm and an initial thickness of $h_o = 1.5$ mm. A preliminary process design suggests a 40% reduction in a single pass on a rolling mill with 150 mm-radius rolls. If the roll surface speed is 1 m/s, and the coefficient of friction is $\mu = 0.1$, calculate the rolling force and power requirements. Repeat the problem if two passes were taken to achieve the desired reduction.

13.44. Assume that you are an instructor covering the topics described in this chapter and you are giving a quiz on the numerical aspects to test the understanding of the students. Prepare two quantitative problems and supply the answers.

Synthesis, Design, and Projects

13.45. A simple sketch of a four-high mill stand is shown in Fig. 13.3c. Make a survey of the technical literature and present a more detailed sketch for such a stand, showing the major components.

13.46. Obtain a piece of soft, round rubber eraser, such as that at the end of a pencil, and duplicate the process shown in Fig. 13.19b. Note how the central portion of the eraser will begin to erode, producing a hole.

13.47. If you repeat the experiment in Problem 13.46 with a harder eraser, such as that used for erasing ink, you will note that the whole eraser will begin to crack and crumble. Explain why.

13.48. Design a set of rolls to produce cross sections other than those shown in Fig. 13.13.

13.49. Design an experimental procedure for determining the neutral point in a flat-rolling operation.

13.50. Using a rolling pin and any available dough (bread, cookie, etc.), measuring 100 by 100 by 8 mm, quantify the spreading in flat rolling for different reductions in thickness.

13.51. Derive an expression for the thickest workpiece that can be drawn between two rolls as a function of roll gap, roll radius and coefficient of friction.

13.52. Make an extensive list of products that could use the benefits of tailor welded coils, and for each product, explain why.

Chapter 14

Metal-forging Processes and Equipment

- This chapter describes the fundamentals of forging and related processes, including design and economic considerations.
- Open-die forging operations for producing simple shapes are described first, followed by impression-die and closed-die forging operations for producing more intricate shapes.
- Various forging operations, such as heading, piercing, coining, swaging, and cold extrusion, are then introduced.
- Factors involved in forging defects and die failures are explained.
- The economics of forging, as it relates to process selection, is introduced.
- The chapter ends with a review of forging design considerations, guidelines for die design and manufacturing, and selection of die materials and lubricants in forging operations.

Typical parts made by forging and related processes: Shafts, gears, bolts, turbine blades, hand tools, dies, and a wide variety of components for machinery, transportation, and farm equipment.

Alternative processes: Casting, powder metallurgy, machining, additive manufacturing, and fabrication.

14.1 Introduction

Forging is a basic process in which the workpiece is shaped by compressive forces applied through various dies and tooling. One of the oldest and most important metalworking operations, dating back at least to 4000 B.C., forging was first used to make jewelry, coins, and various implements, by hammering metal with tools made of stone. Forged parts now include large rotors for turbines, gears, cutlery (Fig. 14.1), hand tools, and miscellaneous components for machinery, aircraft, and transportation equipment.

Unlike rolling operations (Chapter 13) that generally produce continuous plates, sheets, strips, and various structural cross sections, forging operations produce *discrete parts*. Because the metal flow in a die and the material's grain structure can be controlled, forged parts have good strength and toughness, and are very reliable for highly stressed and critical applications (Fig. 14.2). Simple forging operations can be performed with a heavy hammer and an anvil, as has been done traditionally by blacksmiths for centuries. Most forgings require a set of dies and such equipment as presses or powered hammers.

Forging may be carried out at room temperature (*cold forging*) or at elevated temperatures (*warm* or *hot forging*), depending on the homologous temperature, described in Section 1.8. Cold forging requires higher forces, because of the higher strength of the workpiece material. The workpiece material must possess sufficient ductility at room temperature to be able to undergo the required deformation without cracking. Cold-forged parts have good surface finish and dimensional accuracy. Hot forging requires lower forces, but the dimensional accuracy and surface finish of the parts are not as good as those in cold forging.

Forgings generally are subjected to subsequent finishing operations, such as heat treating to modify properties, and machining for dimensional accuracy and good surface finish. The finishing operations can be minimized by *precision forging*, an important example of *net-shape* or *near-net-shape* forming.

As described throughout this book, parts that can be forged successfully also may be manufactured economically by other methods, such as casting (Chapter 11), powder metallurgy (Chapter 17), additive

(a) (b)

Figure 14.1: (a) Illustration of the steps involved in forging a knife. (b) Open die forging of a steel billet. *Source:* (a) Courtesy of Mundial, Inc. (b) Shutterstock/Milos Zvicer.

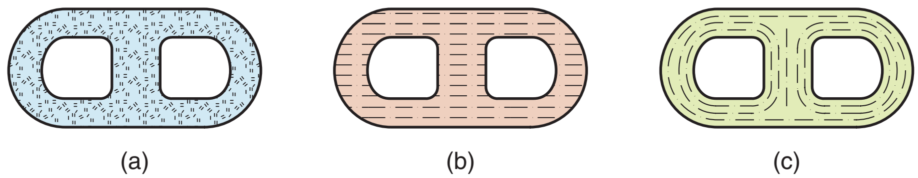


Figure 14.2: Schematic illustration of a part (dragline chain link, approximately 2-m long) made by three different processes and showing grain flow. (a) Casting by the processes described in Chapter 11. (b) Machining from a blank, described in Part IV of this book, and (c) forging. Each process has its own advantages and limitations regarding external and internal characteristics, material properties, dimensional accuracy, surface finish, and the economics of production. *Source:* Courtesy of the Forging Industry Association.

manufacturing (Chapter 20), or machining (Part IV). Each of these will produce a part having different characteristics, particularly with regard to strength, toughness, dimensional accuracy, surface finish, and the possibility of internal or external defects.

14.2 Open-die Forging

Open-die forging is the simplest forging operation (Table 14.1). Although most open-die forgings generally weigh 15 to 500 kg, forgings as heavy as 270 metric tons have been made. Part sizes may range from very small (such as pins, nails, and screws) to very large [up to 23-m-long shafts for ship propellers]. In its simplest form, open-die forging can be described by a metal workpiece *blank*, placed between two flat dies (*platens*), and reduced in height by compressing it (Fig. 14.3), an operation that is also called **upsetting** or **flat-die forging**. The die surfaces may have shallow cavities or features to make relatively simple forgings (see also blocker dies, Section 14.3).

The deformation of a solid cylindrical workpiece under *frictionless* conditions is shown in Fig. 14.3b. Because constancy of volume is to be maintained, any reduction in height increases the diameter of the forged part. Note that the workpiece is deformed *uniformly*. In an actual operation, however, there is friction at the die–workpiece interfaces, whereby the part develops a *barrel* shape (Fig. 14.3c), a deformation mode also called *pancaking*.

Table 14.1: General Characteristics of Forging Processes.

Process	Advantages	Limitations
Open die	Simple and inexpensive dies; wide range of part sizes; good strength characteristics; generally for small quantities	Limited to simple shapes; difficult to hold close tolerances; machining to final shape necessary; low production rate; relatively poor utilization of material; high degree of skill required
Closed die	Relatively good utilization of material; generally better properties than open-die forgings; good dimensional accuracy; high production rates; good reproducibility	High die cost, not economical for small quantities; machining often necessary
Blocker	Low die costs; high production rates	Machining to final shape necessary; parts with thick webs and large fillets
Conventional	Requires much less machining than blocker type; high production rates; good utilization of material	Higher die cost than blocker type
Precision	Close dimensional tolerances; very thin webs and flanges possible; machining generally not necessary; very good material utilization	High forging forces, intricate dies, and provision for removing forging from dies

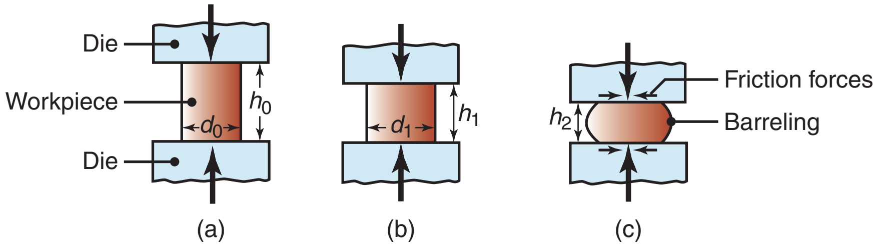


Figure 14.3: (a) Solid cylindrical billet upset between two flat dies. (b) Uniform deformation of the billet without friction. (c) Deformation with friction. Note barreling of the billet caused by friction forces at the billet–die interfaces.

Barreling is caused primarily by frictional forces that oppose the outward flow of the workpiece at the die interfaces, thus it can be minimized by using an effective lubricant. Barreling also can develop in upsetting hot workpieces between cold dies; the material at the die surfaces cools rapidly, while the bulk remains relatively hot. Consequently, the material at the top and bottom of the workpiece has higher resistance to deformation than the material at the center. As a result, the central portion of the workpiece expands laterally to a greater extent than do the ends. Barreling from thermal effects can be reduced or eliminated by using heated dies. Thermal barriers, such as glass cloth placed at the two die–workpiece interfaces also can be used for this purpose.

Cogging, also called *drawing out*, is basically an open-die forging operation in which the thickness of a bar is reduced by successive forging steps (*bites*) at specific intervals (Fig. 14.4a). The thickness of bars and rings can be reduced also by similar open-die forging techniques, as illustrated in Fig. 14.4b and c. Because the contact area between the die and the workpiece is now smaller, a long section of a bar can be reduced in thickness without requiring large forces or heavy machinery. Note that blacksmiths have been performed such an operation for centuries, using a hammer, an anvil, and a periodically heated workpiece. Cogging of larger workpieces is usually done on mechanized equipment and with computer controls. The lateral and the vertical movements of the dies are coordinated to produce the desired shape.

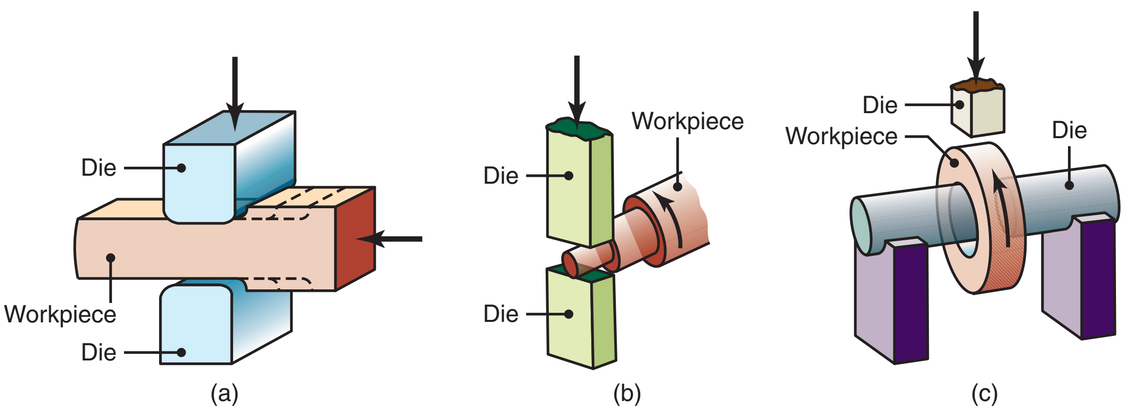


Figure 14.4: (a) Schematic illustration of a cogging operation on a rectangular bar. Blacksmiths use this process to reduce the thickness of bars by hammering the part on an anvil. Reduction in thickness is accompanied by barreling, as in Fig. 14.3c. (b) Reducing the diameter of a bar by open-die forging; note the movements of the dies and the workpiece. (c) The thickness of a ring being reduced by open-die forging.

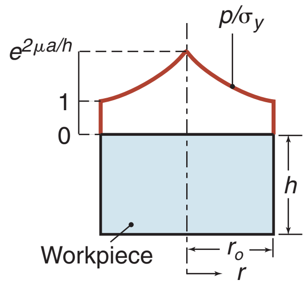


Figure 14.5: Distribution of die pressure in upsetting with sliding friction. Note that the pressure at the outer radius is equal to the flow stress, σ_f, of the material. Sliding friction means that the frictional stress is directly proportional to the normal stress.

Forging Force. The *forging force*, F, in an *open-die forging operation* on a solid cylindrical workpiece can be estimated from the formula

$$F = \sigma_f \pi r^2 \left(1 + \frac{2\mu r}{3h}\right), \tag{14.1}$$

where σ_f is the *flow stress* of the material (see Example 14.1), μ is the coefficient of friction between the workpiece and the die, and r and h are the instantaneous radius and height of the workpiece, respectively.

Friction Hill. Consider the upsetting of a cylinder, as depicted in Fig. 14.3. If the workpiece-die interfaces are frictionless, then the die pressure is the flow stress of the material. If friction is present, as is the case in actual operations, then the die pressure is calculated as follows. For upsetting of a cylinder with outer radius r_o, height, h, and coefficient of friction, μ, the die pressure at any radius can be expressed as

$$p = \sigma_f e^{2\mu(r_o - r)/h}. \tag{14.2}$$

The die pressure distribution is shown in Fig. 14.5. Note that the pressure it is at a maximum at the center of the workpiece, and can be very high especially if the diameter-to-height ratio of the workpiece is high. Because of its shape, the pressure-distribution curve in Fig. 14.5 is referred to as the *friction hill*.

Example 14.1 Calculation of Forging Force in Upsetting

Given: A solid cylindrical workpiece made of 304 stainless steel is 150 mm in diameter and 100 mm in height. It is reduced in height by 50%, at room temperature, in an open-die forging operation with flat dies. Assume that the coefficient of friction is 0.2.

Find: What is the forging force at the end of the stroke?

Solution: The forging force at the end of the stroke is calculated using Eq. (14.1), in which the dimensions pertain to the final dimensions of the forging. The final height is $h = 100/2 = 50$ mm, and the final radius, r, is determined from volume constancy, by equating the volumes before and after deformation. Hence,

$$(\pi)(75)^2(100) = (\pi)(r)^2(50).$$

Thus, $r = 106$ mm. The quantity σ_f in Eq. (14.1) is the flow stress of the material, which is the stress required to continue plastic deformation of the workpiece at a particular true strain. The absolute value of the true strain that the workpiece has undergone at the end of the stroke in this operation is

$$\epsilon = \ln\left(\frac{100}{50}\right) = 0.69.$$

The flow stress can be determined using by referring to Eq. (2.8) and noting from Table 2.3 that, for 304 stainless steel, $K = 1275$ MPa and $n = 0.45$. Thus, for a true strain of 0.69, the flow stress is calculated to be 1100 MPa. Another calculation method is to refer to Fig. 2.5 and note that the flow stress for 304 stainless steel at a true strain of 0.69 is about 1000 MPa. The small difference between the two values is due to the fact that the data in Table 2.3 and Fig. 2.5 are from different sources. Taking the latter value for flow stress, the forging force can now be calculated, noting that in this problem the units in Eq. (14.1) must be in N and m. Thus,

$$\begin{aligned} F &= \left(1000 \times 10^6\right)(\pi)(0.106)^2(1) + \frac{(2)(0.2)(0.106)}{(3)(0.050)} \\ &= 4.5 \times 10^7 \text{ N} = 45 \text{ MN} = 4500 \text{ metric tons.} \end{aligned}$$

14.3 Impression-die and Closed-die Forging

In *impression-die forging*, the workpiece takes the shape of the die cavity while being forged between two shaped dies (Figs. 14.6a through c). This process is usually carried out at elevated temperatures in order to lower the forging forces and to develop enhanced ductility of the workpiece. Note in Fig. 14.6c that, during deformation, some of the material flows outward and forms a **flash**.

The flash has an important role in impression-die forging: The high pressure and the resulting high frictional resistance in the flash present a major constraint on the radially outward flow of the material in the die; this is due to the friction hill effect, described in Sec. 14.2. Thus, based on the principle that the material flows in the direction of least resistance (because it requires less energy), the material flows preferentially *into* the die cavity, eventually filling it completely.

Instead of being made as one piece, forging dies may be made of two or more pieces (*segmented*), including *die inserts* (Fig. 14.7) and particularly for complex part shapes. The inserts can easily be replaced in case of wear or failure in a particular region of the die; they are usually made of stronger and harder wear-resistant materials (see Section 33.5).

The blank to be forged can be prepared by (a) *cropping* (shearing, Section 16.2) from an extruded or drawn bar stock; (b) *powder metallurgy* or *casting*; or (c) it is a preformed blank from a prior forging operation. The blank is placed on the lower die, and as the upper die begins to descend, its shape gradually changes, as shown in Fig. 14.8a.

Preforming operations (Figs.14.8b and c) are typically made to enhance the distribution of the material into various regions of the blank, using simple dies with various contours. In **fullering**, material is distributed away from a die region; in **edging**, it is gathered into a localized region. The part is then formed into a rough shape by a process called **blocking**, using *blocker dies*. The final operation consists of finishing of the forging in *impression dies*, giving the forging its final shape. The flash is later removed by a trimming operation (Fig. 14.9).

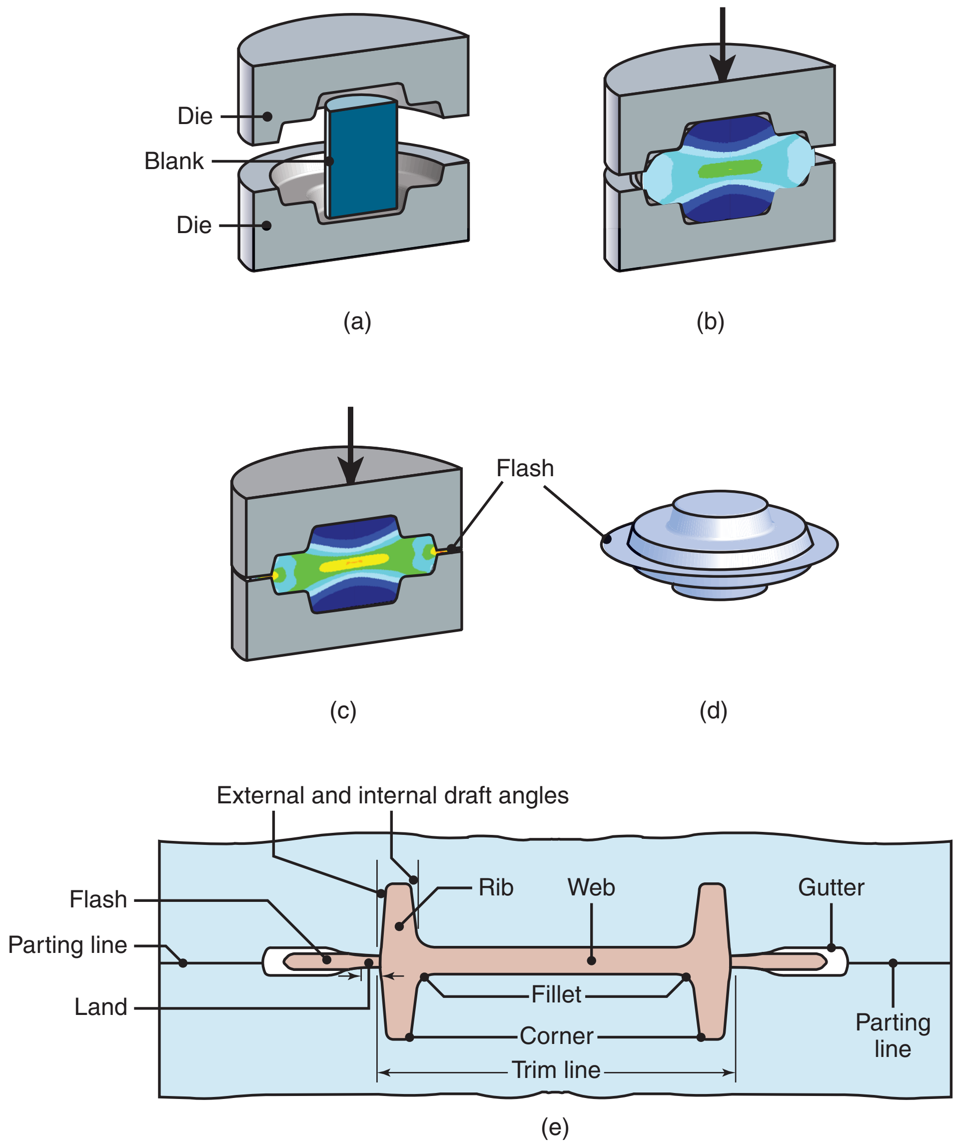


Figure 14.6: (a) through (d) Stages in impression-die forging of a solid round billet, with contours showing effective strain. Note the formation of flash, which is excess metal that is subsequently trimmed off. (e) Standard terminology for various features of a forging die.

Forging Force. The *forging force*, F, required in an *impression-die forging* operation can be estimated from the formula

$$F = k\sigma_f A, \tag{14.3}$$

where k is a multiplying factor, obtained from Table 14.2, σ_f is the flow stress of the material at the forging temperature, and A is the projected area of the forging, including the flash area. In hot-forging operations, the actual forging pressure for most metals typically ranges from 550 to 1000 MPa. As an example, assume that the flow stress of a material at the forging temperature is 300 MPa, and a part (such as that shown in Fig. 14.8a) has a projected area (with flash) of 0.05 m^2. Taking a value of $k = 10$ from Table 14.2, the forging force would be $F = (10)(300 \times 10^6)(0.05) = 150$ MN.

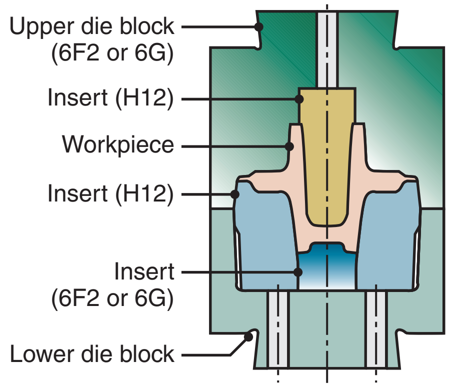


Figure 14.7: Die inserts used in forging an automotive axle housing (see Section 5.7 for die materials).

Closed-die Forging. The process shown in Fig. 14.6 is also referred to as *closed-die forging*. In true closed-die forging, however, a flash does not form (hence the term *flashless forging*), and the workpiece completely fills the die cavity (see right side of Fig. 14.10b). The accurate control of the blank volume and proper die design are essential to producing a forging with the required dimensional tolerances. Undersized blanks prevent the complete filling of the die cavity; conversely, oversized blanks generate excessive pressures and may cause dies to fail prematurely or the forging machine to jam.

Precision Forging. In order to reduce the number of additional finishing operations, hence cost, the trend has been toward greater precision in forged products (net-shape forming). Typical precision-forged products are gears, connecting rods, and turbine blades. Precision forging requires (a) special and more complex dies, (b) precise control of the blank's volume and shape, and (c) accurate positioning of the blank

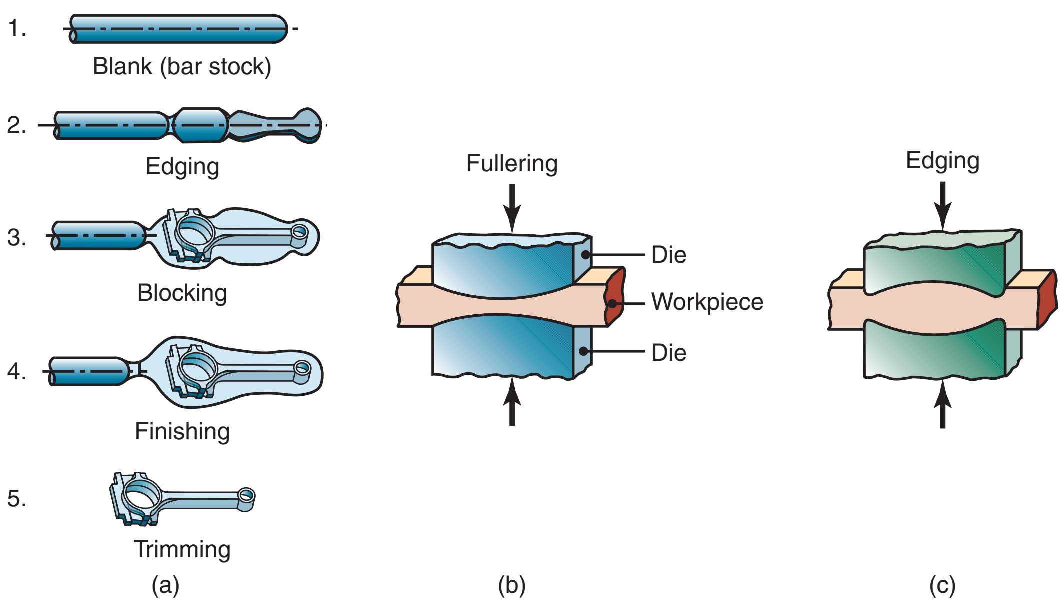


Figure 14.8: (a) Stages in forging a connecting rod for an internal combustion engine. Note the amount of flash required to ensure proper filling of the die cavities. (b) Fullering and (c) edging operations to distribute the material properly when preshaping the blank for forging.

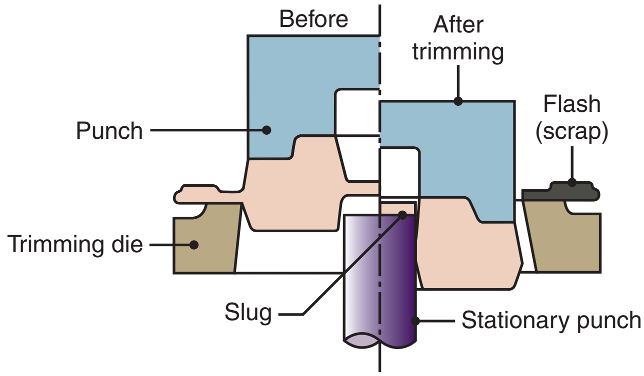


Figure 14.9: Trimming flash from a forged part. Note that the thin material at the center is removed by punching.

Table 14.2: Range of k Values for Eq. (14.3).

Shape	k
Simple shapes, without flash	3–5
Simple shapes, with flash	5–8
Complex shapes, with flash	8–12

in the die cavity. Because of the higher forces required to produce fine details on the part, precision forging requires higher capacity equipment. Aluminum and magnesium alloys are particularly suitable, because of the relatively low forging loads and forging temperatures that they require; however, steels and titanium also can be precision forged economically.

Forging Practice and Product Quality. A hot forging operation typically involves the following sequence of steps:

1. Prepare a slug, billet, or preform; if necessary, clean surfaces by such means as shot blasting (see Section 34.16).

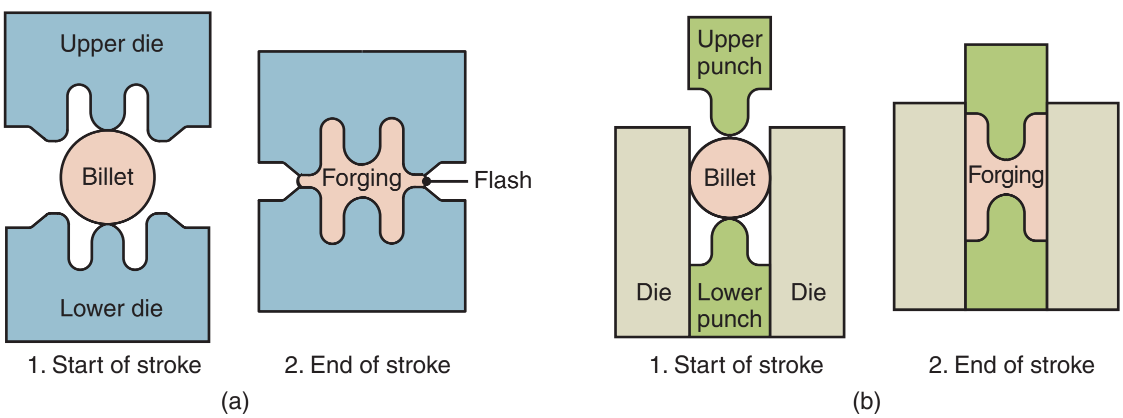


Figure 14.10: Comparison of (a) closed-die forging with flash and (b) precision or flashless forging of a round billet. *Source:* After H. Takemasu, V. Vazquez, B. Painter, and T. Altan.

2. Heat the workpiece in a suitable furnace; then, if necessary, descale it with a wire brush, water jet, steam, or by scraping. Some descaling also may occur during the initial stages of forging, when the thick, brittle scale falls off during forging.

3. Preheat, if necessary, and lubricate the dies.

4. Forge the billet in appropriate dies and in the proper sequence. If necessary, remove any excess material, especially any flash, by trimming, machining, or grinding.

5. Clean the forging, check for dimensional accuracy; if necessary, machine or grind to final dimensions and specified tolerances and surface finish.

6. Perform additional finishing operations, such as straightening and heat treating, for improving mechanical properties.

7. Inspect the forging for any external and internal defects.

The quality, dimensional tolerances, and surface finish of a forging depend on how well these operations have been performed. Generally, dimensional tolerances range between ± 0.5 and $\pm 1\%$ of the dimensions of the forging. In good practice, tolerances for hot forging of steel are usually less than ± 6 mm; in precision forging, they can be as low as ± 0.25 mm. Other factors that contribute to dimensional inaccuracies are draft angles, radii, fillets, die wear, whether the dies have closed properly, and mismatching of the dies.

14.4 Various Forging Operations

Several other operations related to the basic forging process are described below.

Coining. Essentially a closed-die forging process, coining was originally used in the minting of coins, medallions, and jewelry (Fig. 14.11). It is also used to produce a wide variety of parts with high accuracy, such as precision gears, industrial seals, and medical devices. The blank or slug is coined in a completely closed die cavity, in order to produce fine details. The pressures required can be as high as five or six

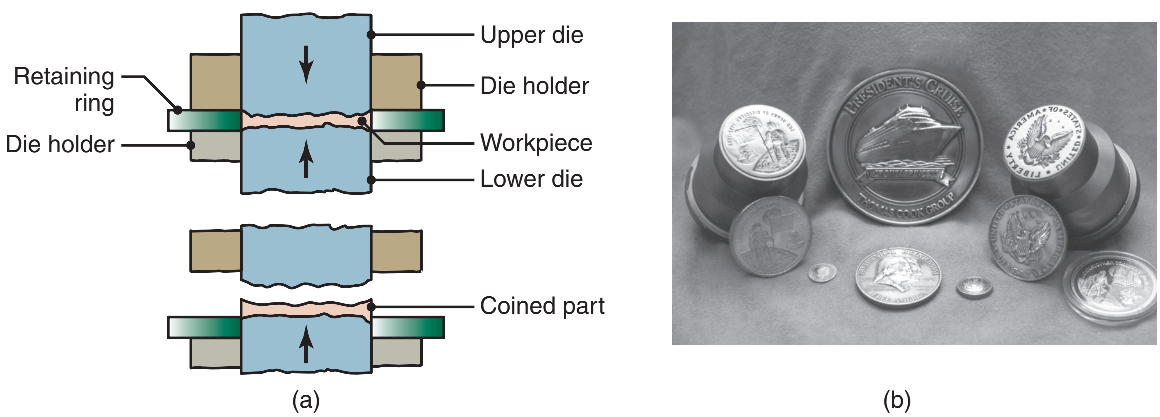


Figure 14.11: (a) Schematic illustration of the coining process. (b) An example of a modern coining operation, showing the coins and tooling. Note the detail and superior surface finish that can be achieved in this process. *Source:* Courtesy of C & W Steel Stamp Co., Inc.

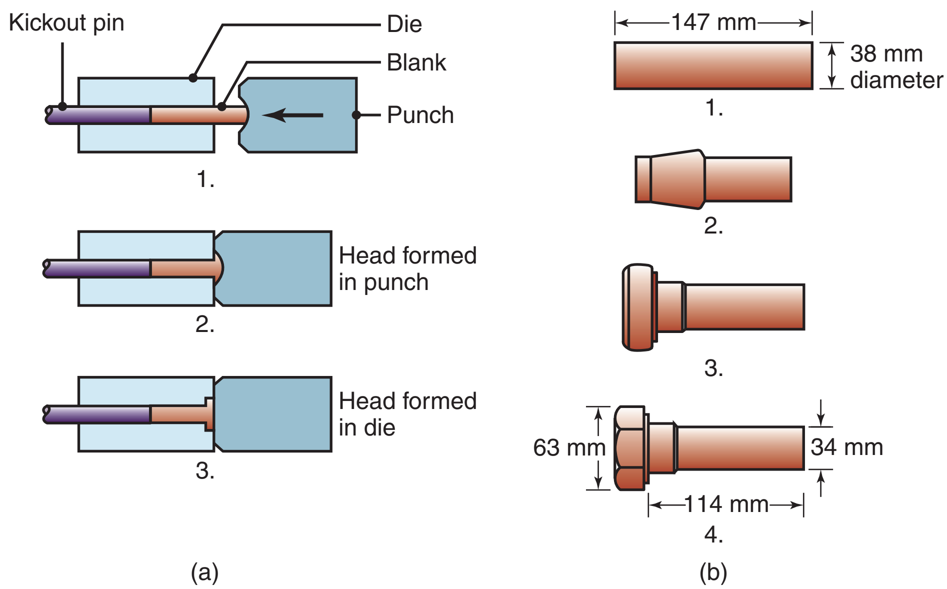


Figure 14.12: (a) Heading operation to form heads on fasteners, such as nails and rivets. (b) Sequence of operations used to produce a typical bolt head by heading.

times the strength of the material. On some parts, several coining operations may be required. Lubricants should not be used in coining because they can become entrapped in die cavities and, being incompressible, prevent the full reproduction of die-surface details and surface finish.

Marking parts with letters and numbers (for identification) also can be done rapidly through coining. **Sizing** is a process used mainly with forged or powder metal blanks (see Chapter 17) and other processes to improve surface finish and to impart the desired dimensional accuracy, with little or no change in part size.

Heading. Also called **upset forging**, *heading* is basically an upsetting operation, performed on the end of a rod or wire in order to increase the cross section. Typical products made are nails, bolt heads, screws, rivets, and fasteners (Fig. 14.12a). Heading can be carried out cold, warm, or hot, and can be combined with cold-extrusion processes to make various parts, as described in Section 15.4. Heading operations are performed on machines called **headers**; they are highly automated, with production rates of hundreds of pieces per minute for small parts. Hot heading operations on larger parts typically are performed on **horizontal upsetters**.

An important consideration in heading is the tendency for the workpiece to *buckle* if its unsupported length-to-diameter ratio is too high. This ratio is generally limited to 3:1, but with appropriate dies, it can be higher. Higher ratios can be accommodated if the diameter of the die cavity is not more than 1.5 times the diameter of a round bar.

Piercing. This is a process of indenting, but not breaking through, the surface of a workpiece with a punch, in order to produce a cavity or an impression (Fig. 14.13). The workpiece may be confined in a container, such as a die cavity, or may be unconstrained. The surface deformation of the workpiece will depend on how much it is constrained from flowing freely as the punch penetrates. Piercing may be followed by punching to produce a hole in the part; see the slug above the stationary punch in the central portion of Fig. 14.9.

The *piercing force* depends on (a) the cross-sectional area and the tip geometry of the punch, (b) the strength of the workpiece material, and (c) friction at the punch-workpiece interfaces. The pressure may range from three to five times the strength of the material, which is about the same level of stress required to make an indentation in hardness testing (see Section 2.6).

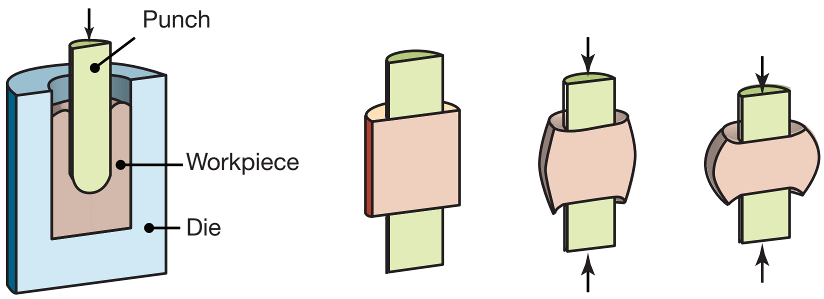


Figure 14.13: Examples of piercing operations.

Case Study 14.1 Manufacture of a Stepped Pin by Heading and Piercing Operations

Figure 14.14a shows a stepped pin made from SAE 1008 steel, and used as a portion of a roller assembly to adjust the position of a car seat. The part is complex and must be produced in a progressive manner, in order to produce the required details and fill the die completely. The cold-forging steps used to produce this part are shown in Fig. 14.14b. First, a solid, cylindrical blank is extruded (Chapter 15) in two operations, followed by upsetting. The upsetting operation uses a conical cross section in the die to produce the preform, and is oriented such that material is concentrated at the top of the part in order to ensure proper die filling. After impression-die forming, a piercing operation is performed to form the bore.

Hubbing. This process consists of pressing a hardened punch, with a specific tip geometry, into the surface of a block of metal. The cavity produced is subsequently used as a die for forming operations, such as those employed in making tableware. The die cavity is usually shallow, but for deeper cavities some material may be removed from the surface of the block by machining prior to hubbing (see Figs. 24.2c and d). The *hubbing force* can be estimated from the equation

$$\text{Hubbing force} = 3(S_{\text{ut}})(A), \tag{14.4}$$

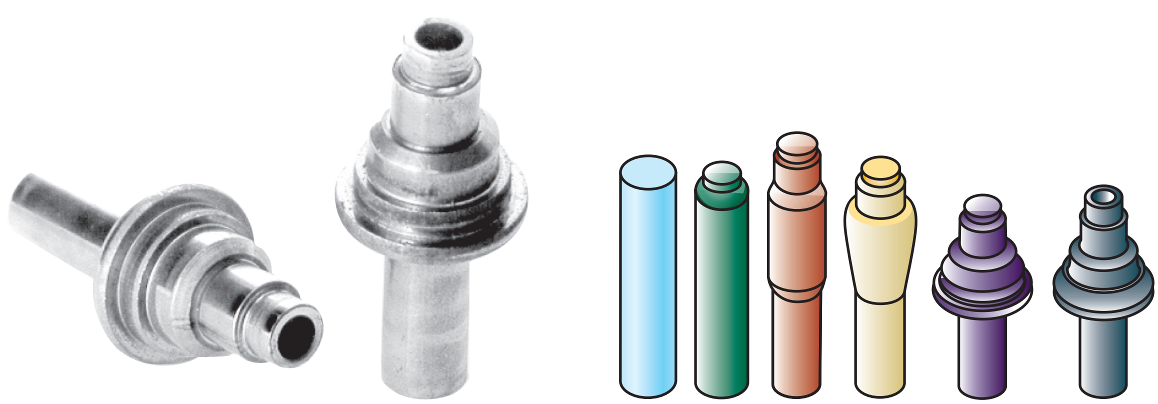


Figure 14.14: (a) The stepped pin used in Case Study 14.1. (b) Illustration of the steps used to produce the stepped pin. *Source:* Courtesy of National Machinery, LLC.

where S_{ut} is taken from Table 2.2, and A is the projected area of the impression. As an example: for high-strength steel, with $S_{\text{ut}} = 1500$ MPa and a part with a projected area of 400 mm^2, the hubbing force is $(3)\left(1500\ \text{N/mm}^2\right)\left(400\ \text{mm}^2\right) = 1.8\ \text{MN} = 183$ metric tons.

Orbital Forging. In this process, the upper die moves along an orbital path and forms the part *incrementally*. The operation is similar to the action of a mortar and pestle, used for crushing herbs and seeds. Typical components made are disk-shaped and conical parts, such as bevel gears and gear blanks. The forging force is relatively small, because at any particular instant, the die contact is concentrated onto a small area of the workpiece (see also *incremental forging* next). The operation is relatively quiet, and parts can be formed within 10 to 20 cycles of the orbiting die.

Incremental Forging. In this process, a tool forges a blank into a particular shape. The operation is somewhat similar to cogging (Fig. 14.4a), in which the die deforms the blank, in several small steps, to a different extent at different positions. Because of the small contact area with the die, the process requires much lower forces then in conventional impression-die forging; the tools are simpler and less costly.

Isothermal Forging. Also known as **hot-die forging**, in this process the dies are heated to the same temperature as that of the hot workpiece. Because the workpiece remains hot, its flow stress and high ductility are maintained during forming. Thus, the forging load is low, and the material flow within the die cavity is improved.

Complex parts can be isothermally forged, with good dimensional accuracy and to near-net shape by one stroke in a hydraulic press. The dies are usually are made of nickel or molybdenum alloys, because of their resistance to high temperature. The process is expensive and the production rate is low. It can, however, be economical for specialized, intricate forgings, made of such materials as superalloys and titanium.

Rotary Swaging. In this process, also known as *radial forging*, *rotary forging*, or *swaging*, a solid rod or tube is subjected to radial impact forces, using a set of reciprocating dies of the machine (Figs. 14.15a and b). The dies are activated by means of a set of rolls within a cage, in an action similar to that of a roller bearing. The workpiece is stationary and the dies rotate while moving radially in their slots, striking the workpiece at rates as high as 20 strokes per second.

In **die-closing swaging machines**, die movements are through the reciprocating motion of the wedges (Fig. 14.15c). The dies can be opened wider than those in rotary swagers, thereby accommodating large-diameter or variable-diameter parts. In other arrangements, the dies do not rotate but move radially in and out.

The swaging process also can be used to *assemble* fittings over cables and wire, in which case the tubular fitting is swaged directly onto the cable. The process is also used for operations such as *pointing* (tapering the tip of a round rod) and *sizing* (finalizing the dimensions of a part).

Swaging generally is limited to a maximum workpiece diameter of about 150 mm, and parts as small as 0.5 mm have been swaged. Dimensional tolerances range from ± 0.05 to ± 0.5 mm. The process is suitable for medium-to-high rates of production, as high as 50 parts per minute, depending on part complexity. Swaging is a versatile process and is limited in length only by the length of the bar supporting the mandrel, if one is needed (see Fig. 14.15b).

Tube Swaging. In this process, the internal diameter and/or the thickness of a tube is reduced, with or without using *internal mandrels* (Figs. 14.16a and b). For small-diameter tubing, high-strength wire can be used as a mandrel. Mandrels also can be made with longitudinal grooves, to allow swaging of internally shaped tubes (Fig. 14.16c). For example, the *rifling* in gun barrels (internal spiral grooves to give gyroscopic effect to bullets) can be produced by swaging a tube over a mandrel with spiral grooves. Special machinery can swage gun barrels, and other parts, with starting diameters as large as 350 mm.

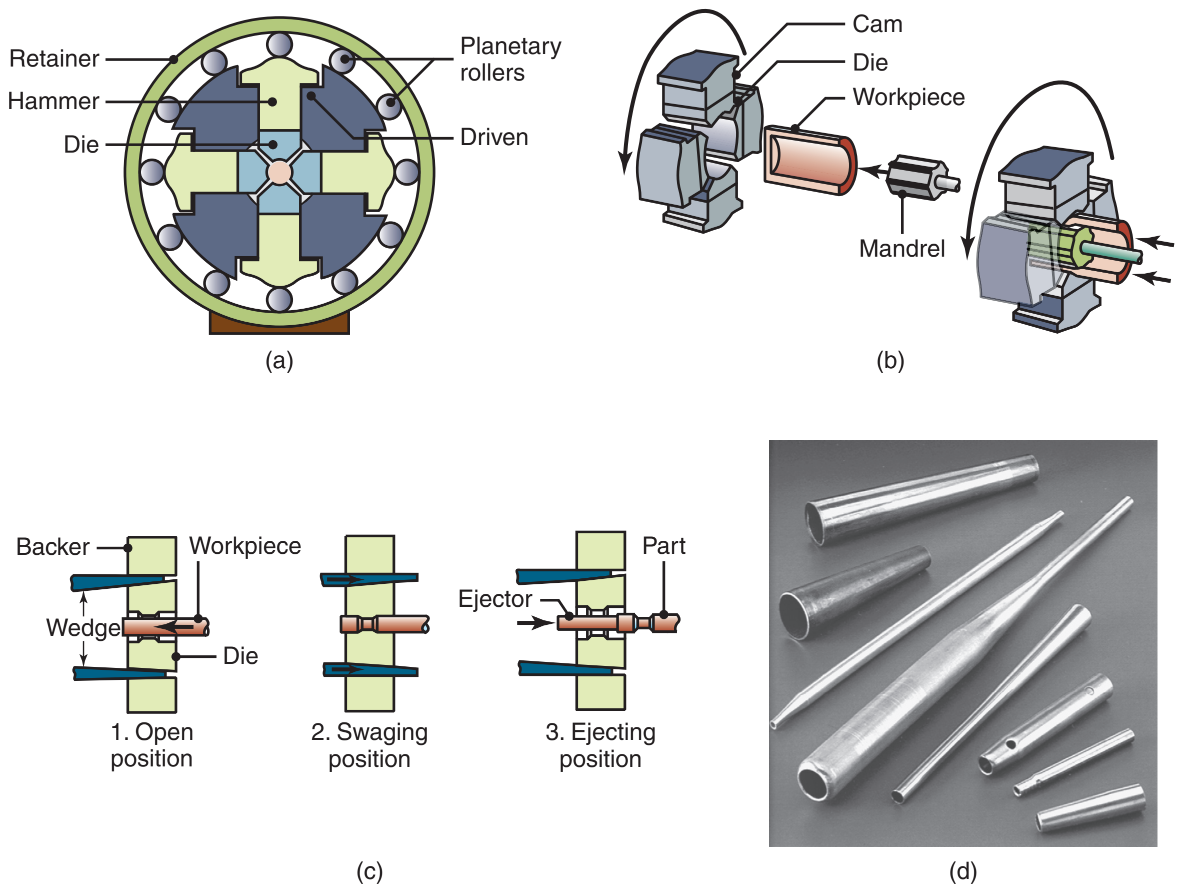


Figure 14.15: (a) Schematic illustration of the rotary-swaging process. (b) Forming internal profiles on a tubular workpiece by swaging. (c) A die-closing swaging machine, showing forming of a stepped shaft. (d) Typical parts made by swaging. *Source:* (d) Courtesy of Woodsage Holdings, LLC.

14.5 Forgeability of Metals; Forging Defects

Forgeability is generally defined as the capability of a material to undergo deformation in forging without cracking. Various tests have been developed over the years to quantify forgeability; however, because of their complex nature, only two simple tests have had general acceptance: upsetting and hot twist.

In the **upsetting test**, a solid, cylindrical specimen is upset between flat dies to the reduction in height at which cracks on the barreled surfaces begin to develop (see also Fig. 2.20d). The greater the deformation prior to cracking, the greater the forgeability of the metal. The second method is the **hot-twist test**, in which a round specimen is twisted continuously and in the same direction until it fails. This test is performed on a number of specimens and at different temperatures, and the number of complete turns that each specimen undergoes before failure at each temperature is plotted. The temperature at which the maximum number of turns occurs then becomes the forging temperature for maximum forgeability. This test has been found to be useful particularly for steels. The forgeability of various metals and alloys is given in Table 14.3, in decreasing order.

More comprehensively, forgeability is rated on such considerations as (a) ductility and strength of the material, (b) forging temperature required, (c) frictional behavior between the die and workpiece, and (d) the quality of the forgings produced. These ratings should be regarded only as general guidelines. Typical

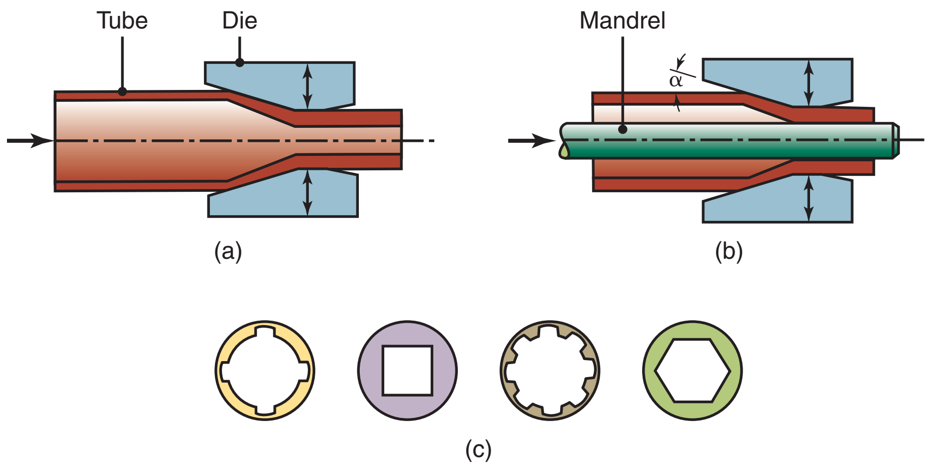


Figure 14.16: (a) Swaging of tubes without a mandrel; note the increase in wall thickness in the die gap. (b) Swaging with a mandrel; note that the final wall thickness of the tube depends on the mandrel diameter. (c) Examples of cross sections of tubes produced by swaging on shaped mandrels. Rifling (internal spiral grooves) in small gun barrels can be made by this process.

hot-forging temperature ranges for various metals and alloys are included in Table 14.3. For *warm* forging, temperatures range from 200° to 300°C for aluminum alloys, and 550° to 750°C for steels.

Forging Defects. In addition to surface cracking, various other defects can develop during forging as a result of the material flow pattern in the die, as described in Section 14.6 regarding die design. For example, if there is an insufficient volume of material to completely fill the die cavity, the web may buckle and develop *laps* (Fig. 14.17a). Conversely, if the web is too thick, the excess material flows past the already formed portions of the forging and develop internal cracks (Fig. 14.17b).

Table 14.3: Forgeability of Metals, in Decreasing Order. See also Table 15.1.

Metal or alloy	Approximate range of hot-forging temperatures (°C)
Aluminum alloys	400–550
Magnesium alloys	250–350
Copper alloys	600–900
Carbon- and low-alloy steels	850–1150
Martensitic stainless steels	1100–1250
Austenitic stainless steels	1100–1250
Titanium alloys	700–950
Iron-based superalloys	1050–1180
Cobalt-based superalloys	1180–1250
Tantalum alloys	1050–1350
Molybdenum alloys	1150–1350
Nickel-based superalloys	1050–1200
Tungsten alloys	1200–1300

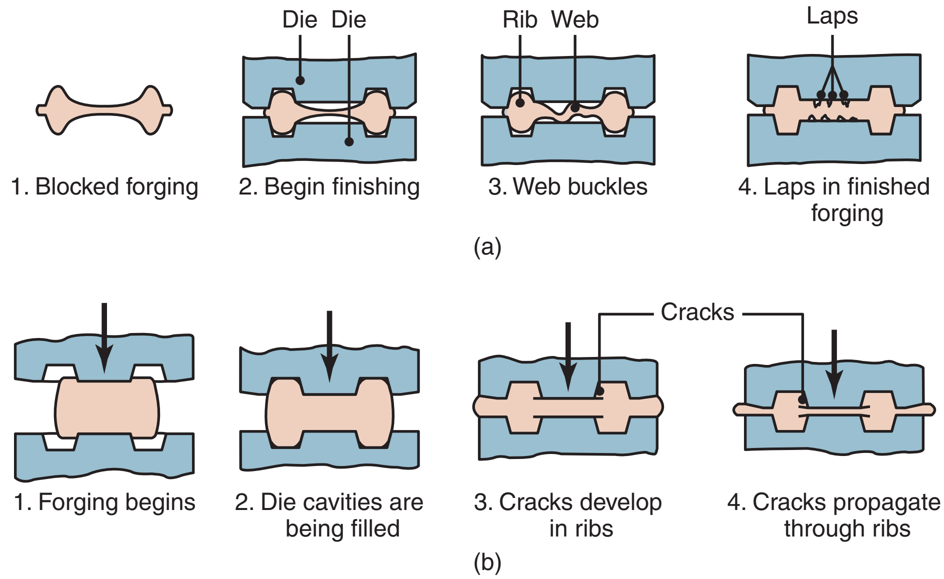


Figure 14.17: Examples of defects in forged parts. (a) Laps formed by web buckling during forging; web thickness should be increased to avoid this problem. (b) Internal defects caused by an oversized billet. Die cavities are filled prematurely, and the material at the center flows past the filled regions as the dies close.

The various radii in the forging-die cavity can significantly influence the formation of defects. Internal defects also may develop because of (a) nonuniform deformation of the material in the die cavity, (b) temperature gradients developed throughout the workpiece during forging, and (c) microstructural changes caused by phase transformations. The *grain-flow pattern* of the material in forging also is important. The flow lines may reach a surface perpendicularly, as shown in Fig. 14.13. In this condition, known as **end grains**, the grain boundaries become directly exposed to the environment and can be attacked by it, developing a rough surface which also acts as stress raisers.

Forging defects can cause fatigue failures, corrosion, and wear during the service life of the forging. The importance of inspecting forgings prior to their placement in service, particularly in critical applications is obvious. Inspection techniques for manufactured parts are described in Chapter 36.

14.6 Die Design, Die Materials, and Lubrication

The design of forging dies requires considerations of (a) the shape and complexity of the workpiece, (b) forgeability, (c) strength and its sensitivity to deformation rate, (d) temperature, (e) frictional characteristics at the die–workpiece interfaces and (f) die distortion under the forging loads. The most important rule in die design is that the part will preferentially flow in the direction of least resistance. Workpiece *intermediate shapes* should be considered so that die cavities can be filled properly and without any defects. An example of the intermediate shapes for a connecting rod is given in Fig. 14.8a.

With continuing advances in reliable simulation of all types of metalworking operations, software is widely available to help predict material flow in die cavities (see Fig. 14.18) and also predict final material microstructure and mechanical properties. The simulations incorporate various conditions, such as workpiece temperature, heat transfer to dies, frictional conditions at die–workpiece contact surfaces, and forging speed. Such software has now become essential in die design, especially for eliminating defects (see also Section 38.7).

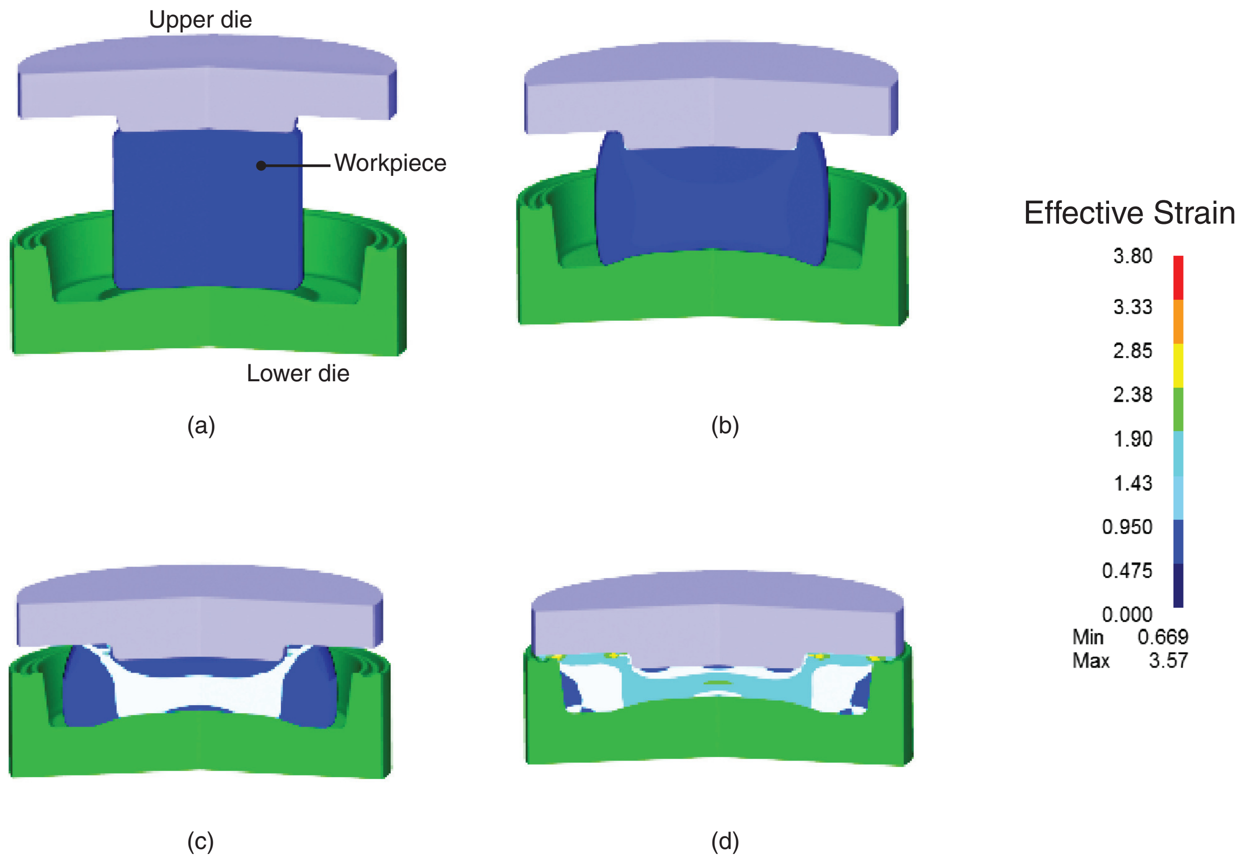


Figure 14.18: Deformation of a blank during forging as predicted by the software program DEFORM, based on the finite-element method of analysis. *Source:* Courtesy of Scientific Forming Technologies Corporation.

Preshaping. The requirements for preshaping a workpiece are: (a) the material should not flow easily into the flash, as otherwise die filling will be incomplete, (b) the grain-flow pattern should be favorable for the strength and reliability of the part made, and (c) sliding at the die–workpiece interface should be minimized in order to reduce die wear. The selection of preshapes involves calculations of cross sectional areas at each location in the forging.

Die Features. The terminology for forging dies is shown in Fig. 14.6d. For most forgings, the **parting line** is located at the largest cross section of the part. For simple symmetric shapes, the parting line is usually a single plane at the center of the forging; for more complex shapes, the line may not lie in a single plane. The dies are then designed in such a way that they properly engage while avoiding side thrust forces and maintaining die alignment during forging.

After constraining lateral flow to ensure proper die filling, the flash material is allowed to flow into a **gutter**, so that the extra flash does not increase the forging load excessively. A general guideline for flash thickness is 3% of the maximum thickness (vertical dimension) of the forging. The length of the **land** is usually 2 to 5 times the flash thickness.

Draft angles are essential in almost all forging dies in order to facilitate removal of the forging. Upon cooling, the forging shrinks both in radial and longitudinal directions, therefore internal draft angles (about 7° to 10°) are made larger than external angles (about 3° to 5°).

Selection of the proper corner and fillet radii is important to ensure smooth flow of the metal into the die cavity and for improving die life. Small radii generally are undesirable because of their adverse effects on metal flow and their tendency to cause rapid die wear (as a result of stress concentration and thermal cycling). Small fillet radii also can cause fatigue cracking of the dies. As a general rule, these radii should be as large as can be permitted by the design of the forging. As with the patterns used in casting (Section 12.2.1), *allowances* are provided in forging-die design when machining or grinding of the forging is necessary. Machining allowance should be provided at flanges, holes, and mating surfaces.

Die Materials. General requirements for die materials are:

- Strength and toughness, especially at elevated temperatures
- Hardenability and ability to be hardened uniformly
- Resistance to mechanical and thermal shock
- Wear resistance, particularly abrasive wear, because of the presence of hard scale on the surfaces of hot forgings.

Common die materials are tool and die steels, containing chromium, nickel, molybdenum, and vanadium (see Tables 5.7 and 5.8). Dies are made from *die blocks*, which themselves are forged from castings, and then machined and finished to the desired shape, dimensional accuracy, and surface finish.

Lubrication. A wide variety of metalworking fluids are available for use in forging (Section 33.7). Lubricants greatly influence friction and wear, in turn affecting the forging forces required (see Eq. (14.1)), die life, and the manner in which the material flows into die cavities. Lubricants can also act as a *thermal barrier* between the hot workpiece and the relatively cool dies, thus slowing the rate of cooling of the workpiece and significantly improving metal flow. An additionally important function of the lubricant is to act as a *parting agent*, preventing the forging from sticking to the dies and to help release it from the die.

14.7 Die-manufacturing Methods and Die Failures

Dies are an important factor in the overall economics of forging, as their cost and the lead time required to produce them can be extensive; some dies can take months to make and cost more than a million dollars. Equally important are the proper maintenance of dies and their modifications and repair.

Several manufacturing methods, either singly or in combination, can be used to make dies for forging. These methods include casting, forging, machining, grinding, and by the advanced machining techniques described in Chapter 27. An important and continuing trend is the production of tools and dies by **rapid tooling**, using rapid prototyping techniques (Section 20.6).

Producing a cavity in a die block is called **die sinking**. The process of **hubbing** (Section 14.4), either cold or hot, also may be used to make smaller dies with shallow cavities. Dies are subsequently heat treated, for higher hardness, toughness, and wear resistance (Chapter 33). If necessary, their surface profile and finish are further improved by finish grinding and polishing, either by hand or using programmable industrial robots (Section 37.6).

The choice of a die-manufacturing method depends on die size and shape, and the particular operation in which the die is to be used, such as casting, forging, extrusion, powder metallurgy, or molding. As in all manufacturing operations, cost often dictates the process selected. Dies can be cast from steels, cast irons, and nonferrous alloys. The processes used for preparing them may range from sand casting (for large dies, weighing several tons) to shell molding (for small dies). Cast steels generally are preferred for large dies because of their strength, toughness, and wear resistance as well as the ease with which the steel composition, grain size, and other properties can be controlled and modified as necessary.

Most commonly, dies are *machined* from forged die blocks using such processes as high-speed milling, turning, grinding, electrical discharge (including wire EDM), and electrochemical machining (see Part IV).

Such an operation is shown in Fig. I.11b for making molds for eyeglass frames. For high-strength and wear-resistant die materials that are hard or are heat treated (and thus difficult to machine), processes such as hard machining and electrical and electrochemical machining are in common practice. An increasing trend is a die produced by *additive manufacturing* (see Section 20.6), although the number of parts made is lower than with conventionally produced tooling.

Typically, a die is machined by milling on a computer-controlled machine tool, using various software packages (see Fig. I.11) that have the capability of optimizing the *cutting-tool path*. Thus, the best surface finish can be produced in the least possible machining time. Equally important is the *setup* for machining, because ideally dies should be machined in one setup, without having to remove them from their fixtures and reorient them for subsequent machining operations.

After heat treating to achieve the desired mechanical properties, dies usually are subjected to *finishing operations* (Section 26.7), such as grinding, polishing, and chemical and electrical methods, for the desired surface finish and dimensional accuracy. Finishing includes *laser surface treatments* and *coatings* (Chapter 34) to improve die life. Laser beams also may be used for die repair and reconfiguration of the worn regions of dies (see also Fig. 33.12).

Die Costs. Some qualitative ranges of tool and die costs are given throughout this book, such as in Table 12.6. Even small and relatively simple dies can cost hundreds of dollars. The cost of a set of dies for automotive body panels can be on the order of $2 million. On the other hand, because a large number of parts usually are made from one set of dies, *die cost per piece made* is generally a small portion of a part's manufacturing cost (see also Section 40.10). The *lead time* required to produce dies also can have a significant impact on productivity and the overall manufacturing cost of parts made.

Die Failures. Failure of dies generally results from one or more of the following causes:

- Improper die design
- Defective or improper selection of die material
- Improper manufacturing, heat-treatment, and finishing operations
- Overheating and heat checking (cracking caused by temperature cycling of dies)
- Excessive die wear
- Overloading (excessive force on the die)
- Improper alignment of die components or segments
- Improper handling or misuse of the die.

Other Considerations. In order to withstand the forces involved, a die must have sufficiently large cross sections and clearances (to prevent jamming). Cooling channels can be machined into a die to help extract heat. Abrupt changes in cross section, sharp corners, radii, fillets, and a coarse surface finish (including grinding marks and their orientation on die surfaces) act as stress raisers, and thus die life. For improved strength and to reduce the tendency for cracking, dies may be made in segments and assembled into a complete die, with rings that prestress the dies. Proper handling, installation, assembly, and alignment of dies are essential. Overloading of tools and dies can cause premature failure. A common cause of damage to dies is the failure of the operator or of a programmable robot, to remove a formed part from the die before another blank is loaded into the die.

Table 14.4: Typical Speed Ranges of Forging Equipment

Equipment	m/s
Hydraulic press	0.06–0.30
Mechanical press	0.06–1.5
Screw press	0.6–1.2
Gravity drop hammer	3.6–4.8
Power drop hammer	3.0–9.0
Counterblow hammer	4.5–9.0

14.8 Forging Machines

Various types of forging machines are available, with a wide range of capacities (tonnage), speeds, and speed–stroke characteristics (Table 14.4).

Hydraulic Presses. These presses operate at constant speeds and are *load limited* (load restricted); the press stops if the load required exceeds its capacity. Large amounts of energy can be transmitted from the press to the workpiece by a constant load throughout the whole stroke, the speed of which can be controlled. Because forging in a hydraulic press takes longer than in other types of forging machines, the workpiece may cool rapidly unless the dies are heated (see *isothermal forging*, Section 14.4). Compared with mechanical presses, hydraulic presses are slower and involve higher initial costs, but they require less maintenance.

A hydraulic press typically consists of a frame with two or four columns, pistons, cylinders (Fig. 14.19), rams, and hydraulic pumps driven by electric motors. The ram speed can be varied during the stroke. Press capacities range up to 125 MN (12,700 metric tons) for open-die forging, and up to around 730 MN (74,000 metric tons), although this is rare (450 MN (45,000 metric tons) is more common for high capacity hydraulic presses). The main landing-gear support beam for the Boeing 747 aircraft is forged in a 450-MN (45,000-metric ton) hydraulic press (Fig. 14.19d), with the forging shown in the forefront. The is made of a titanium alloy and weighs approximately 1350 kg.

Mechanical Presses. These presses are basically of the crank or of the eccentric type (Fig. 14.19a). The speed varies from a maximum at the center of the stroke to zero at the bottom of the stroke, thus are *stroke limited*. The energy in a mechanical press is generated by a large flywheel powered by an electric motor. A clutch engages the flywheel to an eccentric shaft; a connecting rod then translates the rotary motion into a reciprocating linear motion. A *knuckle-joint* mechanical press is shown in Fig. 14.19b. Because of the linkage design, very high forces can be applied (see also Fig. 11.19).

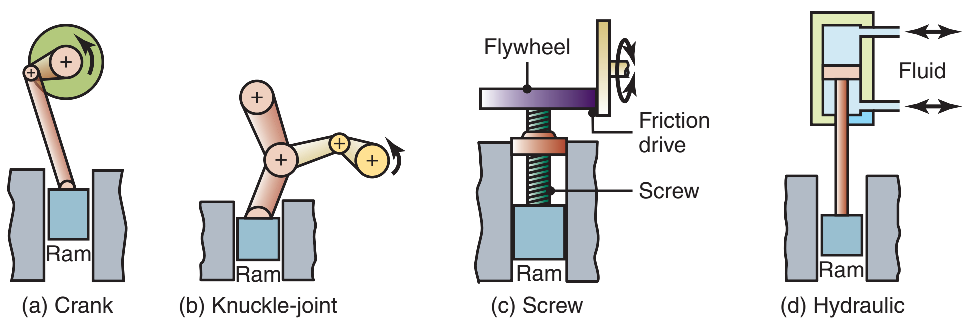


Figure 14.19: Schematic illustration of the principles of various forging machines. (a) Mechanical press with an eccentric drive; the eccentric shaft can be replaced by a crankshaft to give up-and-down motion to the ram. (b) Knuckle-joint press. (c) Screw press. (d) Hydraulic press.

The force available in a mechanical press depends on the stroke position, and it becomes extremely high at the end of the stroke. Thus, proper setup is essential to avoid breaking the dies or equipment components. Mechanical presses have high production rates, are easier to automate, and require less operator skill than do other types of machines. Press capacities generally range from 2.7 to 10^7 MN (270 to 10,800 metric tons). Mechanical presses are preferred for forging parts requiring high precision.

Screw Presses. These presses (Fig. 14.19c) derive their energy from a flywheel, hence they are *energy limited*. The forging load is transmitted through a large vertical screw, and the ram comes to a stop when the flywheel energy has been dissipated. If the dies do not close, the operation is repeated until the forging is completed to its final shape. Screw presses are used for various open-die and closed-die forging operations. They are particularly suitable for small production quantities and for thin parts with high precision, such as turbine blades. Press capacities range from 1.4 to 280 MN (144 to 28,350 metric tons).

Hammers. Hammers derive their energy from the potential energy of the ram, which is converted into kinetic energy, thus they are *energy limited*. Unlike hydraulic presses, hammers operate at high speeds, thus minimizing the cooling of a hot forging, thus allowing forging of complex shapes, particularly those with thin and deep recesses in dies. To complete the forging, several successive blows are usually made in the same die. Hammers are available in a variety of designs, and are the most versatile and the least expensive type of forging equipment.

Drop Hammers. In *power drop hammers*, the ram's downstroke is accelerated by steam, air, or hydraulic pressure. Ram weights range from 225 to 22,500 kg, with energy capacities reaching 1150 kJ. In the operation of *gravity drop hammers*, a process called **drop forging**, the energy is derived from the free-falling ram. The available energy is the product of the ram's weight and the height of its drop. Ram weights range from 180 to 4500 kg, with energy capacities ranging up to 120 kJ.

Counterblow Hammers. These hammers have two rams that simultaneously approach each other, horizontally or vertically, to forge the part. As in open-die forging operations, the workpiece may be rotated between blows to better shape the workpiece. Counterblow hammers operate at high speeds and transmit less vibration to their bases. Capacities range up to 1200 kJ.

Servo Presses. Used for forging and stamping applications (Fig. 14.20), these presses utilize servo drives along with linkage mechanisms, as in mechanical, knuckle joint, or screw presses. There are no clutches or brakes; instead, the desired velocity profile is achieved through a servo motor controller. The servo drive allows considerable flexibility regarding speeds and stroke heights, thus simplifying set up and allows an optimized velocity profile for forging difficult materials or products. In addition, servo presses can produce parts with as little as 10% of the energy consumption of other presses, attributable mainly to their low energy costs when not producing parts (see Section 40.5). Servo presses can develop forces up to 25,000 kN (2500 metric tons); larger forces can be developed by *hybrid* machines that combine servo drives with energy storage in a flywheel.

14.9 Economics of Forging

Several factors are involved in the cost of forgings, depending on the complexity of the forging and tool and die costs, which range from moderate to high. As in other manufacturing operations, these costs are spread out over the total number of parts forged with that particular die set. Thus, referring to Fig. 14.21, even though the cost of workpiece material per forging is constant, setup and tooling costs per piece decrease as the number of pieces forged increases.

The ratio of the cost of the die material to the total cost of forging a part increases with the weight of forgings. Because dies must be made and forging operations must be performed regardless of the size of the forging, the cost of dies and of the forging operation relative to material cost is high for small parts. By contrast, die material costs are relatively low.

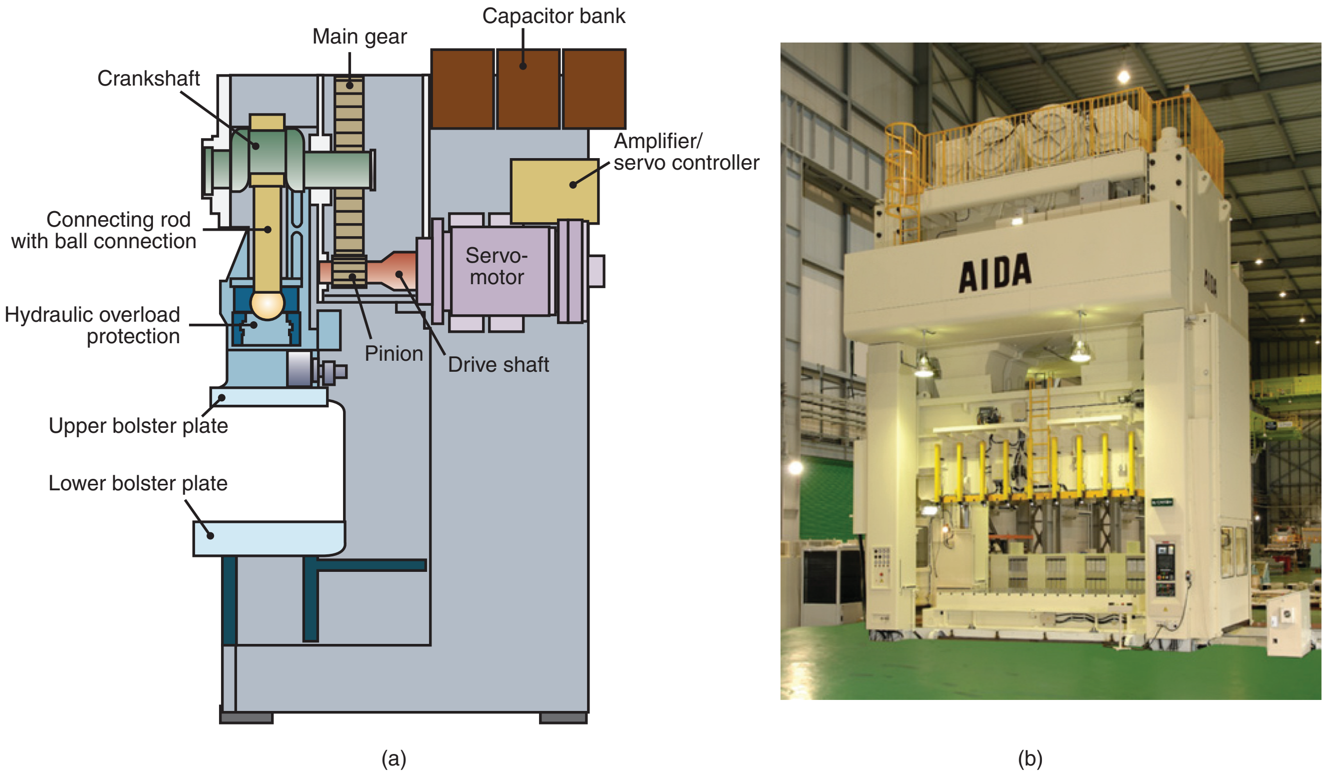


Figure 14.20: (a) Schematic illustration of a servo press, with the power source and transmission components highlighted. (b) An example of a servo press, with a 23,000 kN (2250 metric tons) capacity. *Source:* Courtesy of Aida Engineering, Inc.

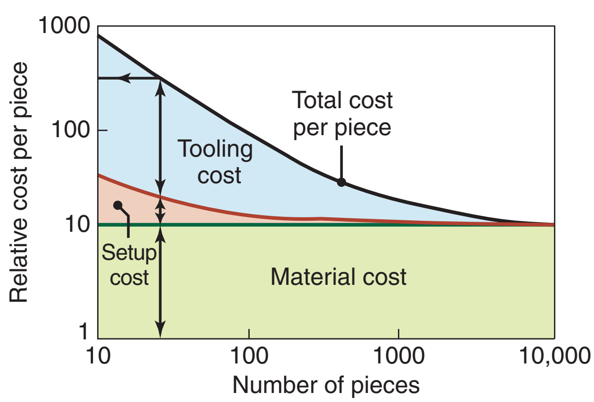


Figure 14.21: Typical cost per piece in forging; note how the setup and the tooling costs per piece decrease as the number of pieces forged increases if all pieces use the same die.

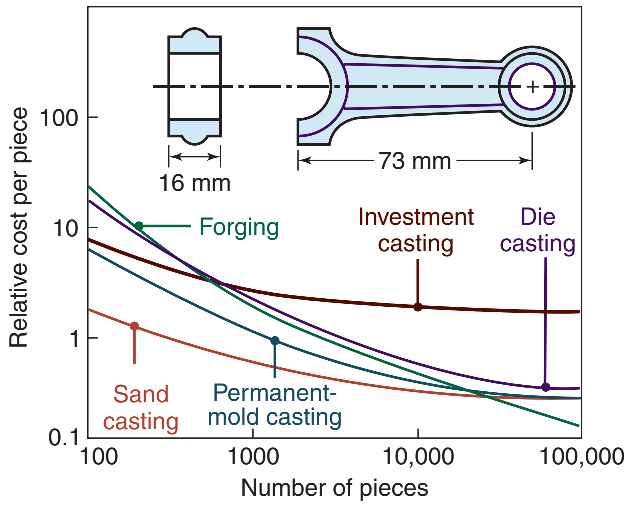


Figure 14.22: Relative unit costs of a small connecting rod made by various forging and casting processes. Note that, for large quantities, forging is more economical, and sand casting is the most economical process for fewer than about 20,000 pieces.

The size of forgings also has some effect on cost. Sizes range from small forgings (utensils and small automotive components) to large ones (gears, crankshafts, and connecting rods for large engines). As the size of the forging increases, the share of material cost in the total cost also increases, but at a lower rate. This is because (a) the incremental increase in die cost for larger dies is relatively small, (b) the machinery and operations involved are essentially the same regardless of forging size, and (c) the labor involved per forging is not that much higher.

The total cost involved is not influenced to any major extent by the type of materials forged. Because they have been reduced significantly by automated and computer-controlled operations, labor costs generally are moderate. Furthermore, die design and manufacturing are now performed by computer-aided design and manufacturing (Chapter 38), resulting in major savings.

The cost of forging a part, compared to that of producing it by other processes, such as casting, powder metallurgy, machining, is an important consideration. For example, for shorter production runs and all other factors being the same, making a certain part by, say, expendable-mold casting may well be more economical than producing it by forging (Fig. 14.22). Recall that this particular casting method does not require expensive molds and tooling, whereas forging typically requires expensive dies. The competitive aspects of manufacturing and process selection are described in greater detail in Chapter 40.

Case Study 14.2 Suspension Components for the Lotus Elise Automobile

The Lotus Elise is a high-performance sports car, designed for superior ride and handling. The Lotus group investigated the use of steel forgings instead of extruded-aluminum suspension uprights in order to reduce cost and improve reliability and performance. Their development efforts consisted of two phases, shown in Fig. 14.23. The first phase involved the development of a forged-steel component that can be used on the existing Elise sports car; the second phase involved the production of a suspension upright for a new model. A new design was developed using an iterative process, with advanced software tools to reduce the number of components and to determine the optimum geometry. The material selected for the upright was an air-cooled forged steel, which gives uniform grain size and microstructure, and uniform high strength without the need for heat treatment. These materials also have approximately 20% higher fatigue strengths than traditional carbon steels, such as AISI 1548-HT used for similar applications.

The revised designs are summarized in Table 14.5. As can be seen, the optimized new forging design (Fig. 14.23d) resulted in significant cost savings. Although it also resulted in a small weight increase, when compared to the aluminum-extrusion design, the weight penalty is recognized as quite small. Furthermore, the use of forged steel for such components is especially advantageous in fatigue-loading conditions, constantly encountered by suspension components. The new design also had certain performance advantages, in that the component stiffness is now higher, which registered as improved customer satisfaction and better "feel" during driving. Furthermore, the new design reduced the number of parts required, thus satisfying another fundamental principle in design.

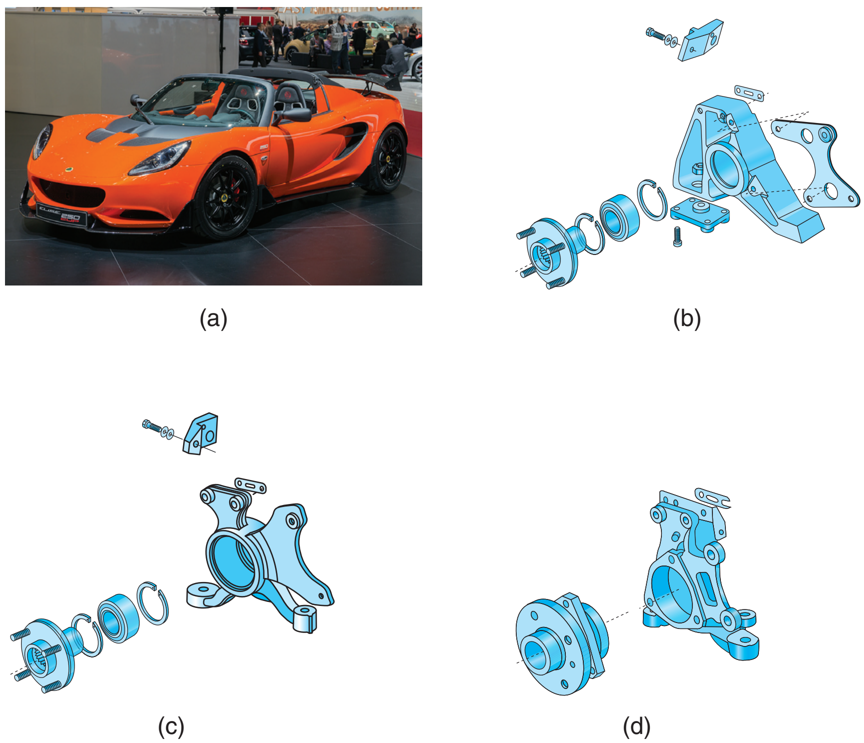


Figure 14.23: (a) The Lotus Elise Series 2 automobile, (b) illustration of the original design for the vertical suspension uprights, using an aluminum extrusion, (c) retrofit design, using a steel forging, (d) optimized steel forging design for new car models. *Source:* (a) Shutterstock/VanderWolf Images.

Table 14.5: Comparison of Suspension Upright Designs for the Lotus Elise Automobile.

Fig. 14.23 sketch	Material	Application	Mass (kg)	Cost
(b)	Aluminum extrusion, steel bracket, steel bushing, housing	Original design	2.105	85
(c)	Forged steel	Phase I	2.685 (+28%)	27.7 (−67%)
(d)	Forged steel	Phase II	2.493 (+18%)	30.8 (−64%)

Summary

- Forging denotes a family of processes in which deformation of the workpiece is carried out by compressive forces applied through a set of dies. The process is capable of producing a wide variety of structural parts, with favorable characteristics, such as higher strength, and improved toughness, dimensional accuracy, and reliability in service.

- The forging operation can be carried out at room, warm, or high temperatures. Workpiece material behavior during deformation, friction, heat transfer, and material-flow characteristics in the die cavity are important considerations, as are the proper selection of die materials, lubricants, workpiece and die temperatures, forging speeds, and equipment.

- Several defects can develop in a forging if the process is not designed or controlled properly. Computer-aided design and manufacturing techniques are used extensively in die design and manufacturing, preform design, predicting material flow, and avoiding the possibility of internal and external defects during forging.

- A variety of forging machines is available, each with its own capabilities, characteristics, and costs. Forging operations are highly automated, using industrial robots and computer controls.

- Swaging is a type of rotary forging in which a solid rod or a tube is reduced in diameter by the reciprocating radial movement of a set of two or four dies. The process is suitable for producing short or long lengths of bar or tubing, with various internal or external profiles.

- Because die failure has a major economic impact on the operation, die design, material selection, and the specific production method are of critical importance. A variety of die materials and manufacturing methods is available, including advanced material-removal and finishing processes.

Key Terms

Barreling
Closed-die forging
Cogging
Coining
Edging
End grain
Flash
Forgeability
Forging
Fullering
Hammers
Heading
Hot-twist test
Hubbing
Impression-die forging
Incremental forging
Isothermal forging
Net-shape forging
Open-die forging
Orbital forging
Piercing
Precision forging
Presses
Sizing
Swaging
Upsetting

Bibliography

Altan, T., Ngaile, G., and Shen, G. (eds.), **Cold and Hot Forging: Fundamentals and Applications**, ASM International, 2004.
ASM Handbook, Vol. 14A: **Metalworking: Bulk Forming**, ASM International, 2005.
ASM Handbook, Vol. 22A: **Fundamentals of Modeling for Metals Processing**, ASM International, 2009.
ASM Handbook, Vol. 22B: **Metals Process Simulation**, ASM International, 2010.
Boljanovic, V., **Metal Shaping Processes**, Industrial Press, 2009.
Dieter, G.E., Kuhn, H.A., and Semiatin, S.L. (eds.), **Handbook of Workability and Process Design**, ASM International, 2003.
Fundamentals of Tool Design, 6th ed., Society of Manufacturing Engineers, 2010.
Hosford, W.F., and Caddell, R.M., **Metal Forming: Mechanics and Metallurgy**, 4th ed., Cambridge, 2011.
Product Design Guide for Forging, Forging Industry Association, 1997.

Review Questions

14.1. What is the difference between cold, warm, and hot forging?

14.2. Explain the difference between open-die and impression-die forging.

14.3. Explain the difference between fullering, edging, and blocking.

14.4. What is flash? What is its function?

14.5. Why is the intermediate shape of a part important in forging operations?

14.6. Describe the features of a typical forging die.

14.7. Explain what is meant by "load limited," "energy limited," and "stroke limited" as these terms pertain to forging machines.

14.8. What type of parts can be produced by rotary swaging?

14.9. Why is hubbing an attractive alternative to producing simple dies?

14.10. What is the difference between piercing and punching?

14.11. What is a hammer? What are the different kinds of hammers?

14.12. Why is there barreling in upsetting?

14.13. What are the advantages and disadvantages of isothermal forging?

14.14. Why are draft angles needed in forging dies?

14.15. Is a mandrel needed in swaging?

Qualitative Problems

14.16. Describe and explain the factors that influence spread in cogging operations on square billets.

14.17. How can you tell whether a certain part is forged or cast? Explain the features that you would investigate.

14.18. Identify casting design rules, described in Section 12.2, that also can be applied to forging.

14.19. Describe the factors involved in precision forging.

14.20. Why is control of the volume of the blank important in closed-die forging?

14.21. Why are there so many types of forging machines available? Describe the capabilities and limitations of each.

14.22. What are the advantages and limitations of cogging operations? Should cogging be performed hot or cold? Explain.

14.23. Describe your observations concerning Fig. 14.17.

14.24. What are the advantages and limitations of using die inserts? Give some examples.

14.25. Review Fig. 14.6d and explain why internal draft angles are larger than external draft angles. Is this also true for permanent-mold casting?

14.26. Comment on your observations regarding the grain-flow pattern in Fig. 14.13.

14.27. Review Fig. 14.13 and make a sketch that identifies the locations where the strain is highest.

14.28. Describe your observations concerning the control of the final tube thickness in Fig. 14.16.

14.29. By inspecting some forged products, such as hand tools, you will note that the lettering on them is raised rather than sunk. Offer an explanation as to why they are made that way.

14.30. Describe the difficulties involved in defining the term "forgeability" precisely.

14.31. Describe the advantages of servo presses for forging and stamping.

14.32. List the general recommendations you would make for forging materials with limited ductility.

14.33. Which would you recommend, hot forging and heat treating a workpiece, or cold forging it and relying upon strain hardening for strengthening? Explain your answer.

Quantitative Problems

14.34. Take two solid, cylindrical specimens of equal diameter, but different heights, and compress them (frictionless) to the same percent reduction in height. Show that the final diameters will be the same.

14.35. Calculate the room-temperature forging force for a solid, cylindrical workpiece made of 5052-O aluminum that is 100 mm high and 125 mm in diameter and is to be reduced in height by 30%. Let the coefficient of friction be 0.15.

14.36. Using Eq. (14.2), estimate the forging force for the workpiece in Problem 14.35, assuming that it is a complex forging and that the projected area of the flash is 30% greater than the projected area of the forged workpiece.

14.37. To what thickness can a cylinder of 1020 steel that is 20 mm in diameter and 40 mm high be forged in a press that can generate 400 kN?

14.38. In Example 14.1, calculate the forging force, assuming that the material is 1100-O aluminum and that the coefficient of friction is 0.10.

14.39. Using Eq. (14.1), make a plot of the forging force, F, as a function of the radius, r, of the workpiece. Assume that the flow stress, σ_f of the material is constant. Remember that the volume of the material remains constant during forging; thus, as h decreases, r increases.

14.40. How would you go about calculating the punch force required in a hubbing operation, assuming that the material is mild steel and the projected area of the impression is 400 mm^2? Explain clearly. (*Hint:* See Section 2.6 on hardness.)

14.41. A mechanical press is powered by a 25 kW motor and operates at 50 strokes per minute. It uses a flywheel, so that the crankshaft speed does not vary appreciably during the stroke. If the stroke is 160 mm, what is the maximum constant force that can be exerted over the entire stroke length?

14.42. A solid cylindrical specimen, made of a perfectly plastic material, is being upset between flat dies with no friction. The process is being carried out by a falling weight, as in a drop hammer. The downward velocity of the hammer is at a maximum when it first contacts the workpiece and becomes zero when the hammer stops at a certain height of the specimen. Establish quantitative relationships between workpiece height and velocity, and make a qualitative sketch of the velocity profile of the hammer. (*Hint:* The loss in the kinetic energy of the hammer is the plastic work of deformation; thus, there is a direct relationship between workpiece height and velocity.)

14.43. Plot the force vs. reduction in height curve in open-die forging of a cylindrical, annealed Ti-6Al-4V specimen that is 10 mm high and 25 mm in diameter, up to a reduction of 50%, for the cases of (a) 20°C with $\mu = 0.2$, and (b) a workpiece preheated to a temperature of 600°C with $\mu = 0.4$.

14.44. Plot the force vs. reduction in height curve in open-die forging of a cylindrical, annealed Ti-6Al-4V specimen that is 10 mm high and 25 mm in diameter, up to a reduction of 50%, with $\mu = 0.2$, for the cases of (a) a hydraulic press with a speed of 0.1 m/s and (b) a mechanical press with a speed of 1 m/s. Assume the temperature is 800°C for both cases.

14.45. Estimate the force required to upset a 6 mm diameter C74500 brass rivet in order to form a 12 mm diameter head. Assume that the coefficient of friction between the brass and the tool-steel die is 0.25 and that the head is 6 mm in thickness. Use $S_y = 175$ MPa.

14.46. A compressor blade is to be forged of Ti-6Al-4V at 900°C, where $K = 140$ MPa and $n = 0.40$. The volume of the compressor blade is 40,000 mm^3, but the blank is oversized so that 20% of the blank volume will go into flash. In the finishing die, the projected area is 4000 mm^2. Use a flash width of 5 mm, and recognize that the compressor blade is a simple shape, but has some complexity because of the thin sections and the detail at the mounting end. Estimate the required forging force if the largest strain in the finish forging is $\epsilon = 0.25$.

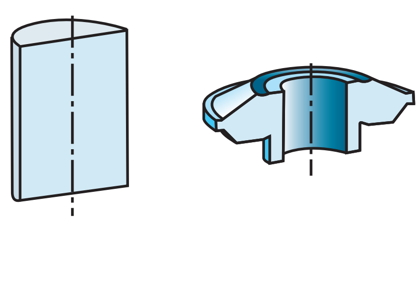
14.47. Assume that you are an instructor covering the topics described in this chapter and you are giving a quiz on the numerical aspects to test the understanding of the students. Prepare two quantitative problems and supply the answers.

Synthesis, Design, and Projects

14.48. Devise an experimental method whereby you can measure only the force required for forging the flash in impression-die forging.

14.49. Assume that you represent the forging industry and that you are speaking with a representative of the casting industry. What would you tell that person about the merits of forging processes?

14.50. The figure below shows a round impression-die forging made from a cylindrical blank, as illustrated on the left. As described in this chapter, such parts are made in a sequence of forging operations. Suggest a sequence of intermediate forging steps to make the part on the right, and sketch the shape of the dies needed.



14.51. In comparing forged parts with cast parts, we have noted that the same part may be made by either process. Comment on the pros and cons of each process, considering factors such as part size, shape complexity, design flexibility, mechanical properties developed, and performance in service.

14.52. From the data given in Table 14.3, obtain the approximate value of the yield strength of the materials listed at hot-forging temperatures. Plot a bar chart showing the maximum diameter of a hot-forged part produced on a press with a 54-metric ton capacity as a function of the material.

14.53. Review the sequence of operations in the production of the stepped pin shown in Fig. 14.14. If the conical-upsetting step is not performed, how would the final part be affected?

14.54. Using a flat piece of wood, perform simple cogging operations on pieces of clay and make observations regarding the spread of the pieces as a function of the original cross sections (for example, square or rectangular with different thickness-to-width ratios).

14.55. Discuss the possible environmental concerns regarding the operations described in this chapter.

14.56. Assume that in upsetting a solid cylindrical specimen between two flat dies with friction, the dies are rotated at opposite directions to each other. How, if at all, will the forging force change from that for nonrotating dies? (*Hint:* Note that the dies will now require torque, because of the change in the direction of frictional forces at the die–workpiece interfaces.)

14.57. List the advantages and disadvantages in using a lubricant in forging operations.

Chapter 15

Metal Extrusion and Drawing Processes and Equipment

- Extrusion and drawing involve pushing or pulling, respectively, a material through a die, for the purpose of modifying its cross section.
- The chapter begins by describing the basic types of extrusion processes, and how the extrusion force can be calculated from material and processing parameters.
- Hot and cold extrusion are then presented, including die design, and describing how cold extrusion is often performed in combination with forging to produce specific part shapes.
- Extrusion practices and die designs to avoid common defects are also presented.

- Drawing of rod, wire, and tubing is in a similar manner, along with die design considerations.
- Equipment characteristics for these processes are also described.

Typical parts made by extrusion and drawing: Long pieces having a wide variety of constant cross sections, rods, shafts, bars for machinery and automotive power-train applications, aluminum ladders, collapsible tubes, wire for numerous electrical and mechanical applications and musical instruments.

Alternative processes: Machining, powder metallurgy, shape rolling, roll forming, pultrusion, additive manufacturing, and continuous casting.

15.1 Introduction

Extrusion and drawing have numerous applications in manufacturing continuous as well as discrete products from a wide variety of metals and alloys. In simple extrusion, a cylindrical billet is forced through a die (Fig. 15.1) in a manner similar to squeezing toothpaste from a tube. A wide variety of solid or hollow cross sections can be produced by extrusion, which basically are semifinished products.

A characteristic of extrusion (from the Latin *extrudere*, meaning to force out) is that large deformations can take place without fracture, because the material is under high triaxial compressive stresses (see Section 2.2.8). Since the die geometry remains unchanged throughout the process, extruded products typically have a constant cross section along their length.

Typical products made by extrusion are railings for sliding doors, window and door frames, tubing, aluminum ladder frames, and structural and architectural shapes. Extrusions can later be cut into desired lengths, which then become discrete parts, such as brackets, small gears, and coat hangers (Fig. 15.2). Commonly extruded materials are aluminum, copper, steel, magnesium, and lead; other metals and alloys also can be extruded, at various levels of difficulty.

Each billet is extruded individually, thus extrusion is a batch or semicontinuous operation. The process can be economical for large as well as short production runs. Tool costs generally are low, particularly for producing simple, solid cross sections. Depending on the ductility of the material, the process can be carried out at room or at elevated temperatures. Extrusion at room temperature is often combined with forging operations, in which case it is called **cold extrusion** (see also Section 14.4), with numerous applications, such as fasteners and components for automobiles, bicycles, motorcycles, machinery, and transportation equipment.

In **drawing**, developed between 1000 and 1500 A.D., the cross section of a solid rod, wire, or tubing is reduced or changed in shape by pulling it through a die. Drawn rods are used for shafts, spindles, and small pistons and as the raw material for fasteners such as rivets, bolts, and screws. In addition to round rods, various profiles also can be drawn.

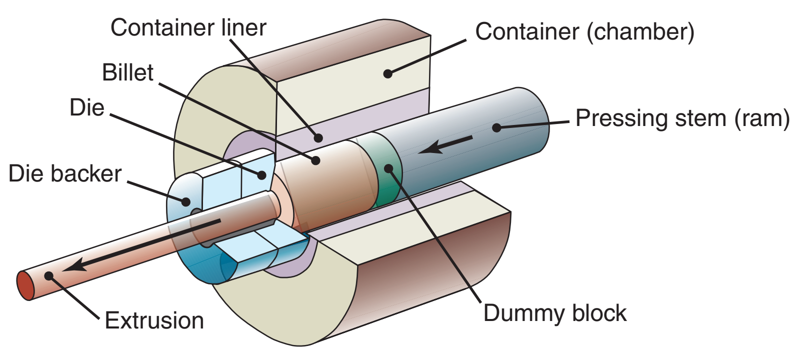


Figure 15.1: Schematic illustration of the direct extrusion process.

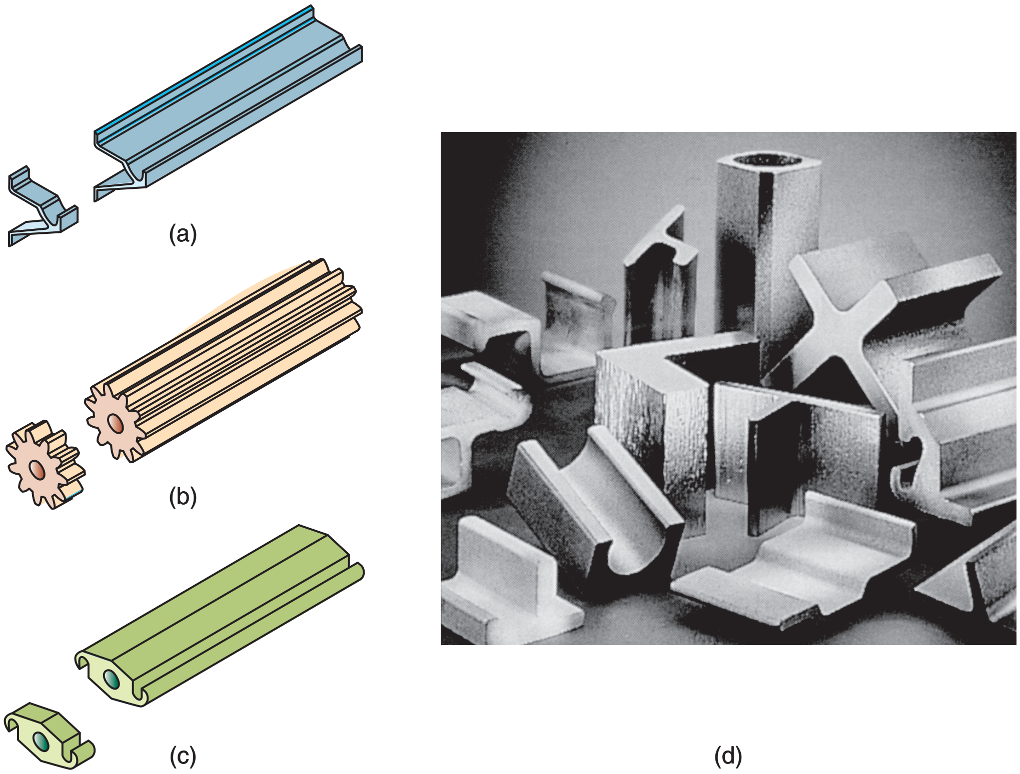


Figure 15.2: Extrusions and examples of products made by sectioning off extrusions. *Source:* Courtesy of Plymouth Engineered Shapes.

The distinction between the terms **rod** and **wire** is somewhat arbitrary, with rod taken to be larger in cross section than wire. In industry, wire is generally defined as a rod that has been drawn through a die at least once, or whose diameter is sufficiently small so that it can be coiled. Wire drawing involves much smaller diameters than rod drawing, with sizes down to 0.01 mm for magnet wire; smaller diameters can be obtained for specialized applications.

15.2 The Extrusion Process

There are three basic types of extrusion processes. In **direct** or **forward extrusion**, a billet is placed in a container (*chamber*) and forced through a die, as shown in Fig. 15.1. The die opening can be round or it may have various shapes, depending on the desired cross section. The function of the *dummy block*, shown in the figure, is to protect the tip of the pressing stem, particularly in hot extrusion.

In **indirect** extrusion, also called *reverse, inverted,* or *backward extrusion*, the die moves toward the stationary billet (Fig. 15.3a). Indirect extrusion has the advantage of having no billet–container friction, since there is no relative motion; thus, it is used on materials with very high friction, such as high-strength steels.

In **hydrostatic extrusion** (Fig. 15.3b), the billet is smaller in diameter than the container, which is filled with a fluid, and the pressure is transmitted to the fluid by a ram. The fluid pressure imparts triaxial compressive stresses on the billet, thus improving its formability (see Section 2.2.8). Furthermore, there is much less workpiece–container friction than in direct extrusion. A less common type of extrusion is *lateral* (or *side*) *extrusion* (Fig. 15.3c).

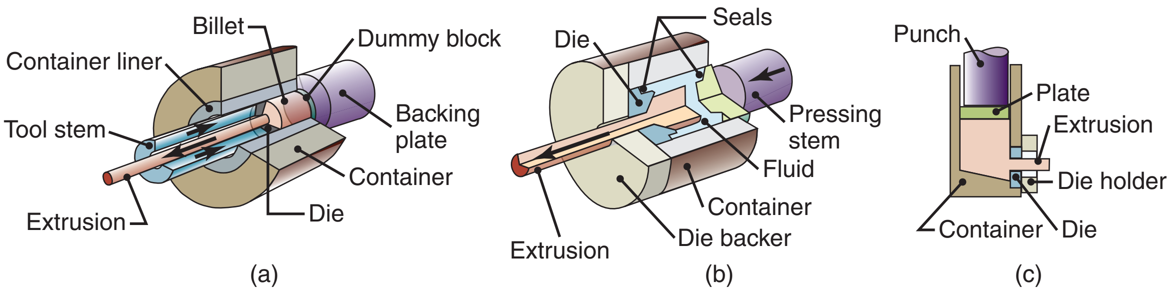


Figure 15.3: Types of extrusion: (a) indirect; (b) hydrostatic; (c) lateral.

In **friction stir extrusion**, the die or, the mandrel in the case of hollow parts, is rotated as the workpiece is being extruded. This has the effect of causing greater frictional heating and redundant work, leading to a hotter and softer workpiece. This process is especially useful for extruding high-strength alloys, but has also been demonstrated as a recycling approach for metal shavings or powder feedstocks.

As can be seen in Fig. 15.4, the basic geometric variables in extrusion are the die angle, α, and the **extrusion ratio,** R (ratio of the cross-sectional area of the billet to that of the extruded part, A_o/A_f); additional processing variables are the billet temperature, the speed at which the ram travels towards the die, and the type of lubricant used, if any.

Extrusion Force. The force required for extrusion depends on (a) the strength of the billet material, (b) extrusion ratio, (c) friction between the billet, container, and die surfaces, and (d) processing variables. It has been shown that for a small die angle, α, the extrusion pressure can be approximated as

$$p = S_y\left(1+\frac{\tan\alpha}{\mu}\right)\left(R^{\mu\cot\alpha}-1\right), \tag{15.1}$$

where μ is the coefficient of friction, S_y is the yield strength of the billet material, and R is the extrusion ratio. The extrusion force can then be obtained by multiplying the pressure by the billet area, and can be simplified as

$$F = A_o k \ln\left(\frac{A_o}{A_f}\right), \tag{15.2}$$

where k is the *extrusion constant*, determined experimentally; thus k is a measure of the strength of the material being extruded and the frictional conditions. Figure 15.5 gives k for several metals and a range of extrusion temperatures.

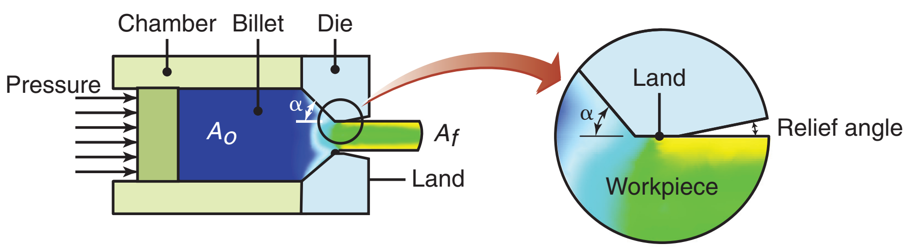


Figure 15.4: Process variables in direct extrusion. The die angle, reduction in cross section, extrusion speed, billet temperature, and lubrication all affect the extrusion pressure. The contour plot shows effective strain as obtained from a finite element simulation for the geometry shown.

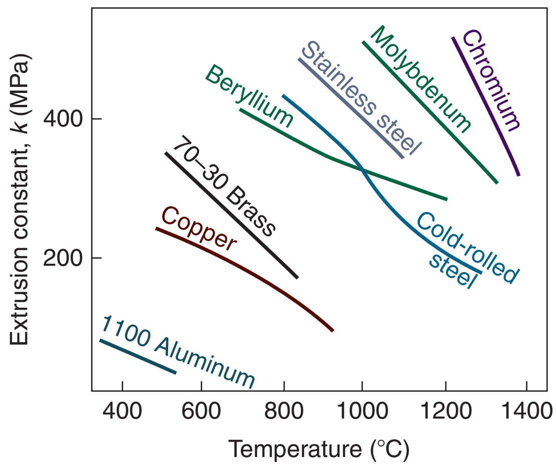


Figure 15.5: Extrusion constant k for various metals at different temperatures, as determined experimentally. *Source:* After P. Loewenstein.

Example 15.1 Calculation of Force in Hot Extrusion

Given: A round billet made of 70–30 brass is extruded at a temperature of 675°C. The billet diameter is 125 mm, and the diameter of the extrusion is 50 mm.

Find: Calculate the extrusion force required.

Solution: The extrusion force is calculated using Eq. (15.2), in which the extrusion constant, k, is obtained from Fig. 15.5. For 70–30 brass, $k = 250$ MPa at the given extrusion temperature. Thus,

$$F = \pi(0.125/2)^2(250 \times 10^6) \ln\left[\frac{\pi(125)^2}{\pi(0.050)^2}\right] = 5.5 \text{ MN}.$$

Metal Flow in Extrusion. The metal flow pattern in extrusion, as in other forming processes, is important because of its influence on the quality and the final properties of the extruded product. The material flows longitudinally, much like an incompressible fluid flows in a channel; thus, extruded products have an elongated grain structure (*preferred orientation*, Section 1.6). Improper metal flow during extrusion can produce various defects in the extruded product, as described in Section 15.5.

A common technique for investigating the flow pattern is to cut the round billet lengthwise in half and mark one face with a square grid pattern. The two halves are then placed together in the chamber and are extruded. Figure 15.6 shows typical flow patterns obtained by this technique, for the case of direct extrusion with square dies (90° die angle).

The conditions under which these different flow patterns occur are described in the caption of Fig. 15.6. Note the **dead-metal zone** in Fig. 15.6b and c, where the metal at the corners essentially remains stationary.

Processing Parameters. In practice, extrusion ratios usually range from about 10 to 100; they may be higher for special applications (such as 400 for softer nonferrous metals) or lower for less ductile materials. The ratio usually has to be at least 4 to deform the material plastically through the bulk of the workpiece.

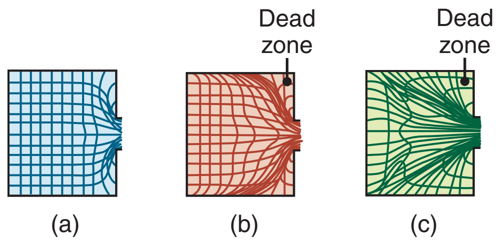


Figure 15.6: Types of metal flow in extruding with square dies. (a) Flow pattern obtained at low friction or in indirect extrusion. (b) Pattern obtained with high friction at the billet–chamber interfaces. (c) Pattern obtained at high friction or with cooling of the outer regions of the billet in the chamber. This type of pattern, observed in metals whereby their strength increases rapidly with decreasing temperature, leads to a defect known as *pipe*, or *extrusion defect*.

Extruded products are typically less than 7.5 m long, because of the difficulty in handling greater lengths; they can be as long as 30 m. Ram speeds range up to 0.5 m/s; generally, lower speeds are preferred for aluminum, magnesium, and copper, and higher for steels, titanium, and refractory alloys. Dimensional tolerances are usually in the range from ± 0.25 to 2.5 mm; they increase with increasing cross sectional area.

Because they have high ductility, aluminum, copper, and magnesium and their alloys, and steels and stainless steels, are extruded with relative ease into various cross sections. Metals such as titanium and refractory metals also can be extruded, but only with some difficulty and significant die wear.

Most extruded products, particularly those with small cross sections, require subsequent straightening and twisting. This is typically done in a hydraulic stretcher equipped with jaws.

The presence of a die angle causes a small portion at the end of the billet to remain in the chamber at the end of the ram stroke. This portion, called *scrap* or the *butt end*, is later removed by cutting it off. Alternatively, a graphite block or another billet may be placed in the chamber behind the previous extrusion.

In **coaxial extrusion**, coaxial billets are extruded together, provided that the strength and ductility of the two metals are compatible. *Stepped extrusions* also are produced, by extruding the billet partially in one die and then in successively larger dies (see also *cold extrusion*, Section 15.4). *Lateral extrusion* (Fig. 15.3c) is used for the sheathing of wire and the coating of electric wire with plastic.

15.3 Hot Extrusion

For metals and alloys that do not have sufficient ductility at room temperature, or in order to reduce the forces required, extrusion is carried out at elevated temperatures (Table 15.1). As in all other elevated-temperature operations, hot extrusion has special requirements because of the high operating temperatures involved. For example, die wear can be excessive, and cooling of the hot billet's surfaces (in the cooler container) and the die can result in highly nonuniform deformation of the billet, as shown in Fig. 15.6c. Thus, extrusion dies may be preheated, as is also done in hot-forging operations (Section 14.1).

Because the billet is hot, it develops an oxide film, unless it is heated in an inert environment. Oxides can be abrasive (see Section 33.2), and can affect the flow pattern of the material. Their presence also results in an extruded product that may be unacceptable when good surface finish is required. To avoid forming of oxide films on the hot extruded product, the dummy block placed ahead of the ram (Fig. 15.1) is made a little smaller in diameter than the container. As a result, a thin shell (*skull*), consisting mainly of the outer oxidized layer of the billet, is left in the container; it is later removed from the chamber.

Table 15.1: Typical Extrusion Temperature Ranges for Various Metals and Alloys (see also Table 14.3).

Material	Extrusion temperature, °C
Lead	200–250
Aluminum and its alloys	375–475
Copper and its alloys	650–975
Steels	875–1300
Refractory alloys	975–2200

Die Design. Die design requires considerable experience, as can be appreciated by reviewing Fig. 15.7. *Square dies*, also called *shear dies*, are used in extruding nonferrous metals, especially aluminum. These dies develop *dead-metal zones*, which in turn form an effective die angle (see Fig. 15.6b and c) along which the material flows. These zones produce extrusions with bright finishes, because of the burnishing action (Section 16.2) that takes place as the material flows past the die; the workpiece also is initially oxide-free.

Tubing can be extruded from a solid or hollow billet (Fig. 15.8). Wall thickness is usually limited to 1 mm for aluminum, 3 mm for carbon steels, and 5 mm for stainless steels. When solid billets are used, the ram is fitted with a mandrel that pierces a hole into the billet. Billets with a previously pierced hole also may be extruded in this manner. Because of friction and the severity of deformation, thin-walled extrusions are more difficult to produce than those with thick walls.

Hollow cross sections (Fig. 15.9a) can be extruded by *welding-chamber* methods and a **porthole die**, **spider die**, or **bridge die** (Fig. 15.9b to d). During extrusion, the metal divides and flows around the supports for the internal mandrel into strands; this is a condition much like that of air or water flowing around an object and rejoining downstream. The strands being extruded then become rewelded, under the high pressure in the welding chamber, before exiting the die. The rewelded surfaces have good strength because they have not been exposed to the environment; otherwise, they would develop oxides on their surfaces, thereby inhibiting good welding. The welding-chamber process is suitable only for aluminum and some of its alloys, because they can develop a strong weld under high pressure, as described in Section 31.2. Lubricants cannot be used because they prevent rewelding of the metal surfaces within the die.

Die Materials. Die materials for hot extrusion usually are hot-work die steels (Section 5.7). Coatings, such as partially stabilized zirconia (PSZ), may be applied to the die surfaces to extend their life. Dies made of

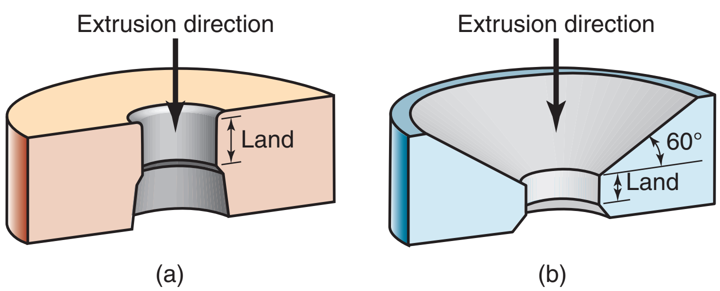


Figure 15.7: Typical extrusion–die configurations: (a) die for nonferrous metals; (b) die for ferrous metals; (c) die for a T-shaped extrusion, made of hot-work die steel and used with molten glass as a lubricant.

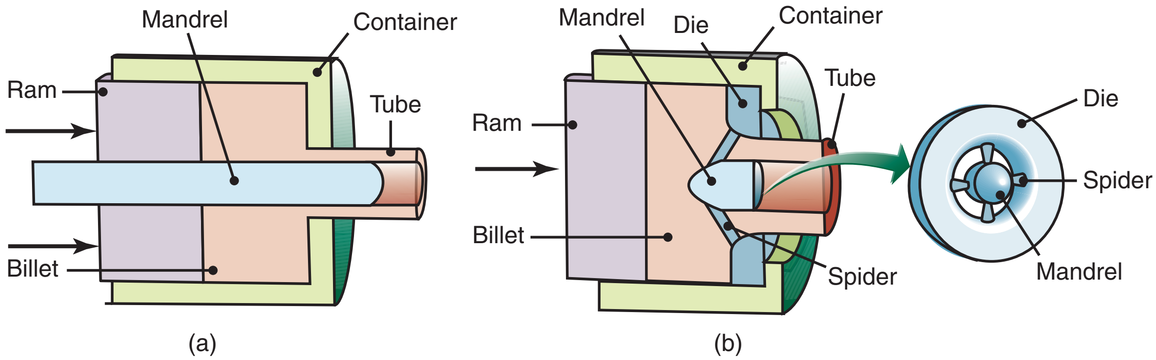


Figure 15.8: Extrusion of a seamless tube (a) using an internal mandrel that moves independently of the ram; an alternative arrangement has the mandrel integral with the ram, (b) using a spider die (see Fig. 15.9) to produce seamless tubing.

PSZ (Section 8.2.2) also are used for hot extrusion of tubes and rods. However, they are not suitable for extruding complex shapes, because of the severe stress gradients that develop in the die, possibly leading to their premature failure.

Lubrication. Lubrication is important in hot extrusion, because of its effects on (a) material flow during extrusion, (b) surface finish and integrity, (c) product quality, and (d) extrusion forces. *Glass* (Section 8.4) is an excellent lubricant for hot extrusion of steels, stainless steels, and high-temperature metals and alloys. In a process developed in the 1940s and known as the **Séjournet process**, a circular glass or fiberglass *pad* is placed in the chamber at the die entrance. The hot billet conducts heat to the glass pad, whereupon a thin layer of glass melts, is entrained, and acts as a lubricant. Before the hot billet is directly placed in the chamber, its cylindrical surface is coated with a layer of powdered glass, to develop a thin glass lubricant layer at the billet–chamber interface.

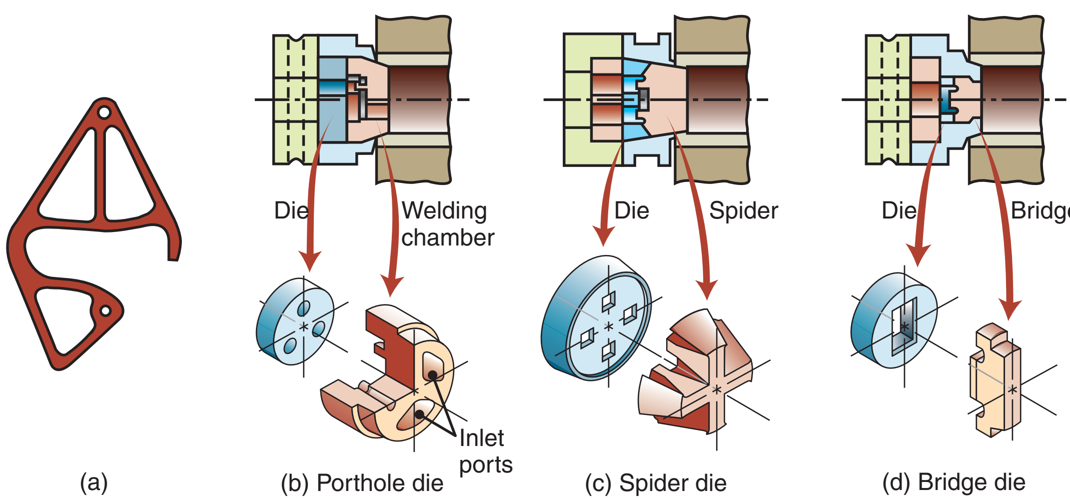


Figure 15.9: (a) An extruded 6063-T6 aluminum-ladder lock for aluminum extension ladders. This part is 8 mm thick and is sawed from the extrusion (see Fig. 15.2). (b) through (d) Components of various dies for extruding intricate hollow shapes. *Source:* (b) through (d) after K. Laue and H. Stenger.

For metals that have a tendency to stick or even weld to the container and the die surfaces, the billet can be enclosed in a *jacket*, a thin-walled container made of a softer and lower strength metal, such as copper or mild steel; this procedure is called **jacketing** or **canning**. In addition to acting as a low-friction interface, the jacket prevents contamination of the billet by the environment. For billet materials that are toxic or radioactive, the jacket also prevents it from contaminating the environment.

Case Study 15.1 Manufacture of Aluminum Heat Sinks

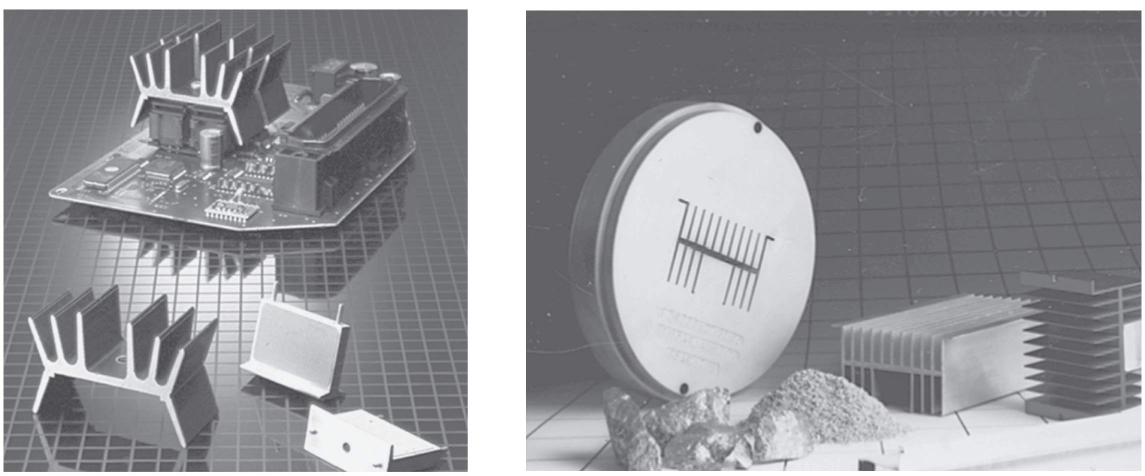
Aluminum is used widely to transfer heat for both cooling and heating applications, because of its very high thermal conductivity. In fact, on a weight-to-cost basis, no other material conducts heat as efficiently as does aluminum.

Hot extrusion of aluminum is preferred for heat-sink applications, such as those in the electronics industry. Fig. 15.10a shows an extruded heat sink, used for removing heat from a transformer on a printed circuit board. Heat sinks usually are designed with a large number of fins that maximize the surface area and assist in heat transfer to a cooler fluid flowing over them. The fins are very difficult and expensive to machine, forge, or roll form, but they can be made economically by hot extrusion, using dies made by electrical-discharge machining (Section 27.5).

Fig. 15.10b shows a die and a typical hot-extruded heat sink cross section. The shapes shown also could be produced through a casting operation, but extrusion is preferred because there is no internal porosity in the part and its thermal conductivity is thus higher.

15.4 Cold Extrusion

Developed in the 1940s, *cold extrusion* is a general term often denoting a *combination* of operations, such as a combination of direct and indirect *extrusion and forging* (Fig. 15.11). Cold extrusion is used widely for components in automobiles, motorcycles, bicycles, appliances, and in transportation and agricultural equipment.



(a) (b)

Figure 15.10: (a) Aluminum extrusion used as a heat sink for a printed circuit board, (b) extrusion die and extruded heat sinks. *Source:* Courtesy of Aluminum Extruders Council.

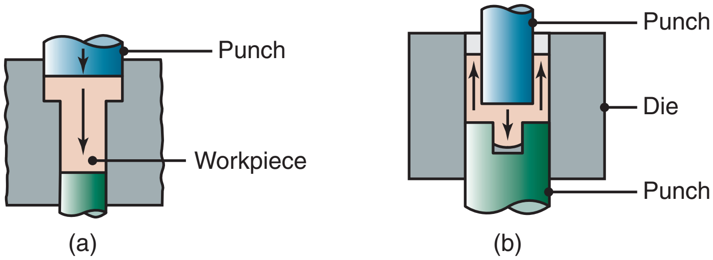


Figure 15.11: Two examples of cold extrusion; arrows indicate the direction of metal flow during extrusion.

The cold extrusion process uses slugs cut or sheared from cold-finished or hot-rolled bars, wire, or plates. Slugs that are less than about 40 mm in diameter are sheared (*cropped*), and, if necessary, their ends are squared off by processes such as upsetting, machining, or grinding. Larger diameter slugs are machined from bars into specific lengths. Cold-extruded parts weighing as much as 45 kg and having lengths of up to 2 m can be made, although most parts weigh much less. Powder-metal slugs (preforms) also may be cold extruded (Section 17.3.3).

The *force*, F, in cold extrusion may be estimated from the formula

$$F = 1.7A_o\sigma_f\epsilon, \tag{15.3}$$

where A_o is the cross-sectional area of the blank, σ_f is the average flow stress of the metal, and ϵ is the true strain that the piece undergoes, based on its original and final cross-sectional area. For example, assume that a round slug 10 mm in diameter and made of a metal with $\sigma_f = 300$ MPa that is reduced to a final diameter of 7 mm by cold extrusion. The force would be

$$F = 1.7(\pi)\left(\frac{0.010^2}{4}\right)\left(300 \times 10^6\right)\left[\ln\left(\frac{10}{7}\right)^2\right] = 28.6 \text{ kN}.$$

Cold extrusion has the following advantages over hot extrusion:

- Improved mechanical properties, resulting from work hardening, provided that the heat generated by plastic deformation and friction does not recrystallize the extruded metal.
- Good control of dimensional tolerances, thus reducing the need for subsequent machining or finishing operations.
- Improved surface finish, due partly to the absence of an oxide film and provided that lubrication is effective.
- Production rates and costs are competitive with those of other methods of producing the same part. Some machines are capable of producing more than 2000 parts per hour.

On the other hand, the stresses acting on the tooling in cold extrusion is very high (especially with steel and specialty-alloys), being on the order of the hardness of the workpiece material. The punch hardness usually ranges between 60 and 65 HRC, and the die hardness between 58 and 62 HRC. Punches are a critical component in cold extrusion, as they must possess not only sufficient strength but also high toughness and resistance to wear and fatigue failure. *Lubrication* is critical, especially with steels, because of the possibility

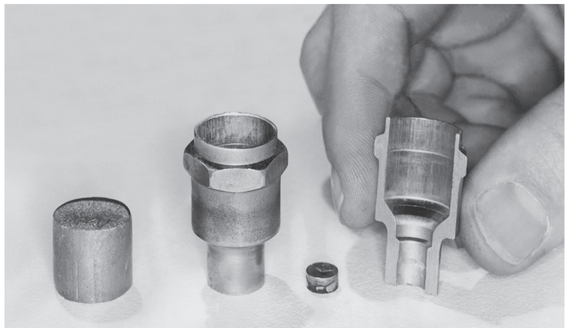


Figure 15.12: Production steps for the metal portion of a cold-extruded spark plug. *Source:* Courtesy of National Machinery Company, LLC.

of sticking (*seizure*) between the workpiece and the tooling. The most effective means of lubrication is the application of a *phosphate-conversion coating* over the workpiece surfaces, followed by a coating of soap or wax, as described in Section 34.10.

Tooling design and the selection of appropriate tool and die materials are essential to the success of cold extrusion. Also important are the selection and control of the workpiece material with regard to its quality and the repeated accuracy of the slug dimensions and its surface condition.

Case Study 15.2 Cold-extruded Part

A typical cold-extruded part, similar to the metal component of an automotive spark plug, is shown in Fig. 15.12. First, a slug is sheared off (*cropped*) the end of a round rod (Fig. 15.12, left). It then is cold extruded (Fig. 15.12, middle) in an operation similar to those shown in Fig. 15.11, but with a blind hole. Then the material at its bottom is punched out, producing the small slug shown. Note the respective diameters of the slug and the hole at the bottom of the sectioned part.

Investigating material flow during the deformation of the slug helps avoid defects and leads to improvements in punch and die design. The part usually is sectioned in the midplane, and then polished and etched to display the grain flow, as shown in Fig. 15.13 (see also Fig. 14.13).

15.4.1 Impact Extrusion

Impact extrusion is similar to indirect extrusion, and the process often is included in the cold-extrusion category. The punch descends rapidly on the blank (*slug*), which is extruded backwards (Fig. 15.14). Because of volume constancy, the thickness of the tubular extruded region is a function of the clearance between the punch and the die cavity.

Typical products made by this process are shown in Fig. 15.15a to c. Other examples are collapsible tubes, some cylindrical computer housings, light fixtures, automotive parts, and small pressure vessels. Most nonferrous metals can be impact extruded in vertical presses and at production rates as high as two parts per second.

The maximum diameter of parts made is about 150 mm. The impact-extrusion process can produce thin-walled tubular sections, with thickness-to-diameter ratios as small low as 0.005. Consequently, the symmetry of the part and the concentricity of the punch and the blank are important.

Figure 15.13: A cross section of the metal part in Fig. 15.12, showing the grain-flow pattern. *Source:* Courtesy of National Machinery Company, LLC.

15.4.2 Hydrostatic Extrusion

In *hydrostatic extrusion*, the pressure required in the chamber is supplied via a piston and through an incompressible fluid medium surrounding the billet (Fig. 15.3b). Pressures are typically on the order of 1400 MPa. The high pressure in the chamber transmits some of the fluid to the die surfaces, where it significantly reduces friction. Hydrostatic extrusion is usually carried out at room temperature, typically using vegetable oils as the fluid.

Brittle materials can be extruded successfully by this method, because the hydrostatic pressure, along with low friction and the use of small die angles and high extrusion ratios, increases the ductility of the material (Section 2.2.8). Hydrostatic extrusion has had limited industrial applications, mainly because of the complex nature of the tooling, the design of specialized equipment, and the long cycle times required, making the process uneconomical for most materials and applications.

15.5 Extrusion Defects

Depending on workpiece material condition and process variables, extruded products can develop several types of defects that can affect significantly their strength and product quality. Some defects are visible to

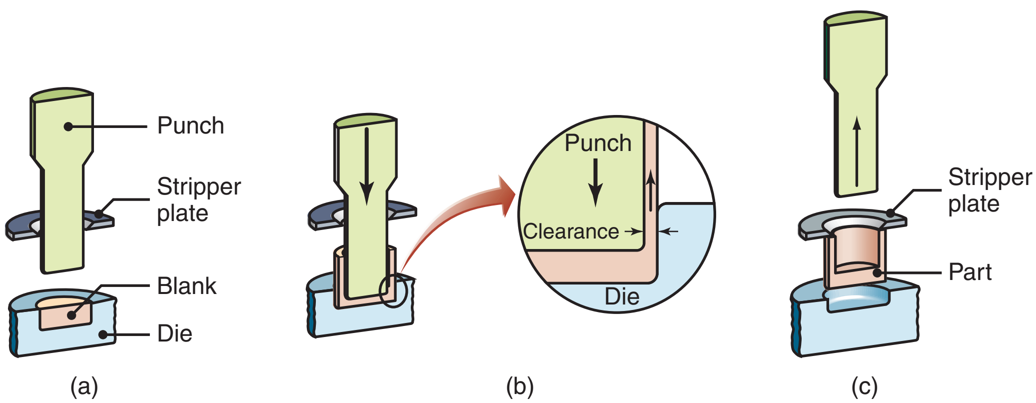


Figure 15.14: Schematic illustration of the impact-extrusion process. The extruded parts are stripped by the use of a stripper plate, because they tend to stick to the punch.

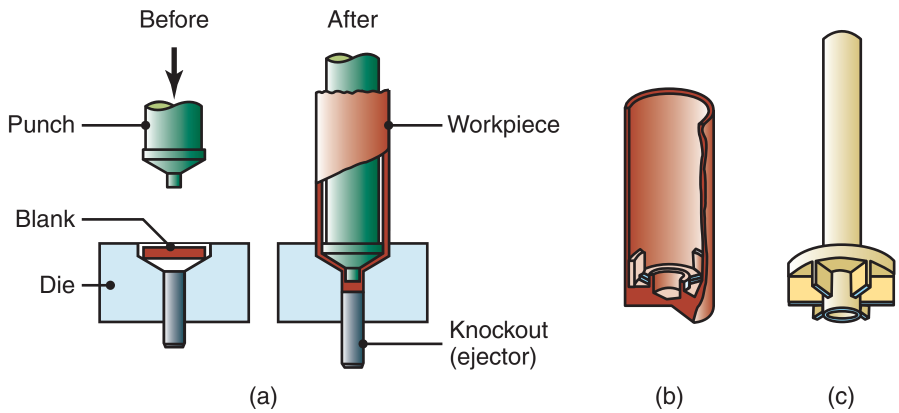


Figure 15.15: (a) Impact extrusion of a collapsible tube by the *Hooker process*. (b) and (c) Two examples of products made by impact extrusion. These parts also may be made by casting, forging, or machining. The choice of process depends on the materials involved, part dimensions and wall thickness, and the properties desired. Economic considerations also are important in final process selection.

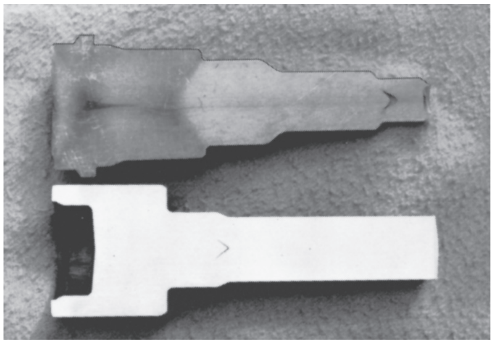
the naked eye, while others can be detected only by the techniques described in Section 36.10. There are three principal *extrusion defects* of importance: surface cracking, pipe, and internal cracking.

Surface Cracking. If extrusion temperature, friction, and speed are too high, surface temperatures can rise significantly, which may cause surface cracking and tearing (*fir-tree cracking* or *speed cracking*). These cracks are intergranular (along the grain boundaries; see Fig. 2.27), and usually are caused by **hot shortness** (Section 1.5.2). Such defects occur especially in aluminum, magnesium, and zinc alloys; they can be avoided by lowering the billet temperature and the extrusion speed.

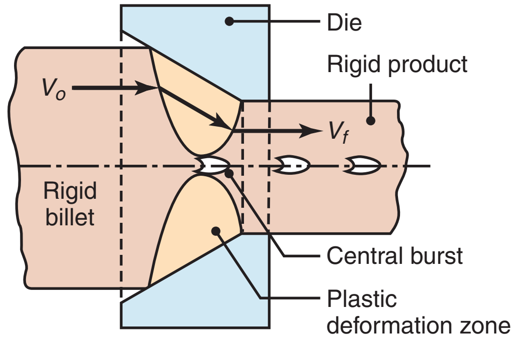
Surface cracking may occur also at lower temperatures, attributed to periodic sticking of the extruded part along the die land. Because of its similarity in appearance to the surface of a bamboo stem, it is known as a **bamboo defect**. The explanation is that, when the product being extruded temporarily sticks to the die land (see Fig. 15.7), the extrusion pressure increases rapidly; shortly thereafter, it moves forward again, and the pressure is released. The cycle is repeated continually, producing periodic circumferential cracks on the surface.

Pipe. The type of metal-flow pattern in extrusion shown in Fig. 15.6c tends to draw surface oxides and impurities toward the center of the billet, much like a funnel. This defect is known as *pipe defect*, *tailpipe*, or *fishtailing*; as much as one-third of the length of the extruded product may contain this type of defect, and has to be cut off as scrap. Piping can be minimized by modifying the flow pattern to be more uniform, such as by controlling friction and minimizing temperature gradients within the part. Another method is to machine the billet's surface prior to extrusion (so that scale and surface impurities are removed) or by chemical etching of the surface oxides prior to extrusion.

Internal Cracking. The center of the extruded product can develop cracks, variously called *center cracking*, *center-burst*, *arrowhead fracture*, or *chevron cracking* (Fig. 15.16a). Cracking has been attributed to a state of hydrostatic tensile stress that develops at the centerline in the deformation zone in the die (Fig. 15.16b); this condition is similar to the necked region in a tensile-test specimen (see Fig. 2.23). Such cracks also have been observed in tube extrusion and in tube spinning (see Fig. 16.49b and c), appearing on the *inside* surfaces of tubes. The tendency for center cracking (a) increases with increasing die angle, (b) increases with increasing amount of impurities in the material, and (c) decreases with increasing extrusion ratio and friction.



(a)



(b)

Figure 15.16: (a) Chevron cracking (central burst) in extruded round steel bars. Unless the products are inspected, such internal defects may remain undetected and later cause total failure of the part in service. This defect can also develop in the drawing of rod, wire, and tubes. (b) Schematic illustration of rigid and plastic zones in extrusion. The tendency toward chevron cracking increases if the two plastic zones do not meet. Note that the plastic zone can be made larger either by decreasing the die angle, increasing the reduction in cross section, or both. *Source:* After B. Avitzur.

15.6 Design Considerations

Extrusion of *constant cross sections* is often a more economical method of producing a part than by forging, casting, or machining. While there is considerable freedom in designing the cross sections, there are several general rules that should be followed to simplify production and reduce defects. Before laying out the cross section, the designer should consider the following:

- Some guidelines for proper die design in extrusion are illustrated in Fig. 15.17. Note the (a) importance of symmetry of the cross section, (b) avoiding sharp corners, (c) maintaining uniform wall thickness, and (d) avoiding severe changes in die dimensions within the cross section.
- Solid shapes are the easiest to extrude. When possible, the cross section should avoid hollow sections, although such sections can be extruded using porthole, bridge, or spider dies, as illustrated in Fig. 15.9.
- If there is a critical dimension in a cross section, it should not be located at the end of a gap. Figure 15.17 shows the use of a metal web to decrease the tolerance on a critical dimension. Note that this design approach requires the extrusion of a hollow cross section; if the cross section is complex, it can be extruded in two sections, and then assembled using the geometries shown in Fig. 15.18.
- Extrusions will usually develop some curvature, which may require straightening. Wide, thin sections can be difficult to straighten, hence the need for ribs as shown in Fig. 15.17.

Impact extrusions should incorporate the following considerations:

- They should be symmetrical about the punch. External and internal bosses can be used as long as they are in the part axis.
- The maximum length-to-diameter ratio should not exceed 8 or so, to avoid punch failure.
- For reverse extrusion, the outer radius can be small, but the inner radius should be as small as possible, and should preferably incorporate a chamfer.

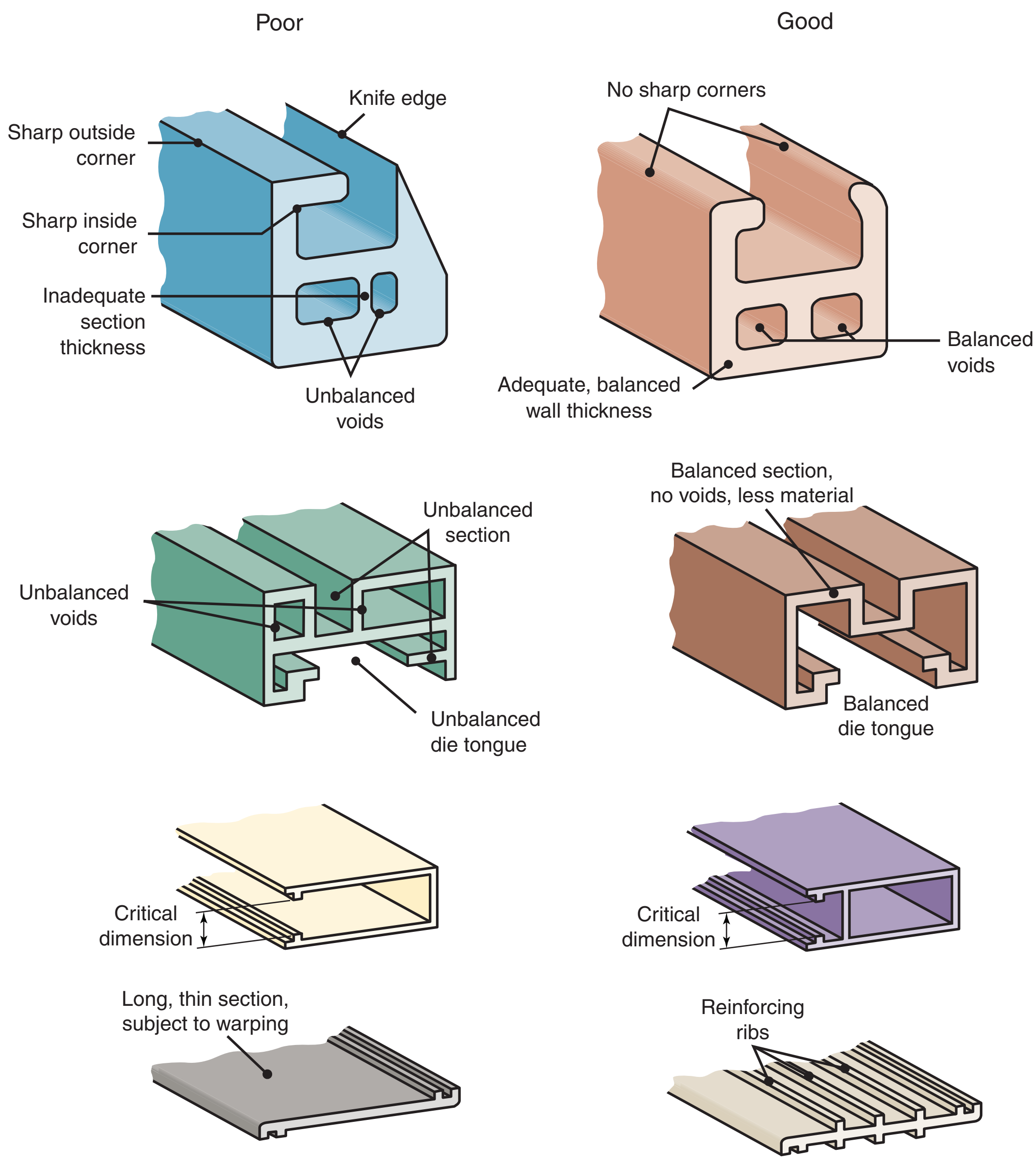


Figure 15.17: Examples of poor and good design practices for extrusion. Note the importance of eliminating sharp corners and of keeping section thicknesses uniform.

15.7 Extrusion Equipment

The basic equipment for extrusion is a *horizontal hydraulic press* (Fig. 14.19d). These presses are suitable for extrusion because the stroke and speed of the operation can be controlled, and they are capable of applying a constant force over a long stroke. Consequently, long billets can be used, correspondingly larger extrusions can be produced per setup, and the production rate is thus increased. Hydraulic presses with a ram-force capacity as high as 120 MN (12,600 metric tons) have been built for hot extrusion of large-diameter billets.

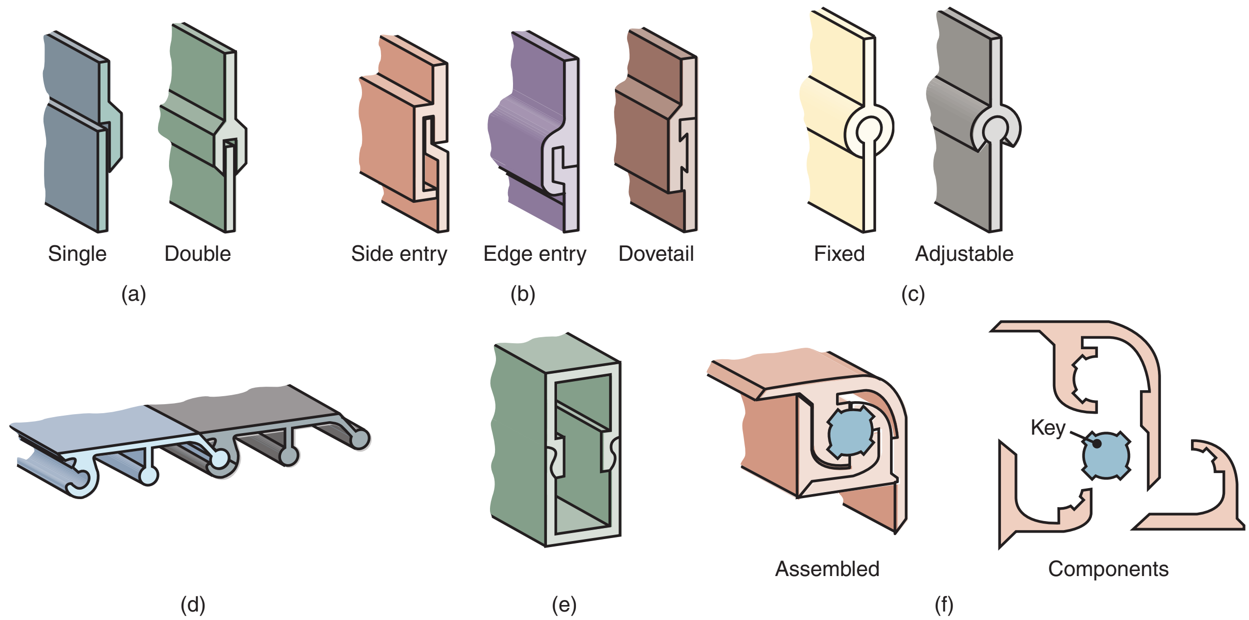


Figure 15.18: Examples of part geometries that allow assembly of extruded sections. (a) Lap joints; (b) lap-lock joints; (c) cylindrical sliding fits; (d) cylindrical sliding lock joints; (e) snap fit; (f) keyed assembly.

Vertical hydraulic presses typically are used for cold extrusion, and generally have lower capacity than those for hot extrusion, but they take up less floor space. *Crank-joint* and *knuckle-joint* mechanical presses (Fig. 14.19a and b) are used for cold extrusion and for impact extrusion to mass-produce small components. Multistage operations, where the cross-sectional area is reduced in a number of individual steps, are carried out on specially designed presses.

Figure 15.19: A 27-MN (2721-metric ton) Sutton aluminum extrusion press. This is the first of the Sutton MK-V extrusion press series built in the United States to SMS group Engineering specifications. Photo courtesy of SMS Group, Inc., Pittsburgh, PA.

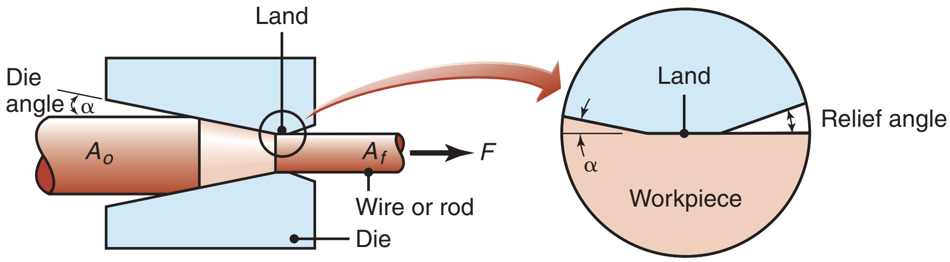


Figure 15.20: Process variables in wire drawing. The major processing variables in drawing are similar to those in extrusion, that is, reduction in cross-sectional area, die angle, frictional conditions along the die–workpiece interfaces, and drawing speed.

15.8 The Drawing Process

In *drawing*, the cross section of a rod or wire is reduced or changed in shape by pulling (hence the term drawing) it through a die, called a *draw die* (Fig. 15.20). Drawn rod and wire products cover a very wide range of applications, including shafts for power transmission, machine and structural components, blanks for bolts and rivets, electrical wiring, cables, tension-loaded structural members, welding electrodes, springs, paper clips, spokes for bicycle wheels, and stringed musical instruments.

Drawing Force. The expression for the *drawing force*, F, under *ideal and frictionless* conditions is similar to that for extrusion, and is given by the equation

$$F = \sigma_f A_f \ln\left(\frac{A_o}{A_f}\right), \tag{15.4}$$

where σ_f is the average true stress of the material in the die gap. Since more work has to be done to overcome friction, the force increases with increasing friction. Furthermore, because of nonuniform deformation within the die zone, additional energy, known as the *redundant work of deformation* is required. Although several equations have been developed over the years to estimate the force (described in greater detail in advanced texts), a formula that includes friction and the redundant work is

$$F = \sigma_f A_f \left[\left(1 + \frac{\mu}{\alpha}\right)\ln\left(\frac{A_o}{A_f}\right) + \frac{2}{3}\alpha\right], \tag{15.5}$$

where α is the die angle, in radians.

As can be seen from the two equations above, the drawing force increases as reduction increases. However, there is a limit to the drawing force, because when the drawing stress reaches the yield strength of the metal drawn, the wire will yield and eventually fracture. It can be shown that, *ideally* and *without friction*, the maximum reduction in cross-sectional area per pass is 63%. Thus, for example, a 10-mm-diameter rod can be reduced to a diameter of 6.1 mm in one pass without failure.

It can also be shown that, for a certain reduction in diameter and a frictional condition, there is an *optimum die angle* at which the drawing force is a minimum. Often, however, the die force is not the major product quality concern, and die angle in practice may deviate from this value.

Drawing of other Shapes. Solid cross sections can be produced by drawing through dies with various profiles. Proper die design and the selection of reduction sequence per pass require considerable experience to ensure proper material flow in the die, reduce the development of internal or external defects, and improve surface quality.

The wall thickness, diameter, or shape of tubes that have been produced by extrusion or by other processes described in this book can be further reduced by *tube drawing* processes (Fig. 15.20). Tubes as large as 0.3 m in diameter can be drawn by these techniques.

Wedge-shaped dies are used for drawing *flat strips*. Although practiced only in specific applications, the principle behind this process is the fundamental deformation mechanism in **ironing**, used extensively in making aluminum beverage cans, as shown in Fig. 16.31.

15.9 Drawing Practice

Successful drawing requires proper selection of process parameters. Reductions in cross-sectional area per pass range up to about 45%; usually, the smaller the initial diameter, the smaller the reduction per pass. Fine wires are drawn at 15–25% reduction per pass, and larger sizes at 20–45%. Reductions higher than 45% may result in lubricant breakdown, leading to deterioration of surface finish. Although most drawing is done at room temperature, drawing large solid or hollow sections can be done at elevated temperatures in order to reduce forces.

A light reduction, known as **sizing pass**, may be taken on rods to improve their surface finish and dimensional accuracy. However, because light reductions basically deform only the surface layers, they usually produce highly nonuniform deformation of the material and its microstructure. Consequently, the local properties of the material will vary with radial distance within the cross section.

Note in Fig. 15.19 that a rod or wire has to have its tip reduced in cross section in order to be fed through the die opening to be pulled. Typically, this is done by **swaging** (see Section 14.4) the tip of the rod or wire in a manner similar to that shown in Fig. 14.15a and b, in an operation called *pointing*.

Drawing speeds depend on the material and the reduction in cross-sectional area. They may range from 1 to 2.5 m/s for heavy sections to as much as 50 m/s for very fine wire, such as that used for electromagnets. Because the product does not have sufficient time to dissipate the heat generated in drawing, temperatures can rise significantly at high drawing speeds, with detrimental effects on product quality, such as surface finish and dimensional tolerances.

Drawn copper and brass wires are designated by their *temper*, such as 1/4 hard and 1/2 hard, because of work hardening (see Section 1.4.2). *Intermediate annealing* between passes may be necessary to maintain sufficient ductility of the material during cold drawing. High-carbon steel wires, for springs and musical instruments, are made by **patenting**; this is a heat treating operation on the drawn wire, whereby the microstructure developed becomes fine pearlite (see Fig. 4.9). The wires have ultimate tensile strengths as high as 5 GPa, with a tensile reduction of area of about 20%.

Bundle Drawing. Although very fine wire can be made by drawing, the cost can be high because the volume of metal produced per unit time is low. One method employed to increase productivity is to draw several wires simultaneously as a *bundle*. The interfaces between a hundred or more of such wires are kept separate from one another by a suitable metallic material, with similar properties but with lower chemical resistance, so that it subsequently can be *leached out* from the drawn wire surfaces.

Bundle drawing produces wires with cross sections that are somewhat polygonal, rather than round. The wires produced can be as small as 4 μm in diameter; they can be made from such materials as stainless steels, titanium, and high-temperature alloys. Techniques have been developed to produce fine wire that is subsequently broken or chopped into various sizes and shapes. These wires are then used in applications such as electrically conductive plastics, heat-resistant and electrically conductive textiles, filter media, radar camouflage, and medical implants.

Die Design. The characteristic features of a typical drawing die are shown in Fig. 15.21. Note that there are two different angles, entering and approach; approach angles usually range from 6° to 15°, with the entering angle usually larger. The bell and the entering angles are used to control lubricant supply and the thickness of the film. The purpose of the bearing surface (*land*) is to set the final diameter of the product (*sizing*), and also to maintain this diameter even when the die–workpiece interface wears over time.

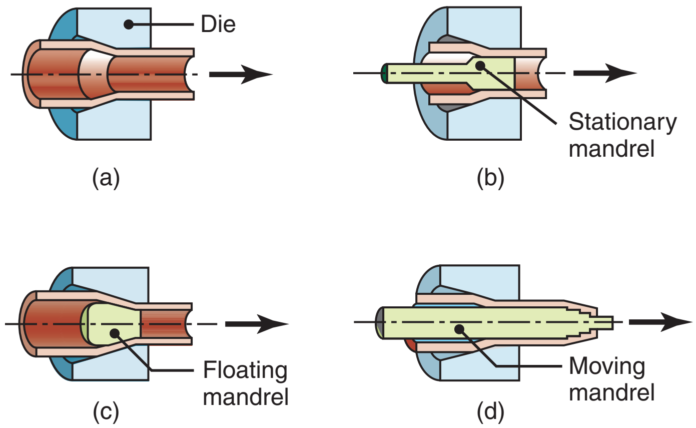


Figure 15.21: Examples of tube-drawing operations, with and without an internal mandrel. Note that a range of diameters and wall thicknesses can be produced from the same initial tube stock which has been made by other processes.

A set of dies is required for **profile drawing**; they involve various stages of deformation to produce the final profile. The dies may be made in one piece or, depending on the complexity of the cross-sectional profile, with several segments; they are held together in a retaining ring. Computer-aided design techniques are implemented to design dies, to ensure smooth material flow and to minimize any defect formation. A set of *idling* cylindrical or *shaped rolls* also may be used in drawing rods or bars of various shapes. Such an arrangement, called a **Turk's head**, is more versatile than common draw dies, because the rolls can be adjusted to different positions and angles for drawing specific profiles.

Die Materials. Die materials for drawing (see Table 5.8) typically are tool steels and carbides. For hot drawing, cast-steel dies can be used because of their high resistance to wear at elevated temperatures. Diamond dies are used for drawing fine wire, with diameters ranging from 2 μm to 1.5 mm. They may be either a *single-crystal* diamond or *polycrystalline*, with diamond particles embedded in a metal matrix, called *compacts*. Because of their very low tensile strength and toughness, carbide and diamond dies are typically used as **inserts** or **nibs**, which are supported in a steel casing (Fig. 15.22).

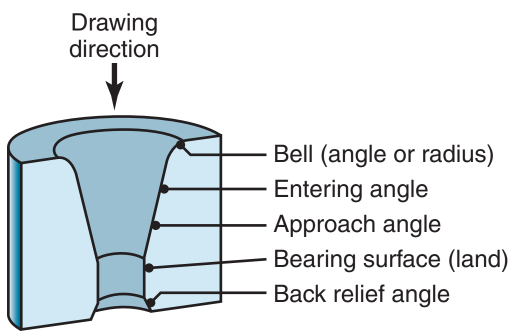


Figure 15.22: Terminology pertaining to a typical die used for drawing a round rod or wire.

Lubrication. Lubrication is essential in drawing operations in order to improve die life and surface finish, and to reduce drawing forces and temperature rise. Lubrication is critical particularly in tube drawing, because of the difficulty of maintaining a sufficiently thick lubricant film at the mandrel–tube interfaces. In drawing rods, a common method of lubrication is phosphate **conversion coatings** (see Section 33.7).

The basic methods of lubrication in wire drawing are (see also Section 33.7):

- **Wet drawing:** the dies and the rod are immersed completely in the lubricant.
- **Dry drawing:** the surface of the rod to be drawn is coated with a lubricant, by passing it through a box filled with the lubricant (*stuffing box*).
- **Metal coating:** the rod is coated with a soft metal, such as copper or tin, that acts as a solid lubricant.
- **Ultrasonic vibration:** vibrations of the dies and mandrels improve surface finish and die life, and reduce drawing forces, thus allowing higher reductions per pass without failure.

15.10 Drawing Defects and Residual Stresses

Typical defects in a drawn rod or wire are similar to those observed in extrusion, especially **center cracking** (see Fig. 15.16). Another major type of defect in drawing is **seams**, which are longitudinal folds in the drawn product. Seams may later open up during subsequent forming operations, such as upsetting, heading, thread rolling, or bending of the rod or wire, and may cause serious quality-control problems. Various other surface defects, such as scratches and die marks, may be due to improper selection of process parameters, poor lubrication, or poor die condition.

Because they undergo nonuniform deformation during drawing, cold-drawn products usually have *residual stresses*. For light reductions, such as only a few percent, the longitudinal-surface residual stresses are compressive while the bulk is in tension, and fatigue life is thus improved. Conversely, heavier reductions induce tensile surface stresses, while the bulk is in compression. Residual stresses can be significant in causing stress-corrosion cracking (Section 2.10.2) of the part over time. Moreover, they cause the component to *warp*, if a layer of material is subsequently removed (see Fig. 2.30), such as by slitting, machining, or grinding. Rods and tubes that are not sufficiently straight, or are supplied as coil, can be straightened by passing them through an arrangement of rolls placed at different axes, a process similar to roller leveling shown in Fig. 13.8b.

15.11 Drawing Equipment

Although it is available in several designs, the equipment for drawing is basically of two types: the draw bench and the bull block. A **draw bench** contains a single die, and its design is similar to that of a long, horizontal tension-testing machine (Fig. 15.23). The pulling force is supplied by a chain drive or hydraulic cylinder. Draw benches are used for a single-length drawing of straight rods and tubes, with diameters larger than 20 mm and lengths up to 30 m. Machine capacities reach 1.3 MN (136 metric tons) of pulling force, with a speed range of 6–60 m/min.

Very long rods and wire of smaller cross sections, usually less than 13 mm, are drawn by a rotating *drum* (**bull block** or **capstan**, Fig. 15.24). The tension in this setup provides the force required for drawing the wire, usually through multiple dies (*tandem drawing*).

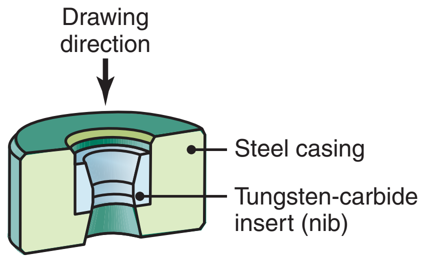


Figure 15.23: Tungsten-carbide die insert in a steel casing. Diamond dies used in drawing thin wire are encased in a similar manner.

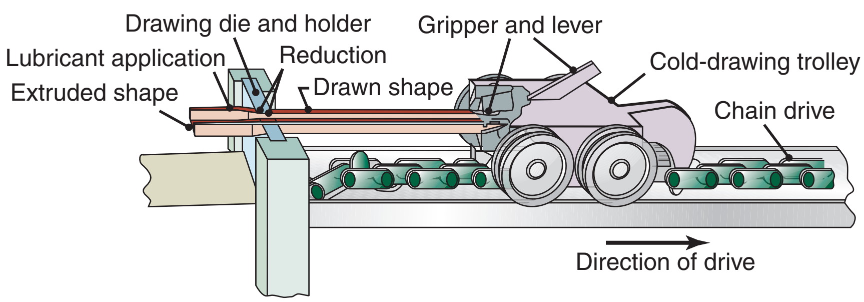


Figure 15.24: Cold drawing of an extruded channel on a draw bench to reduce its cross section. Individual lengths of straight rods or of cross sections are drawn by this method.

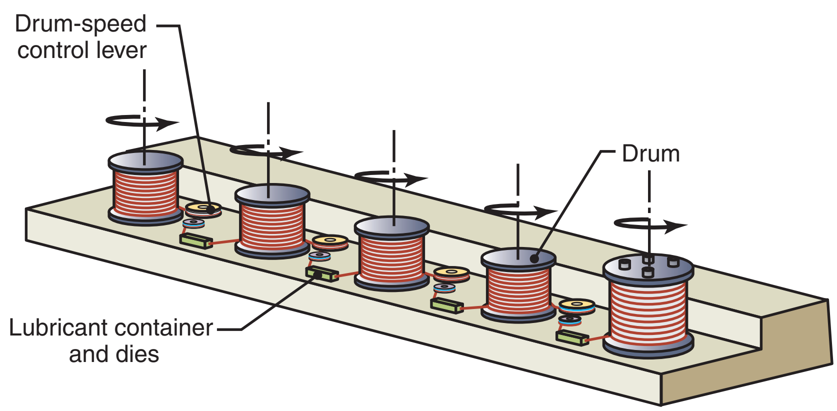


Figure 15.25: An illustration of multistage wire drawing typically used to produce copper wire for electrical wiring. Shown is a five bull block configuration; wire drawing machines can incorporate 15 or more drums, depending on the material and wire size. *Source:* After H. Auerswald.

Summary

- Extrusion is the process of forcing a billet through a die to reduce its cross section or to produce various solid or hollow cross sections. The process is generally carried out at elevated temperatures in order to reduce the extrusion force and improve the ductility of the material.
- Important factors in extrusion are die design, extrusion ratio, billet temperature, lubrication, and extrusion speed. Although the term cold extrusion applies to extrusion at room temperature, it is also the name for a combination of extrusion and forging operations. Cold extrusion is capable of economically producing discrete parts in various shapes and with good mechanical properties and dimensional tolerances.
- Rod, wire, and tube drawing operations basically involve pulling the material through a die or a set of dies in tandem. The cross sections of most drawn products are round, but other shapes also can be drawn. Drawing tubular products, to reduce either their diameter or their thickness, usually requires the use of internal mandrels.
- Die design, reduction in cross-sectional area per pass, and selection of die materials and lubricants are all important parameters in making drawn products of high quality and with good surface finish. External and internal defects can develop both in extrusion and in drawing. The significant factors are the die angle, reduction per pass, and quality of the workpiece material.

Key Terms

Bamboo defect
Bridge die
Bull block
Bundle drawing
Canning
Capstan
Center cracking
Chevron cracking
Cold extrusion
Conversion coating
Dead-metal zone
Draw bench
Drawing
Extrusion
Extrusion constant
Extrusion defects
Extrusion ratio
Fir-tree cracking
Hydrostatic extrusion
Impact extrusion
Ironing
Jacketing
Patenting
Pipe defect
Porthole die
Rod
Seam
Séjournet process
Shear die
Sizing pass
Speed cracking
Spider die
Turk's head
Wire

Bibliography

Altan, T., Ngaile, G., and Shen, G. (eds.), **Cold and Hot Forging: Fundamentals and Applications**, ASM International, 2004.

ASM Handbook, Vol. 14A: **Metalworking: Bulk Forming**, ASM International, 2005.
Bauser, M., Sauer, G., and Siegert, K. (eds.), **Extrusion**, 2nd ed., ASM International, 2006.
Hosford, W.F., and Caddell, R.M., **Metal Forming: Mechanics and Metallurgy**, 4th ed., Prentice Hall, 2014.
Michaeli, W., **Extrusion Dies**, 2nd ed., Hanser, 1992.
Saha, P., **Aluminum Extrusion Technology**, ASM International, 2000.
Sheppard, T., **Extrusion of Aluminum Alloys**, Chapman & Hall, 1997.
Wright, R.N., **Wire Technology: Process Engineering and Metallurgy**, 2nd ed., Butterworth-Heinemann, 2016.

Review Questions

15.1. How does extrusion differ from rolling and forging?

15.2. Explain the difference between extrusion and drawing.

15.3. What is a spider die? What is it used for?

15.4. Why are wires sometimes drawn in bundles?

15.5. What is a dead-metal zone?

15.6. Define the terms (a) cladding, (b) dummy block, (c) shear dies, (d) skull, and (e) canning.

15.7. Why is glass a good lubricant in hot extrusion?

15.8. What types of defects may occur in (a) extrusion and (b) drawing?

15.9. Describe the difference between direct and reverse extrusion.

15.10. What is land? What is its function in a die?

15.11. How are tubes extruded? Can they also be drawn? Explain.

15.12. It is possible to extrude straight gears; can helical gears also be extruded? Explain.

15.13. What is the difference between piping and bambooing?

15.14. What is impact extrusion?

15.15. What is the pipe defect in extrusion?

Qualitative Problems

15.16. List the similarities and differences between direct extrusion and drawing.

15.17. Explain why extrusion is a batch, or semicontinuous, process. Do you think it can be made into a continuous process? Explain.

15.18. The extrusion ratio, die geometry, extrusion speed, and billet temperature all affect the extrusion pressure. Explain why.

15.19. Explain why cold extrusion is an important manufacturing process.

15.20. What is the function of a stripper plate in impact extrusion?

15.21. Explain the different ways by which changing the die angle affects the extrusion process.

15.22. Glass is a good lubricant in hot extrusion. Would you use glass for impression-die forging also? Explain.

15.23. How would you go about avoiding center-cracking defects in extrusion? Explain why your methods would be effective.

15.24. Table 15.1 gives temperature ranges for extruding various metals. Describe the possible consequences of extruding at a temperature (a) below and (b) above these ranges.

15.25. Will the force in direct extrusion vary as the billet becomes shorter? If so, why?

15.26. Comment on the significance of grain-flow patterns, such as those shown in Fig. 15.6.

15.27. In which applications could you use the type of impact-extruded parts shown in Fig. 15.15?

15.28. What is the purpose of the land in a drawing die? Is there a limit to the size of the land that should be used? Explain your answer.

15.29. Can spur gears be made by (a) drawing and (b) extrusion? Can helical gears? Explain.

15.30. How would you prepare the end of a wire in order to be able to feed it through a die so that a drawing operation can commence?

15.31. What is the purpose of a dummy block in extrusion? Explain.

15.32. Describe your observations concerning Fig. 15.9.

15.33. Occasionally, steel wire drawing will take place within a sheath of a soft metal, such as copper or lead. What is the purpose of this sheath?

15.34. Explain the advantages of bundle drawing.

15.35. Under what circumstances would backwards extrusion be preferable to direct extrusion?

15.36. Why is lubrication detrimental in extrusion with a porthole die?

15.37. In hydrostatic extrusion, complex seals are used between the ram and the container, but not between the extrusion and the die. Explain why.

15.38. Describe the purpose of a container liner in direct extrusion, as shown in Fig. 15.1. What is the liner's function in reverse extrusion?

Quantitative Problems

15.39. Estimate the force required in extruding 70–30 brass at 700°C if the billet diameter is 250 mm and the extrusion ratio is 25.

15.40. Assuming an ideal drawing process, what is the smallest final diameter to which a 75-mm diameter rod can be drawn?

15.41. If you include friction in Problem 15.40, would the final diameter be different? Explain.

15.42. Calculate the extrusion force for a round billet 300 mm in diameter, made of stainless steel, and extruded at 1000°C to a diameter of 90 mm.

15.43. A planned extrusion operation involves steel at 1000°C with an initial diameter of 100 mm and a final diameter of 25 mm. Two presses, one with capacity of 20 MN and the other with a capacity of 10 MN, are available for the operation. Is the smaller press sufficient for this operation? If not, what recommendations would you make to allow the use of the smaller press?

15.44. A round wire made of a perfectly plastic material with a flow stress of 200 MPa is being drawn from a diameter of 2.5 to 1.5 mm in a draw die of 15°. Let the coefficient of friction be 0.15. Using both Eqs. (15.4) and (15.5), estimate the drawing force required. Comment on the differences in your answer.

15.45. Assume that you are an instructor covering the topics described in this chapter and you are giving a quiz on the numerical aspects to test the understanding of the students. Prepare two quantitative problems and supply the answers.

Synthesis, Design, and Projects

15.46. Assume that the summary to this chapter is missing. Write a one-page summary of the highlights of the wire-drawing process.

15.47. Review the technical literature, and make a detailed list of the manufacturing steps involved in the manufacture of common metallic hypodermic needles.

15.48. Figure 15.2 shows examples of discrete parts that can be made by cutting extrusions into individual pieces. Name several other products that can be made in a similar fashion.

15.49. The parts in Fig. 15.2 are economically produced by extrusion, but difficult to produce otherwise. List the processes that could be used to produce these parts, and explain why they are not as attractive as extrusion.

15.50. Survey the technical literature, and explain how external vibrations can be applied to a wire-drawing operation to reduce friction. Comment also on the possible directions of vibration, such as longitudinal or torsional.

15.51. How would you go about making a stepped extrusion that has increasingly larger cross sections along its length? Is it possible? Would your process be economical and suitable for high production runs? Explain.

15.52. List the processes that are suitable for producing an aluminum tube. For each process in your list, make a sketch of the grain structure you would expect to see in the finished product.

15.53. Assume that you are the technical director of trade associations of (a) extruders and (b) rod- and wire-drawing operations. Prepare a technical leaflet for potential customers, stating all of the advantages of these processes.

Chapter 16

Sheet-metal Forming Processes and Equipment

Example:

Case Studies:

- This chapter describes the characteristics of sheet metals and the forming processes employed to produce a wide variety of products.
- The chapter opens with a description of the shearing operation, to cut sheet metal into blanks with desired shapes or to remove portions of the material, such as for holes or slots.
- A review of sheet-metal formability follows, with special emphasis on the specific metal properties that affect formability.
- The chapter then presents various bending operations for sheets, plates, and tubes and such operations as stretch forming, rubber forming, spinning, peen forming, and superplastic forming.
- Deep drawing is then described, along with drawability, as it relates to the production of containers with thin walls.
- The chapter ends with an introduction to sheet-metal parts design, equipment characteristics, and the economic considerations for all these operations.

Typical parts made by sheet-metal forming: Truck bodies, aircraft fuselages, trailers, office furniture, appliances, fuel tanks, and cookware.

Alternative process: Die casting, thermoforming, pultrusion, injection molding, blow molding.

16.1 Introduction

Products made of **sheet metals** are all around us. They include a very wide range of consumer and industrial products: beverage cans, cookware, file cabinets, metal desks, appliances, car and truck bodies, trailers, and aircraft fuselages (Fig. 16.1). Sheet forming dates back to 5000 B.C., when household utensils and jewelry were made by hammering and stamping gold, silver, and copper. Compared to those made by casting and forging, sheet-metal parts offer the advantages of versatile shapes, light weight, and high stiffness-to-weight ratios.

As described throughout this chapter, there are numerous processes employed for making sheet-metal parts. The terms **pressworking** or **press forming** are commonly used to describe these operations, because

(a) (b)

Figure 16.1: Examples of sheet-metal parts. (a) Stamped parts. (b) Parts produced by spinning. *Source:* Courtesy of Williamsburg Metal Spinning & Stamping Corp.

Table 16.1: General Characteristics of Sheet-metal Forming Processes (in alphabetic order).

Forming process	Characteristics
Drawing	Shallow or deep parts with relatively simple shapes, high production rates, high tooling, and equipment costs
Explosive	Large sheets with relatively simple shapes, low tooling costs but high labor cost, low-quantity production, long cycle times
Hot Stamping	Simple shapes; used for advanced high-strength steels and materials with limited formability; in-die quenching can lead to superior mechanical properties
Incremental	Simple to moderately complex shapes with good surface finish; low production rates, but no dedicated tooling required; limited materials
Magnetic-pulse	Shallow forming, bulging, and embossing operations on relatively low-strength sheets, requires special tooling
Peen	Shallow contours on large sheets, flexibility of operation, generally high equipment costs, process also used for straightening formed parts
Roll	Long parts with constant simple or complex cross sections, good surface finish, high production rates, high tooling costs
Rubber	Drawing and embossing of simple or relatively complex shapes, sheet surface protected by rubber membranes, flexibility of operation, low tooling costs
Spinning	Small or large axisymmetric parts; good surface finish; low tooling costs, but labor costs can be high unless operations are automated
Stamping	Includes a wide variety of operations, such as punching, blanking, embossing, bending, flanging, and coining; simple or complex shapes formed at high production rates; tooling and equipment costs can be high, but labor cost is low
Stretch	Large parts with shallow contours, low-quantity production, high labor costs, tooling and equipment costs increase with part size
Superplastic	Complex shapes, fine detail and close dimensional tolerances, long forming times (hence production rates are low), parts not suitable for high-temperature use

they typically are performed on *presses* (Sections 14.8 and 16.15), typically using a set of dies. A sheet-metal part produced in presses is called a **stamping** (after the word *stamp*, first used around 1200 A.D., and meaning to force downward or to pound). Low-carbon steel is the most commonly used sheet metal, because of its low cost and generally good strength and formability characteristics. TRIP and TWIP steels (Section 5.5.5) have become more common for automotive applications, because they are strong and provide good crash protection in a lightweight design.

Aluminum is the most common material for such applications as beverage cans, packaging, kitchen utensils, and where corrosion resistance is an important requirement. Common metallic materials for aircraft and aerospace applications are aluminum, titanium, and, more recently, composite materials (Chapters 9 and 19).

Most processes involve sheet metals at room temperature, although hot stamping is also done in order to increase formability and decrease springbuck and forming loads on machinery. Typical sheet metals in hot-stamping operations are titanium alloys and various high-strength steels.

This chapter first describes the methods by which blanks are cut from large sheets, then further processed into desired shapes. The chapter also includes a review of the characteristic features of sheet metals, the techniques employed to determine their formability, and the construction of forming-limit diagrams. All major processes of sheet forming and the equipment involved are also described, as outlined in Table 16.1.

16.2 Shearing

All sheet-metal forming operations begin with a **blank** of suitable dimensions, cut from a large sheet, usually from a *coil* (see Fig. 13.11) by **shearing**. Shearing subjects the sheet to shear stresses, generally using a punch and a die (Fig. 16.2a). The typical features of the sheared edges of the sheet metal and of the slug are

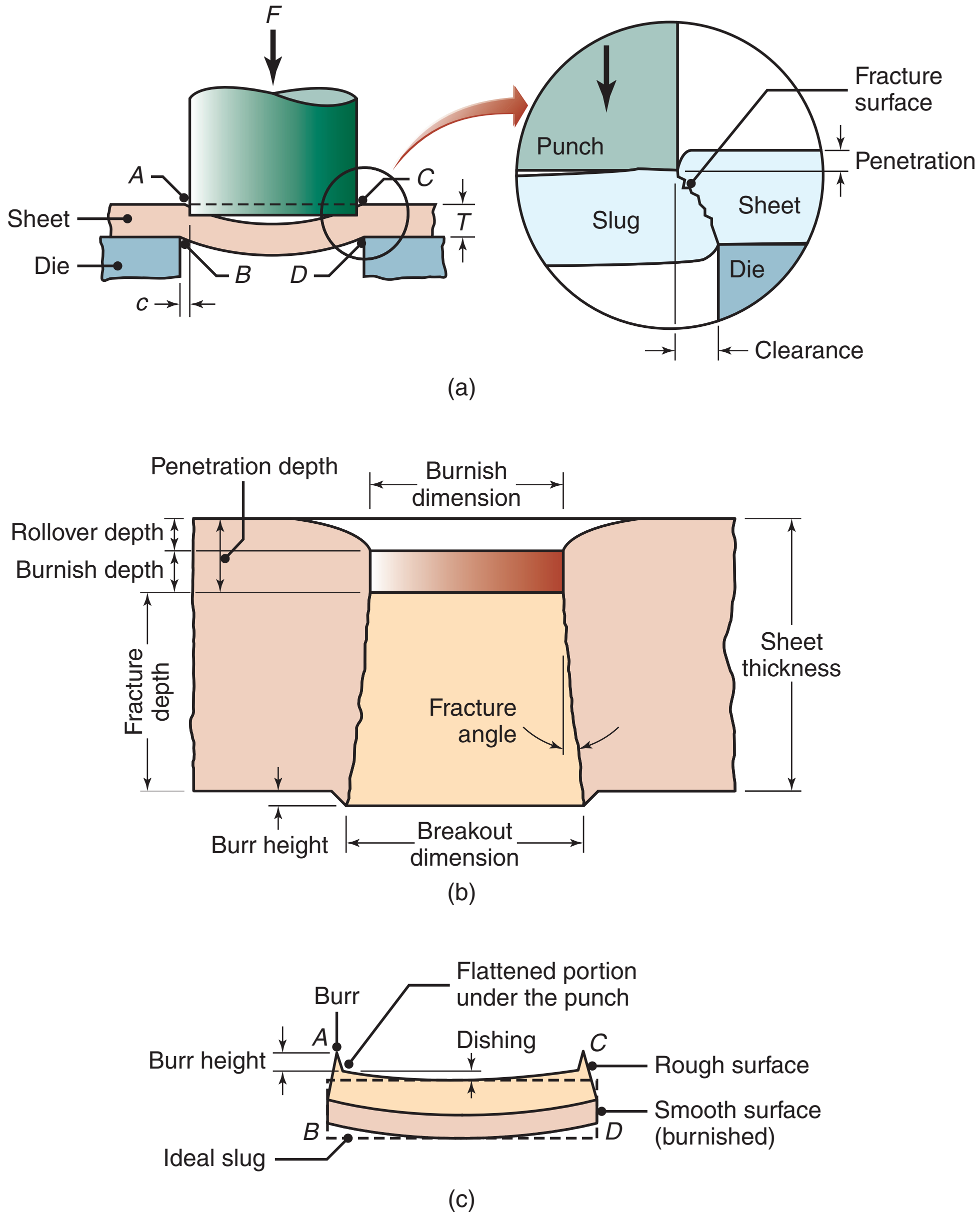


Figure 16.2: (a) Schematic illustration of shearing with a punch and die, indicating some of the process variables. Characteristic features of (b) a punched hole and (c) the slug. (Note that the scales of (b) and (c) are different.)

shown in Fig. 16.2b and c, respectively. Note that the edges are not smooth nor are they perpendicular to the plane of the sheet in this illustration.

Shearing generally starts with the formation of cracks on both the top and bottom edges of the workpiece, at points A and B, and C and D, in Fig. 16.2a. These cracks eventually meet each other, leading to complete separation. The rough *fracture surfaces* are due to the cracks; the smooth and shiny *burnished surfaces* on the hole and the slug are from contact and rubbing of the sheared edge against the walls of the punch and die, respectively.

The major processing parameters in shearing are:

- Shape of the punch and die
- Clearance, c, between the punch and the die
- Punching speed
- Lubrication.

The **clearance** is a major factor in determining the shape and the quality of the sheared edge. As clearance increases, the deformation zone (Fig. 16.3a) becomes larger and the surface of the sheared edge becomes rougher. With excessive clearances, the sheet tends to be pulled into the die cavity, and the perimeter or edges of the sheared zone become rougher. Unless such edges are acceptable, secondary operations may be necessary to make them smoother, increasing the production cost (see also *fine blanking*, Section 16.2.1).

Edge quality can be improved with increasing punch speed, to as high as 10 to 12 m/s. As shown in Fig. 16.3b, sheared edges can undergo severe cold working due to the high shear strains involved. Work hardening of the edges then will reduce the ductility of the edges, thus adversely affecting the formability of the sheet during such subsequent operations as bending and stretching.

The ratio of the burnished area to the rough areas along the sheared edge increases with increasing ductility of the sheet metal, and decreases with increasing sheet thickness and clearance. The extent of the deformation zone, shown in Fig. 16.3, depends also on the punch speed. With increasing speed, the heat generated by plastic deformation becomes confined to a smaller and smaller zone. Consequently, the sheared zone becomes narrower, and the sheared surface is smoother and exhibits less burr formation.

A **burr** is a thin edge or ridge (Fig. 16.2b and c). Its height increases with increasing clearance and ductility of the sheet metal; dull tool edges contribute greatly to large burr formation. The height, shape, and size of the burr can significantly affect subsequent forming operations. Several **deburring** processes are described in Section 26.8.

Punch Force. The force required to punch out a blank is basically the product of the shear strength of the sheet metal and the total area being sheared. The *maximum punch force*, F, can be estimated from the equation

$$F = 0.7TLS_{\mathrm{ut}}, \tag{16.1}$$

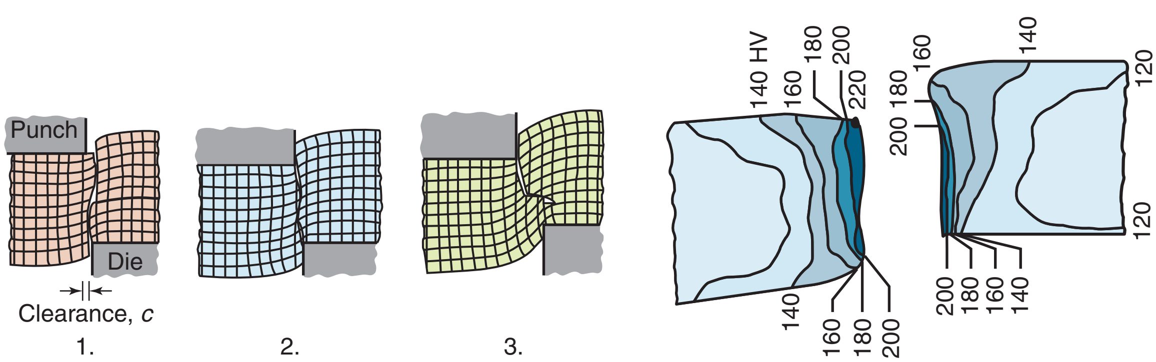


Figure 16.3: (a) Effect of the clearance, c, between punch and die on the deformation zone in shearing. As the clearance increases, the material tends to be pulled into the die rather than be sheared. In practice, clearances usually range between 2% and 10% of the thickness of the sheet. (b) Microhardness (HV) contours for a 6.4-mm thick AISI 1020 hot-rolled steel in the sheared region.

where T is the sheet thickness, L is the total length sheared, such as the perimeter of a hole, and $S_{\rm ut}$ is the ultimate tensile strength of the material. As the clearance increases, the punch force decreases, and the wear on dies and punches also is reduced. The effects of punch shape and die shape on punch forces are described in Section 16.2.3.

Friction between the punch and the sheet significantly increases the punch force. Furthermore, a force is required to strip the punch from the sheet during its return stroke. This force, which is in opposite direction to the punch force, is difficult to estimate because of the several factors involved in the punching operation.

Example 16.1 Calculation of Punch Force

Given: A 25-mm diameter hole is to be punched through a 3.2-mm thick annealed titanium-alloy Ti-6Al-4V sheet at room temperature.

Find: Estimate the force required.

Solution: The force is estimated from Eq. (16.1), where $S_{\rm ut}$ for this alloy is found from Table 6.10 to be 1000 MPa. Thus,

$$F = 0.7\,(0.0032)\,(\pi)(0.025)(1000 \times 10^6) = 0.17\ \text{MN}.$$

16.2.1 Shearing Operations

The most common shearing operations are **punching**, where the sheared slug is scrap (Fig. 16.4a) or it may be used for some other purpose, and **blanking**, where the slug is the part to be used and the rest is scrap. As in most other processes, shearing operations are now carried out on computer-numerical-controlled machines with *quick-change toolholders* (Section 16.15).

Die Cutting. This is a shearing operation that consists of the following basic processes, as shown in Fig. 16.4b:

- Perforating: punching a number of holes in a sheet
- Parting: shearing the sheet into two or more pieces
- Notching: removing pieces from edges
- Lancing: producing a tab without removing any material.

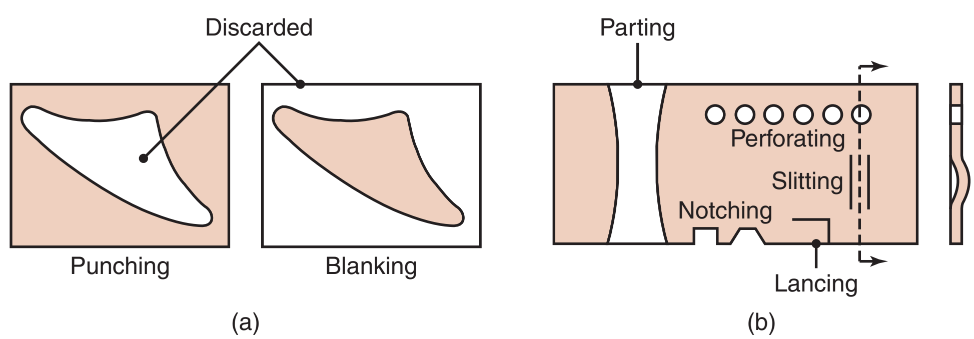
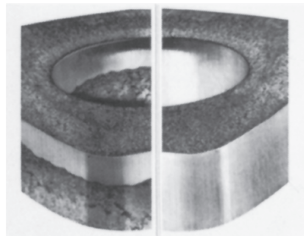
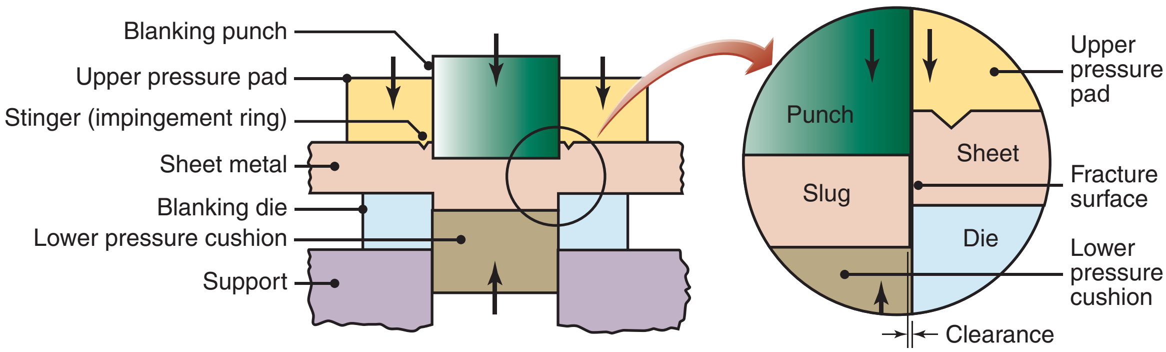


Figure 16.4: (a) Punching (piercing) and blanking. (b) Examples of various die-cutting operations on sheet metal. Lancing involves slitting the sheet to form a tab.



(a)



(b)

Figure 16.5: (a) Comparison of sheared edges produced by conventional (left) and by fine-blanking (right) techniques. (b) Schematic illustration of one setup for fine blanking. *Source:* Courtesy of Feintool U.S. Operations Inc.

Parts made by these processes have numerous important uses, particularly in assembly with other sheet-metal components. Perforated sheet metals, for example, with hole diameters ranging from 1 mm to 75 mm have uses as filters, screens, ventilation, guards for machinery, noise abatement, and in weight reduction of fabricated parts and structures. They are punched in crank presses (Fig. 14.19a), at rates as high as 300,000 holes per minute, using special dies and equipment.

Fine Blanking. Square edges with very smooth sheared surfaces can be produced by *fine blanking* (Fig. 16.5a). One basic die design is shown in Fig. 16.5b. A V-shaped *stinger* or impingement mechanically locks the sheet in place, thus preventing the type of distortions shown in Figs. 16.2b and 16.3. Fine-blanking involves clearances on the order of 1% of the sheet thickness, and may range from 0.5 to 13 mm in most applications. Dimensional tolerances are typically on the order of ± 0.05 mm.

Slitting. Shearing operations can be carried out by means of a pair of circular blades (Fig. 16.6), similar to those in a can opener. In *slitting*, the blades follow a straight line, or a circular or curved path. A slit edge normally has a burr, which may be *folded over* the sheet surface by rolling it (flattening) between two cylindrical rolls. If not performed properly, slitting operations can cause various distortions of sheared edges.

Steel Rules. Soft metals, paper, leather, and rubber can be blanked with a *steel-rule die*. Such a die consists of a thin strip of hardened steel bent into the shape to be cut (similar to that of a cookie cutter), and pressed against the sheet, which rests on the flat surface.

Nibbling. In this process, a machine called a *nibbler*, moves a small straight punch up and down rapidly into a die. A sheet is fed through the gap and several overlapping holes are made. With manual or automatic

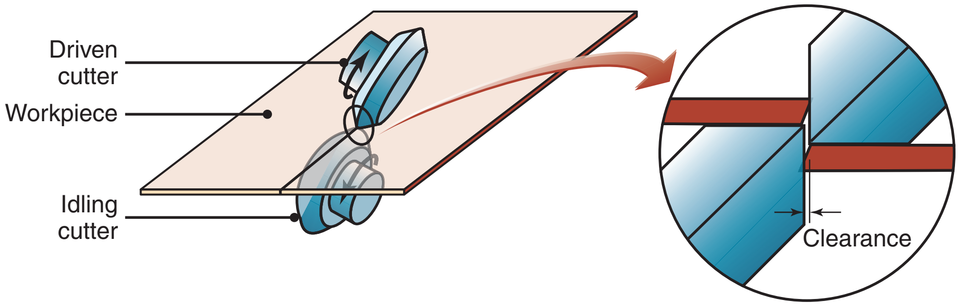


Figure 16.6: Slitting with rotary knives; this process is similar to opening cans.

control, sheets can be cut along any desired path. In addition to its flexibility, nibbling makes intricate slots and notches, such as those shown in Fig. 16.4b, using standard punches. Because no special dies are required, the process is economical for small production runs.

Scrap in Shearing. The amount of scrap (*trim loss*) produced in shearing operations can be as high as 30%, on large stampings (see Table 40.4). Scrap can be a significant factor in manufacturing costs, but it can be reduced substantially by efficient arrangement of the shapes on the sheet to be cut (**nesting**, see Fig. 16.60). Also, computer-aided design techniques are available to minimize scrap.

16.2.2 Tailor-welded Blanks

In the sheet forming processes to be described throughout this chapter, the blank is typically a one-piece sheet of constant thickness, and cut (blanked) from a large sheet. An important variation from this practice involves *laser-beam butt welding* (Section 30.7) of two or more pieces of sheet metal with different shapes and thicknesses. The strips are welded to obtain a locally thicker sheet or add a different material (see Case Study 16.1).

Because of the thicknesses involved are very small, proper alignment of the sheets prior to welding is important. The welded assembly is subsequently formed into a final shape by various processes. This technique has become increasingly important, particularly to the automotive industry. Because each piece now can have a different thickness, composition, coating, or other characteristics, tailor-welded blanks have such advantages as:

- Reduction in scrap
- Elimination of the need for subsequent spot welding operations (see Fig. I.9)
- Better control of dimensions
- Increased productivity.

Case Study 16.1 Tailor-welded Sheet Metal for Automotive Applications

An example using tailor-welded sheet metals is shown in Fig. 16.7. Note that five different pieces are first blanked, which includes cutting by laser beams (Section 16.2.4). Four of these pieces are 1-mm thick, and one is 0.8 mm thick. The pieces are laser butt welded (Section 30.7), then stamped into the final shape. In this manner, the blanks can be tailored to a particular application, not only as to shape and thickness, but also by using different-quality metals, with or without coatings.

Laser-welding techniques are now highly developed and the joints are very strong and reliable. The combination of welding and forming sheet-metal pieces makes possible significant flexibility in product design, formability, structural stiffness, and crash behavior of a vehicle. It also makes possible the use of different materials in one product, weight savings, and cost reductions in materials, scrap, assembly, equipment, and labor.

The various components shown in Fig. 16.8 utilize the advantages outlined above. For example, note in Fig. 16.8b that the strength and stiffness required for the support of the shock absorber are achieved by welding a round piece onto the surface of the large sheet. The sheet thickness in such components varies, depending on its location and on its contribution to such characteristics as stiffness and strength, resulting in significant weight savings without loss of structural strength and stiffness.

Recent advances include *friction stir welding* (Section 31.4) to produce tailor-welded blanks, and the production of *tailor-welded* or tailor-rolled coils, where the material and/or thickness can be made differently at a given location in the sheet (see Fig. 13.7). Such blanks have also been used in hot stamping of automotive space frame pillars (Section 16.11). In this application, a steel grade is used to minimize deflections and to protect occupants, and a more ductile steel that absorbs energy (see *toughness*, Sections 2.2.4 and 2.10) is used where the pillar is attached to the car frame.

16.2.3 Characteristics and Types of Shearing Dies

Clearance. Because the formability of a sheared part can be influenced by the quality of its sheared edges, clearance control is important. An appropriate clearance depends on

- Type of material and its temper
- Thickness and size of the blank
- Proximity to the edges of other sheared edges or to the edges of the original blank.

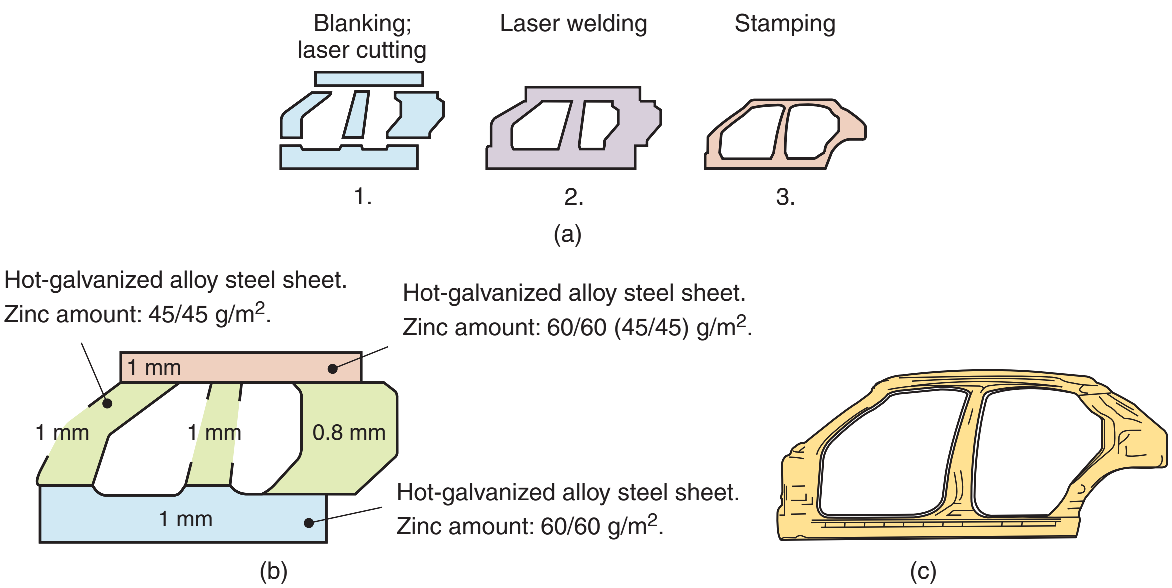


Figure 16.7: Production of an outer side panel of a car body by laser butt welding and stamping. The thickness of each section is as indicated. *Source:* After M. Geiger and T. Nakagawa.

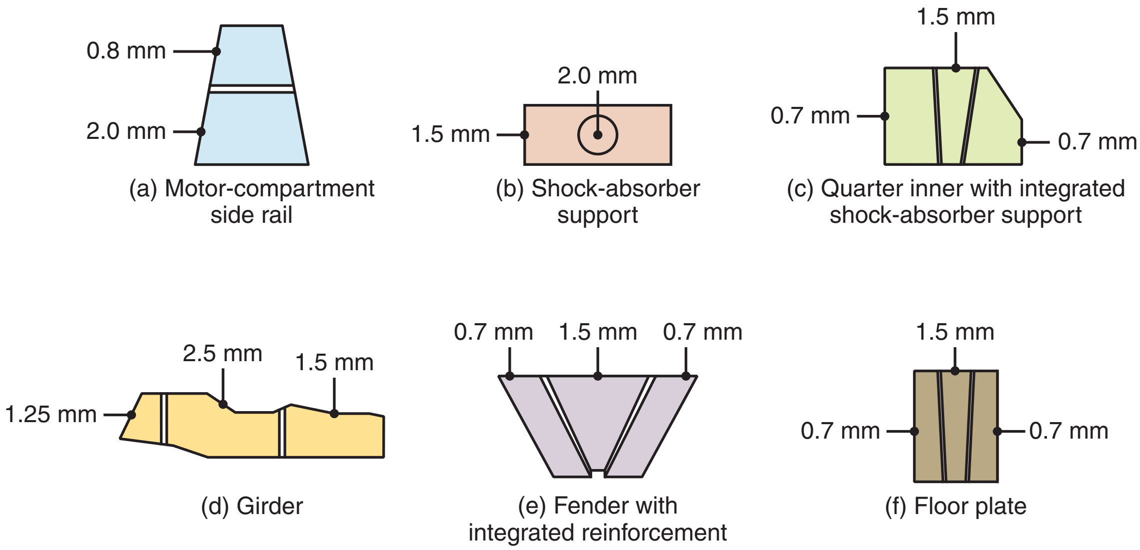


Figure 16.8: Examples of laser butt-welded and stamped automotive-body components. *Source:* After M. Geiger and T. Nakagawa.

Clearances generally range between 2% and 8% of the sheet thickness, although they may be as small as 1% (as in *fine blanking*, Section 16.2.1) or as large as 30%. The smaller the clearance, the better is the edge quality. If the sheared edge is rough and not acceptable, it can be subjected to **shaving** (Fig. 16.9a), a process whereby the extra material from the edge is trimmed by cutting, as depicted in Fig. 21.3.

As a general guideline, (a) clearances for softer materials are smaller than those for harder grades; (b) the clearance has to be larger for thicker sheets; and (c) as the ratio of hole diameter to sheet thickness decreases, clearances must be larger. However, in using larger clearances, attention must be paid to the rigidity and the alignment of the presses, the dies, and their setups.

Punch and Die Shapes. Note in Fig. 16.2a that the surfaces of the punch and of the die are both flat. Because the entire thickness will be sheared at the same time, the punch force increases very rapidly. This force can be controlled by *beveling* the punch and die surfaces (Fig. 16.10). Note also shape of the tip of some common paper punches. Beveling is particularly suitable for shearing thick sheets, because it reduces the force at the beginning of the stroke.

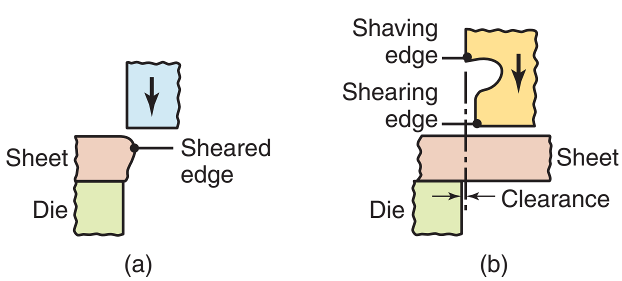


Figure 16.9: Schematic illustrations of the shaving process. (a) Shaving a sheared edge. (b) Shearing and shaving combined in one stroke.

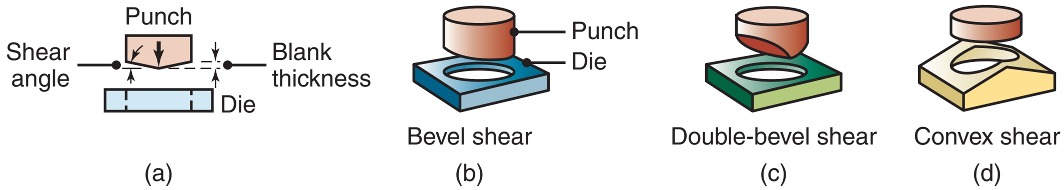


Figure 16.10: Examples of the use of shear angles on punches and dies.

Note in Fig. 16.10c that the punch tip is symmetrical, and in Fig. 16.10d that the die is also symmetrical; thus, there are no lateral forces that could cause distortion. By contrast, the punch in Fig. 16.10b has a single taper, and thus it is subjected to a lateral force. Consequently, the punch and press setups must both have sufficient lateral stiffness, so that they neither produce a hole that is located improperly nor allow the punch to hit the edge of the lower die and cause damage (as it might at point B or D in Fig. 16.2a).

Compound Dies. Several operations can be performed in one stroke and on the same sheet and at one station, using *compound die* (Fig. 16.11). Such combined operations usually are limited to relatively simple shapes, because (a) the process is somewhat slow and (b) the dies become much more expensive, especially for complex dies.

Progressive Dies. Parts requiring multiple forming operations can be made, and at high production rates, using *progressive dies*. The sheet metal is fed through as a coil strip, and a different operation (such as punching, blanking, and notching) is performed at the same station of the machine with each stroke and using a series of punches (Fig. 16.11c). An example of a part made in progressive dies is shown in Fig. 16.11d. The part made is the small round metal tip that supports the plastic nozzle in spray cans.

Transfer Dies. In a *transfer die*, the sheet metal undergoes different operations and at different stations of the machine. They are typically arranged along a straight line or as circular path. After each step in a station, the part is transferred (hence the name) to the next station for further sequential operations.

Tool and Die Materials. Tool and die materials for shearing generally are tool steels; carbides are also used for high production rates because of their higher mechanical properties (see Table 5.7). Lubrication is important for reducing tool and die wear and for maintaining edge quality.

16.2.4 Miscellaneous Methods of Cutting Sheet Metal

There are several other methods of cutting metal sheets and plates:

- **Laser-beam cutting** is an important process (Section 27.6), and typically used with computer-controlled equipment to cut (a) any shape consistently, (b) various thicknesses, and (c) without the use of any punches or dies. The process can also be combined with punching and shearing operations. Some parts with certain features may be produced best by one process, while others, with various features, may be produced best by the other process. Combination machines, incorporating both capabilities, have been designed and built, for this reason.
- **Water-jet cutting** is effective on metallic as well as nonmetallic materials (Section 27.8)
- Cutting with a **band saw** (Section 24.5)
- **Friction sawing**, which involves a disk or a blade that rubs against the sheet or plate at high surface speeds, thus raising the temperature and separating the sheet (Section 24.5)
- **Flame cutting**, a common method, particularly for thick plates, and used widely in shipbuilding and on heavy structural components (Section 30.8).

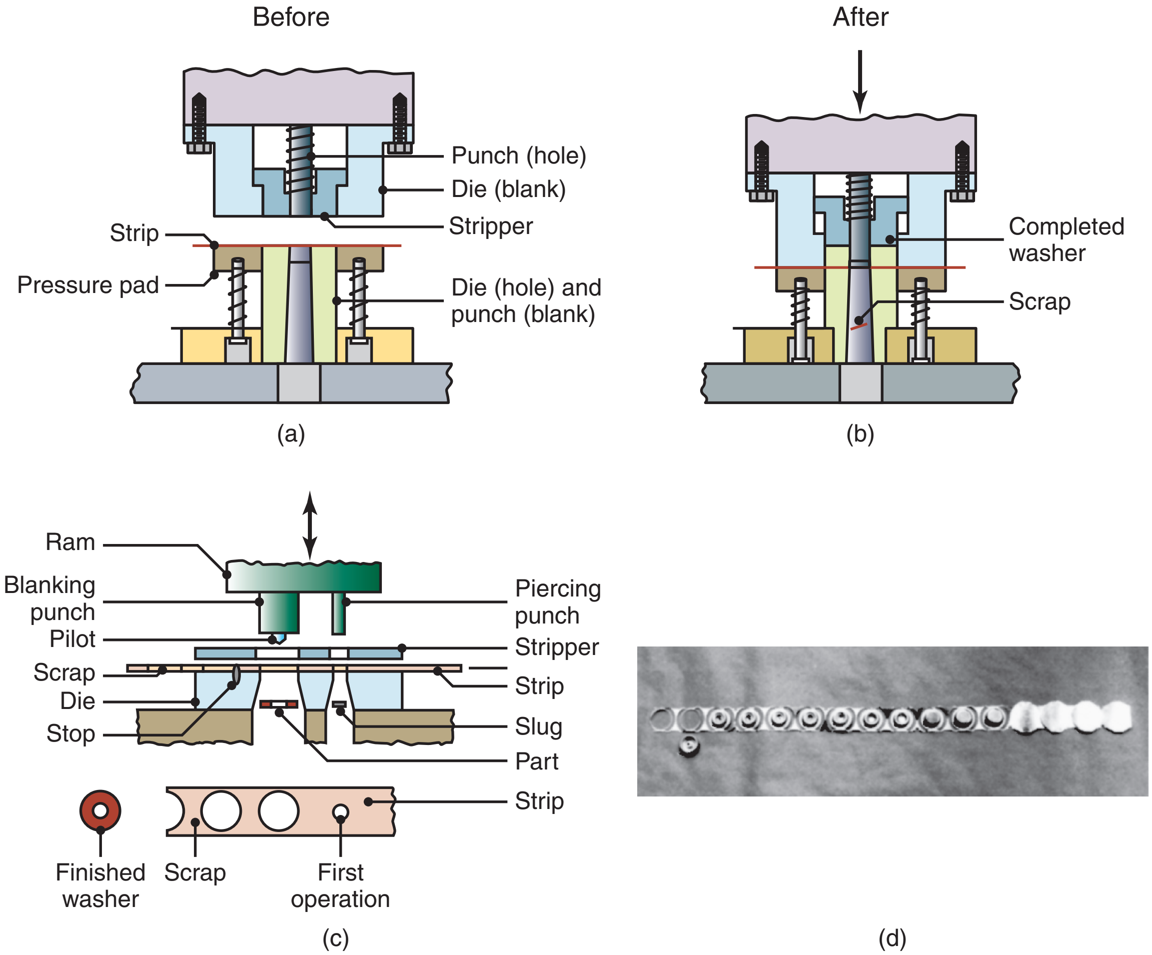


Figure 16.11: Schematic illustrations (a) before and (b) after blanking a common washer in a compound die. Note the separate movements of the die (for blanking) and the punch (for punching the hole in the washer). (c) Schematic illustration of making a washer in a progressive die. (d) Forming of the top piece of an aerosol spray can in a progressive die. Note that the part is attached to the strip until the last operation is completed.

16.3 Sheet-metal Characteristics and Formability

After a blank is cut from a larger sheet or coil, it can be formed into a wide variety of shapes by using several processes, described in the rest of this chapter. This section presents a brief review of the characteristics of sheet metals that have significant effects on forming operations, as outlined in Table 16.2.

Elongation. Sheet-metal forming processes rarely involve simple uniaxial stretching, as in a tension test; however, observations during tensile testing are useful and necessary for understanding the behavior of metals. Recall from Section 2.2 that a specimen subjected to tension first undergoes **uniform elongation**, and that when the load exceeds the ultimate tensile strength, the specimen begins to neck and the elongation is no longer uniform.

Because in sheet forming the material usually is being stretched, high uniform elongation is essential for good formability. The true strain at which necking begins is numerically equal to the *strain-hardening*

Table 16.2: Important Metal Characteristics for Sheet-metal Forming Operations.

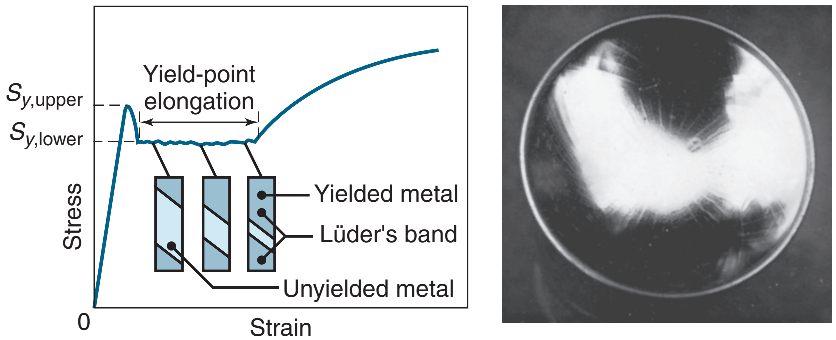
Characteristic	Importance
Elongation	Determines the capability of the sheet metal to be stretched without necking and failure; high strain-hardening exponent (n) and strain-rate sensitivity exponent (m) are desirable
Yield-point elongation	Typically observed with mild-steel sheets, also called Lüder's bands or stretcher strains; results in shallow depressions on the sheet surface; can be eliminated by *temper rolling*, but sheet must be formed within a specific time after rolling
Anisotropy (planar)	Exhibits different behavior in different planar directions; present in cold-rolled sheets because of preferred orientation or mechanical fibering, causes earing in deep drawing, can be reduced or eliminated by annealing, but at lowered strength
Anisotropy (normal)	Determines thinning behavior of sheet metals during stretching, important in deep drawing
Grain size	Determines surface roughness on stretched sheet metal; the coarser the grain, the rougher is the appearance (such as orange peel); also affects material strength and ductility
Residual stresses	Typically caused by nonuniform deformation during forming, results in part distortion when sectioned, can lead to stress-corrosion cracking; reduced or eliminated by stress relieving
Springback	Due to elastic recovery of the plastically deformed sheet after unloading; causes distortion of part and loss of dimensional accuracy; can be controlled by such techniques as *overbending* and *bottoming of the punch*
Wrinkling	Caused by compressive stresses in the plane of the sheet; can be objectionable; depending on its extent, can be useful in imparting stiffness to parts by increasing their section modulus; can be controlled by proper tool and die design
Quality of sheared edges	Depends on process used; edges can be rough, not square, and develop cracks, residual stresses, and a work-hardened layer, which are all detrimental to the formability of the sheet; edge quality can be improved by fine blanking, reducing the clearance, shaving, and improvements in tool and die design and lubrication
Surface condition of sheet	Depends on sheet-rolling practice; important in sheet forming, as it can cause tearing and poor surface quality

exponent, n, shown in Eq. (2.8). Thus, a high n value indicates large uniform elongation (see also Table 2.3). Necking may be *localized* or it may be *diffuse*, depending on the *strain-rate sensitivity*, m, of the material, as given in Eq. (2.9); the higher m is, the more diffuse the neck becomes. A diffuse neck is desirable in sheet-forming operations because it is associated with higher formability. In addition to uniform elongation and necking, the **total elongation** of the specimen, in terms of that for a 50-mm gage length, is also a significant factor in the formability of sheet metals.

Yield-point Elongation. Low-carbon steels and some aluminum–magnesium alloys exhibit a behavior called *yield-point elongation*, having both upper and lower yield points (Fig. 16.12a). This phenomenon results in **Lüder's bands** (also called *stretcher-strain marks* or *worms*) on the sheet (Fig. 16.12b). They are elongated depressions on the surface of the sheet, such as can be observed by looking at the bottom of steel cans containing common household products (Fig. 16.12c). The marks may be objectionable, because coarseness on the surface degrades appearance and may also cause difficulties in subsequent coating and painting operations.

The usual method of avoiding Lüder's bands is to eliminate or reduce yield-point elongation, by an additional reduction in the sheet thickness of 0.5% to 1.5% by cold rolling, known as **temper** or **skin rolling**. However, because of *strain aging*, the yield-point elongation reappears after a few days at room temperature, or after a few hours at higher temperatures. Consequently, the material should be formed within a certain time limit, which depends on the type of sheet.

Anisotropy. An important property that influences sheet-metal forming is *anisotropy* (*directionality*) of the sheet (Fig. 16.17). Anisotropy is acquired during the thermomechanical processing of the sheet. There are



(a) (b)

Figure 16.12: (a) Yield-point elongation in a sheet-metal specimen. (b) Stretcher strains at the bottom of a steel can for household products.

two types of anisotropy: *crystallographic anisotropy*, also called *preferred orientation* of the grains) and *mechanical fibering*, which is the alignment of impurities, inclusions, and voids throughout the thickness of the sheet. The relevance of anisotropy is described further in Section 16.4.

Grain Size. As described in Section 1.5, grain size affects mechanical properties and influences the surface appearance of the formed part (*orange peel*). The smaller the grain size, the stronger is the metal, and the coarser the grain, the rougher is the surface appearance. An ASTM grain size of 7 or finer (see Table 1.1) is preferred for general sheet-forming operations.

Dent Resistance of Sheet Metals. Dents are a common feature on vehicles, appliances, and office furniture. They usually are caused by *dynamic forces* from moving objects hitting the sheet metal. In typical automotive panels, for example, impact velocities range up to 45 m/s. Thus, it is the *dynamic yield strength* (yield strength under high deformation rates), rather than the static yield strength, that is the significant strength parameter.

The factors significant in dent resistance have been shown to be yield stress, S_y, sheet metal thickness, T, and shape of the panel. *Dent resistance* is then expressed by a combination of material and geometrical parameters:

$$\text{Dent resistance} = \frac{S_y^2 T^4}{S},$$

where S is the panel stiffness, which, in turn, is defined as

$$S = (E)(T^a)(\text{shape}),$$

where a ranges from 1 to 2 for most panels. As for shape, the flatter the panel, the higher is dent resistance, because of the sheet's flexibility. Thus, dent resistance (a) increases with increasing strength and thickness of the sheet, (b) decreases with increasing elastic modulus and stiffness, and (c) decreases with decreasing curvature of the sheet. Consequently, panels rigidly held at their edges have lower dent resistance, because of their higher stiffness, than those held with a set of springs.

Dynamic forces tend to cause *localized* dents, whereas static forces tend to *diffuse* the dented area. This phenomenon may be demonstrated by trying to dent a piece of flat sheet metal by pushing a ball-peen hammer against it as contrasted to striking it with the hammer. Note how localized the dent will be in the latter case.

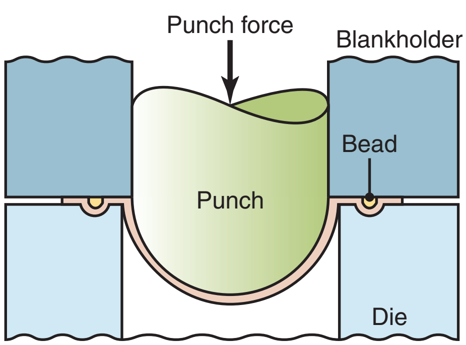
16.4 Formability Tests for Sheet Metals

Sheet-metal formability is generally defined as the ability of the sheet metal to undergo a required shape change without failure, such as cracking, wrinkling, necking, or tearing. As will be noted throughout the rest of this chapter, and depending on part shape, sheet metals may undergo two basic modes of deformation: (1) *stretching* and (2) *drawing*. There are important distinctions between these two modes, and different parameters are involved in determining formability under different conditions. This section describes the methods generally used to estimate formability.

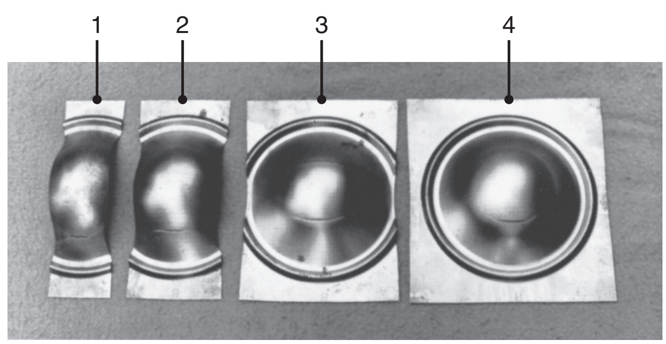
Cupping Tests. The earliest tests developed to predict sheet-metal formability were cupping tests (Fig. 16.13a). In the *Erichsen* test, the sheet specimen is clamped between two circular, flat dies, and a steel ball or a round-tipped punch is forced into the sheet until a crack begins to appear on the bottom of the stretched specimen. The *punch depth*, d, at which a crack appears is a measure of the formability of the sheet. Although this and other similar tests are easy to perform, they do not simulate the exact and often complex conditions of actual forming operations, and hence are not particularly reliable.

Forming-limit Diagrams. An important test for determining the formability of sheet metals is the development of *forming-limit diagrams* (FLD), as shown in Fig. 16.14. For a particular sheet metal, this diagram is constructed by first marking the flat sheet with a grid pattern of circles (Fig. 16.15), using chemical or photoprinting techniques. The blank is then stretched over a round punch (Fig. 16.13a), and the deformation of the circles is observed and measured in the region where failure (*necking* or *tearing*) has occurred. The circles typically are 2.5 to 5 mm in diameter; for better accuracy of measurement, they could be made as small as is practical.

In order to simulate the typically unequal stretching encountered in actual sheet-forming operations, the flat specimens are cut to varying widths (Fig. 16.13b), and then tested. Note that a square specimen (farthest right in the figure) produces *equal biaxial stretching* (such as that achieved in blowing up a spherical balloon), whereas a narrow specimen (farthest left in the figure) basically undergoes a state of *uniaxial stretching* (that is, simple tension). After a series of such tests is performed on a particular type of sheet metal, a forming-limit diagram is constructed, identifying the boundaries between failure and safe zones (Fig. 16.14b).



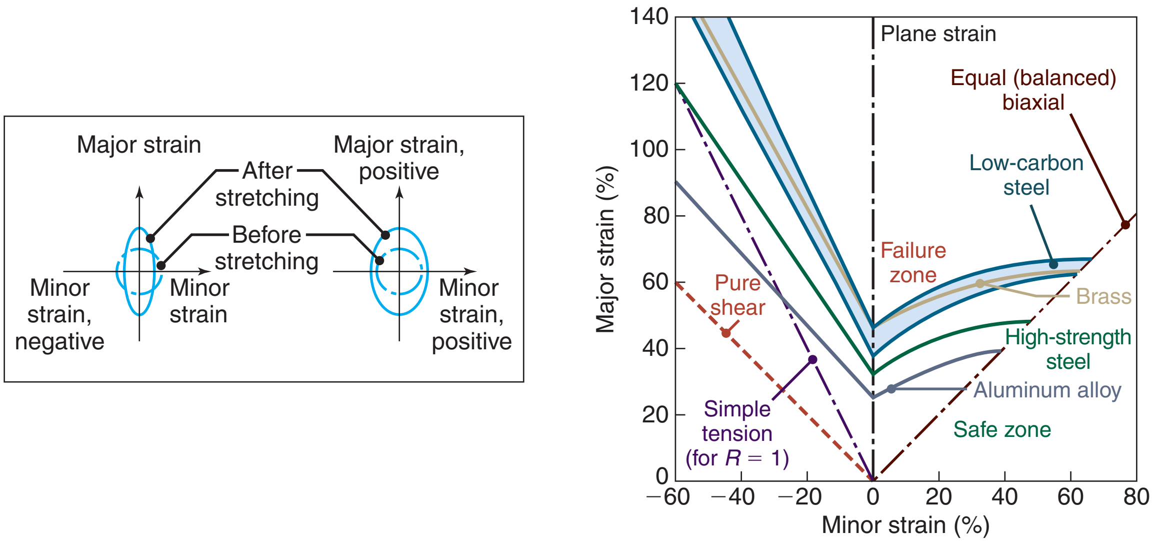
(a)



(b)

Figure 16.13: (a) A cupping test (the Erichsen test) to determine the formability of sheet metals. (b) Bulge-test results on steel sheets of various widths. The specimen farthest left is subjected to, basically, simple tension. The specimen that is farthest right is subjected to equal biaxial stretching. *Source:* (b) Courtesy of ArcelorMittal.

To develop such a diagram, the major and minor engineering strains are obtained, as measured from the deformation of the original circles. Note in Fig. 16.14a that an original (round) circle has deformed into an ellipse, the *major axis* of which represents the major direction and magnitude of stretching. The major strain is the *engineering strain* in this direction, and is always *positive*, because the sheet is being stretched. The *minor axis* of the ellipse represents the minor direction and magnitude of strain in the *transverse* direction, which may have undergone either stretching or shrinking.



(a) (b)

Figure 16.14: (a) Strains in deformed circular grid patterns. (b) Forming-limit diagrams (FLD) for various sheet metals. Although the major strain is always positive (stretching), the minor strain may be either positive or negative. R is the normal anisotropy of the sheet, as described in Section 16.3. *Source:* After S.S. Hecker and A.K. Ghosh.

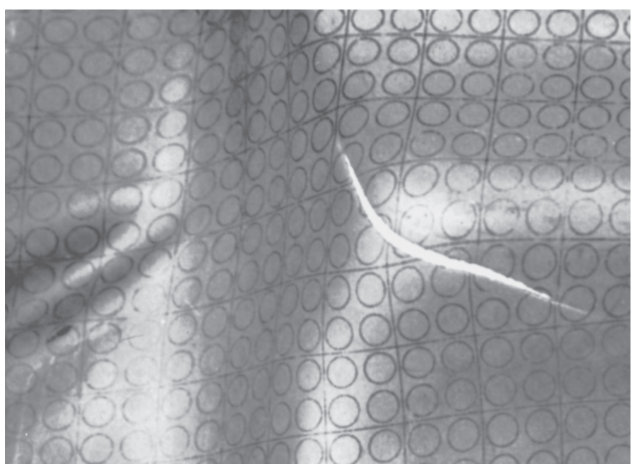


Figure 16.15: The deformation of the grid pattern and the tearing of sheet metal during forming. The major and minor axes of the circles are used to determine the coordinates on the forming-limit diagram in Fig. 16.14b. *Source:* After S.P. Keeler.

The latest systems develop the FLD using cameras and software that track surface strains during the test. The specimens generally have a spackle or spotted surface sprayed onto them, thereby avoiding the printing of circles. Moreover, this method generates many more data points than the printed circle method and reduces uncertainty in constructing the FLD.

Note in the diagrams that the minor strain can be either *positive* or *negative*. For example, if a circle is placed in the center of a tensile-test specimen, and then stretched uniaxially (simple tension), the specimen becomes narrower as it is stretched, due to the Poisson effect (Section 2.2.1); thus, the minor strain is negative. This behavior can easily be demonstrated by stretching a common rubber band and observing the dimensional changes it undergoes. On the other hand, if a circle is placed on a *spherical* rubber balloon and inflated, the minor and major strains will both be positive and equal in magnitude.

By comparing the surface areas of the original and deformed circles on the formed sheet, it can also be determined whether the *thickness* of the sheet has changed during deformation. Because the volume remains constant in plastic deformation, if the area of a deformed circle is larger than the original, the sheet has become thinner. This phenomenon can be demonstrated easily by blowing up a spherical balloon and noting that it becomes more translucent, hence thinner, as it is stretched.

The data thus obtained from different locations in each of the samples shown in Fig. 16.13b are then plotted, as shown in Fig. 16.14b. The curves represent the boundaries between *failure zones* and *safe zones* for each type of sheet metal. As can be noted, the higher the curve, the better is the formability of that particular sheet metal. Different materials and conditions, such as cold worked or heat treated, will have different forming-limit diagrams. Taking the aluminum alloy in Fig. 16.14b as an example, if a circle in a particular location on the specimen has undergone major and minor strains of plus 20% and minus 10%, respectively, there would be no tear in that location of the specimen. On the other hand, if at another location on the sheet the major and minor strains were plus 80% and minus 40%, respectively, at another location, there would be a tear in that particular location of the specimen. An example of a formed sheet-metal part with a grid pattern is shown in Fig. 16.15; note the deformation of the circular patterns in the vicinity of the tear.

It is important to note in forming-limit diagrams that a compressive minor strain of, say, 20% is associated with a higher major strain than is a tensile (positive) minor strain of the same magnitude. In other words, it is desirable for the minor strain to be negative, that is, shrinking in the minor direction. In forming complex parts, special tooling can be designed to take advantage of the beneficial effect of negative minor strains on formability.

The effect of sheet thickness on FLD is to *raise* the curves in Fig. 16.14b. Thus, the thicker the sheet, the higher is its formability curve, and the more formable the sheet. In actual forming operations, however, a thick blank will not bend around small radii as easily without cracking, as described in Section 16.5 on *bending*.

Friction and lubrication at the interface between the punch and the sheet metal also are important factors in the test results. With well-lubricated interfaces, the strains in the sheet become distributed more uniformly over the punch. Also, as expected, and depending on the material and surface defects such as notch sensitivity, surface scratches (see *notch sensitivity*, Section 2.9), deep gouges, and blemishes can significantly reduce formability and, thereby, lead to premature tearing and failure of the part.

A procedure that has been followed with some success to improve sheet-metal formability is to control and vary process parameters *during* forming. For example, deep drawability (Section 16.7.1) can be improved by varying the *blankholder force* (see Fig. 16.32) during deep drawing. This force can be changed with position in the die if, for example, *multiple actuators* are used for the blankholder, or it can be modified with respect to time. Optimized press velocity profiles programmed into servo presses (Section 14.8) can also improve formability.

16.5 Bending Sheets, Plates, and Tubes

Bending is one of the most common forming operations, as evidenced by observing automobile bodies, exhaust pipes, appliances, paper clips, or file cabinets. Bending also imparts stiffness to a part, by increasing its moment of inertia. Note, for example, how corrugations, flanges, beads, and seams improve the stiffness

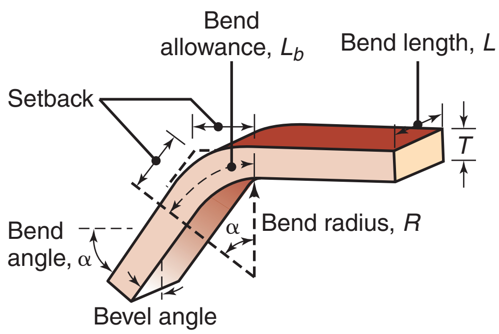


Figure 16.16: Bending terminology; note that the bend radius is measured to the inner surface of the bent part.

of cans, containing a variety of food and liquids, without adding weight. As a specific example, observe the diametral stiffness of a can with and without circumferential beads (see also Section 16.7).

The terminology used in bending sheet or plate is given in Fig. 16.16. Note that the outer fibers of the material are in tension, while the inner fibers are in compression. Because of the Poisson effect, the width of the part (*bend length*, L) becomes smaller in the outer region, and larger in the inner region compared to the original width, as can be seen in Fig. 16.17c. This phenomenon can easily be illustrated by bending a rectangular rubber eraser and observing the changes in its shape.

As shown in Fig. 16.16, the **bend allowance**, L_b, is the length of the *neutral axis* in the bend; it is used to determine the length of the blank for a part to be bent. The position of the neutral axis depends on bend radius and bend angle, as described in texts on mechanics of solids. An approximate formula for the bend allowance is

$$L_b = \alpha\,(R + kT), \tag{16.2}$$

where α is the bend angle (in radians), T is the sheet thickness, R is the bend radius, and k is a constant, which in practice typically ranges from 0.33 (for $R < 2T$) to 0.5 (for $R > 2T$). Note that for the ideal case, the neutral axis is at the center of the sheet thickness, $k = 0.5$, and hence,

$$L_b = \alpha\left[R + \left(\frac{T}{2}\right)\right]. \tag{16.3}$$

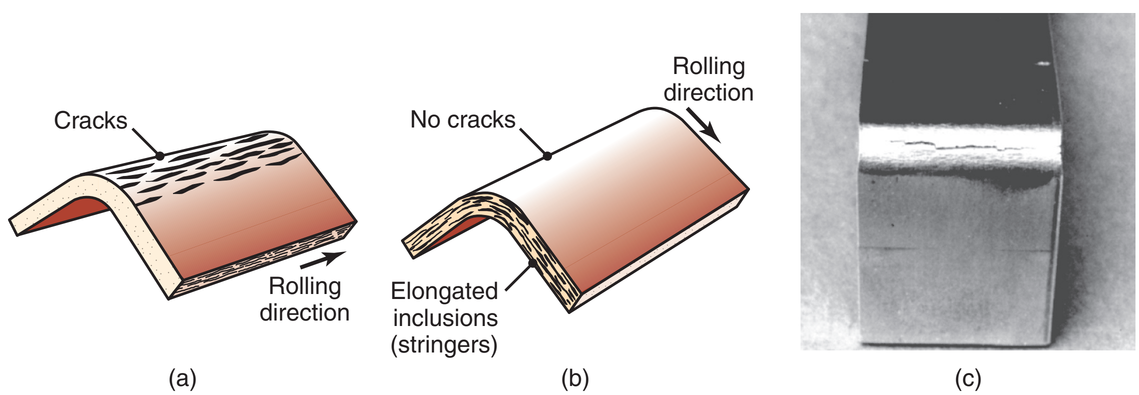


Figure 16.17: (a) and (b) The effect of elongated inclusions (stringers) on cracking as a function of the direction of bending with respect to the original rolling direction of the sheet. (c) Cracks on the outer surface of an aluminum strip bent to an angle of 90°. Note also the narrowing of the top surface in the bend area due to the Poisson effect.

Table 16.3: Minimum Bend Radius for Various Metals at Room Temperature.

	Condition	
Material	**Soft**	**Hard**
Aluminum alloys	0	$6T$
Beryllium copper	0	$4T$
Brass (low-leaded)	0	$2T$
Magnesium	$5T$	$13T$
Steels		
Austenitic stainless	$0.5T$	$6T$
Low-carbon, low-alloy, and HSLA	$0.5T$	$4T$
Titanium	$0.7T$	$3T$
Titanium alloys	$2.6T$	$4T$

Minimum Bend Radius. The radius at which a crack first appears at the outer fibers of a sheet being bent is referred to as the *minimum bend radius*. It can be shown that the engineering strain on the outer and inner fibers of a sheet during bending is given by the expression

$$e = \frac{1}{(2R/T) + 1}. \tag{16.4}$$

Thus, as R/T decreases (i.e., as the ratio of the bend radius to the thickness becomes smaller), the tensile strain at the outer fiber increases, and the material eventually develops cracks (Fig. 16.17). The bend radius usually is expressed in terms of the thickness, such as $2T$, $3T$, and $4T$ (see Table 16.3). Thus, a $3T$ minimum bend radius indicates that the smallest radius to which the sheet can be bent, without cracking, is three times its thickness.

It has been shown that there is an inverse relationship between *bendability* and the tensile reduction of the area, r, of the material (Fig. 16.18). The *minimum bend radius*, R_{min}, is, approximately,

$$R_{\text{min}} = T\left(\frac{50}{r} - 1\right). \tag{16.5}$$

Thus, for $r = 50$, the minimum bend radius is zero; that is, the sheet can be folded over itself, called *hemming* (see Fig. 16.23), in much the same way as a piece of paper is folded. To increase the bendability of metals, their tensile reduction of area can be increased either by heating or by bending it in a high-pressure environment, which improves the ductility of the material (see *hydrostatic stress*, Section 2.2.8).

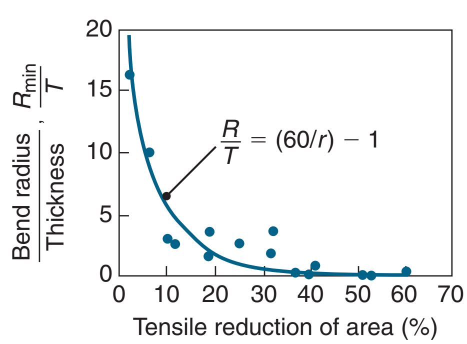


Figure 16.18: Relationship between R_{min}/T and tensile reduction of area for sheet metals. Note that sheet metal with a 50% tensile reduction of area can be bent over itself in a process like the folding of a piece of paper without cracking. *Source:* After J. Datsko and C.T. Yang.

Bendability also depends on the *edge condition* of the sheet. Since rough edges are areas of stress concentration, bendability decreases as edge roughness increases. Another significant factor is the amount, shape, and hardness of *inclusions* present in the sheet metal and the amount of cold working that the edges have undergone during shearing. Because of their pointed shape, inclusions in the form of stringers are more detrimental than globular-shaped inclusions (see also Fig. 2.23). The resistance to edge cracking during bending can be significantly increased by removing the cold-worked regions, by (a) shaving or machining the edges of the part (see Fig. 16.9) or (b) annealing the sheet to improve its ductility.

Anisotropy of the sheet is another important factor in bendability. Cold rolling increases the anisotropy of the sheet by *preferred orientation* or by *mechanical fibering* due to the alignment of impurities, inclusions, and voids (Fig. 1.12). Prior to laying out or *nesting* the blanks (see Fig. 16.60) for subsequent bending or forming, caution should be exercised to cut, as much as possible, in the optimum direction from a rolled sheet.

Springback. Because all materials have a finite modulus of elasticity, plastic deformation is always followed by some elastic recovery when the load is removed (Fig. 2.3). In bending, this recovery is called *springback*; it can easily be demonstrated by bending and then releasing a piece of sheet metal or wire. As noted in Fig. 16.19, the final bend angle of a sheet metal after springback is smaller than the angle to which the sheet was bent, and the final bend radius is larger than before springback.

Springback can be calculated approximately in terms of the radii R_i and R_f (Fig. 16.19) as

$$\frac{R_i}{R_f} = 4\left(\frac{R_i S_y}{ET}\right)^3 - 3\left(\frac{R_i S_y}{ET}\right) + 1. \tag{16.6}$$

Note from this formula that springback increases as the R/T ratio and the yield strength, S_y, of the material increase, and as the elastic modulus, E, decreases.

In V-die bending (Figs. 16.20 and 16.21), it is possible for the material to also exhibit *negative springback*. This is a condition caused by the nature of the deformation occurring within the sheet metal just when the punch completes the bending operation at the end of the stroke. Negative springback does not occur in *air bending*, also called *free bending* (Fig. 16.22a), because of the absence of constraints that a V-die imposes on the bend area.

Compensation for Springback. Springback in forming operations usually is compensated by *overbending* the part (Fig. 16.20a and b), although several trials may be necessary to obtain the desired results. Another method is to *coin* the bend area, by subjecting it to highly localized compressive stresses between the tip of the punch and the die surface (Fig. 16.20c and d); the technique is also called *bottoming the punch*. In another method, the part is subjected to *stretch bending*, in which the part is under external tension while being bent (see also *stretch forming*, Section 16.6).

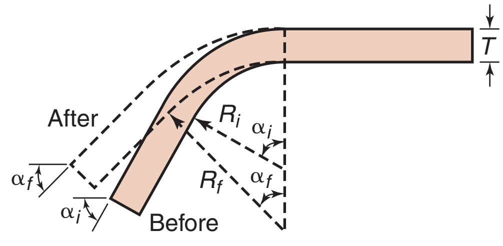


Figure 16.19: Springback in bending. The part tends to recover elastically after bending, and its bend radius becomes larger. Under certain conditions, it is possible for the final bend angle to be smaller than the original angle (negative springback).

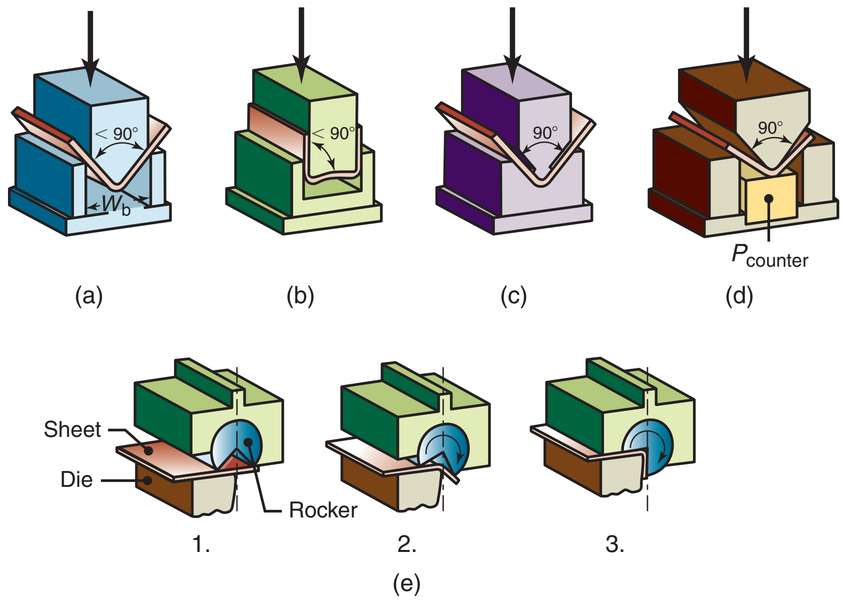


Figure 16.20: Methods of reducing or compensating for springback in bending operations.

Bending Force. The bending force for sheets and plates can be estimated by assuming that the process is one of simple bending of a rectangular beam (as described in texts on mechanics of solids). Thus, the bending force is a function of the yield strength of the material, S_y, the length of the bend, L, the thickness of the sheet, T, and the die opening, W (see Fig. 16.21). Excluding friction, the *maximum bending force*, P, is

$$P = \frac{kS_y LT^2}{W}, \tag{16.7}$$

where the factor k ranges from about 0.3 for a wiping die, to about 0.7 for a U-die, to about 1.3 for a V-die (Fig. 16.21), and S_y is the yield strength of the material. For situations where the punch-tip radius and the sheet thickness are relatively small compared to the die opening, W, the maximum bending force is given by

$$P = \frac{S_{\text{ut}} LT^2}{W} \tag{16.8}$$

where S_{ut} is the ultimate tensile strength of the sheet.

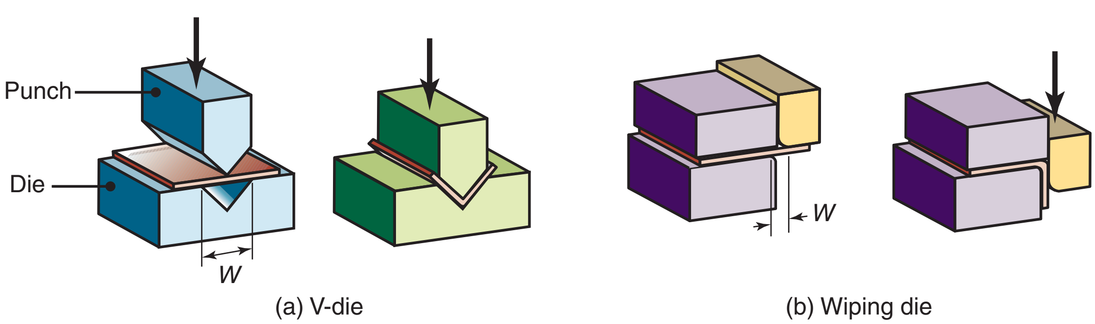


Figure 16.21: Common die-bending operations showing the die-opening dimension, W, used in calculating bending forces.

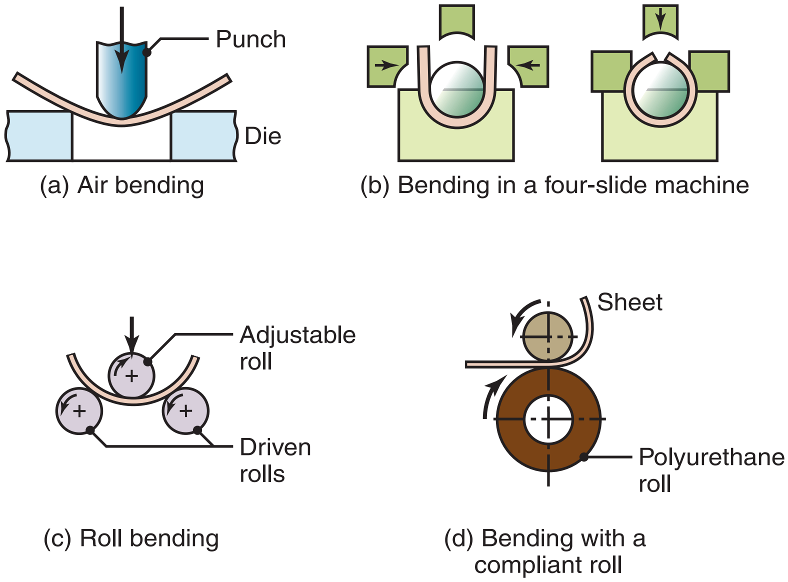


Figure 16.22: Examples of various bending operations.

The force in die bending varies throughout the bending cycle. It increases from zero to a maximum, and it may even decrease as the bend is being completed, but then it increases sharply as the punch reaches the bottom of its stroke (bottoming). However, in air bending (Fig. 16.22a) the force does not increase again after it begins to decrease, because the sheet is not subjected to any resistance in its movement downward.

16.6 Miscellaneous Bending and Related Forming Operations

Press-brake Forming. Sheet metal or plate can easily be bent using simple fixtures in a press. Sheets or narrow strips that are 7 m or even wider usually are bent in a *press brake* (Fig. 16.23). The machine utilizes long dies, in a mechanical or hydraulic press, and is particularly suitable for small production runs. As can be seen in Fig. 16.23, the tooling is simple, the motions are only up and down, and the process is easily adaptable to a wide variety of part shapes. The operation can be easily automated for low-cost, high-production runs. *Die materials* for press brakes range from hardwood for low-strength materials and small-production runs, to carbides, for strong and abrasive sheet metals (such as carbon steel). For most applications, the dies are made of carbon steel or gray iron.

Bending in a Four-slide Machine. Relatively short pieces can be bent on a machine such as the one shown in Fig. 16.22b. The lateral movements of the dies are controlled and synchronized with the vertical die movement. This process is typically used for making seamed tubing and conduits, bushings, fasteners, and various machinery components.

Roll Bending. In this process (Fig. 16.22c), plates are bent using a set of rolls, where curvatures are controlled by adjusting the distance between the three rolls. Roll bending is used extensively for bending plates, applications as boilers, cylindrical pressure vessels and tanks, and curved structural members. Figure 16.22d shows the bending of a strip, with a *compliant roll*, made of polyurethane, which conforms to the shape of the strip as the hard upper roll presses upon it.

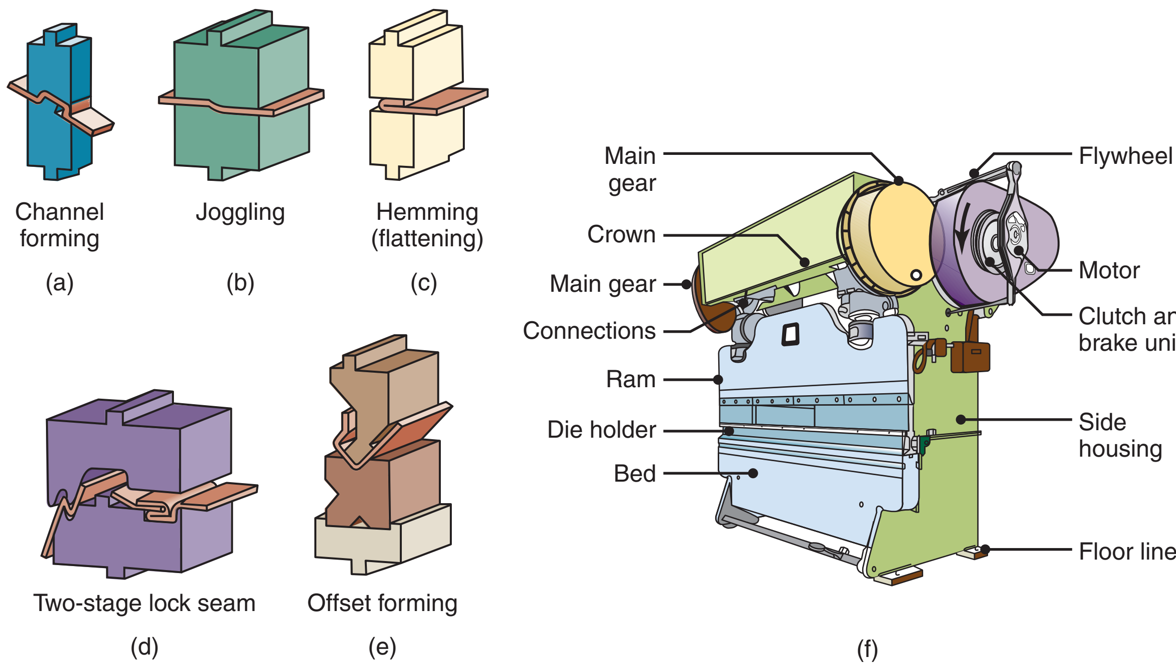


Figure 16.23: (a) through (e) Schematic illustrations of various bending operations in a press brake. (f) Schematic illustration of a press brake. *Source:* Courtesy of Verson Allsteel Company.

Beading. In *beading*, the periphery of the sheet metal is bent in the cavity of a die (Fig. 16.24). The bead imparts stiffness to the part by increasing the moment of inertia of that section. Moreover, beads improve the appearance of parts and eliminate exposed sharp edges, which may be hazardous.

Flanging. This is a process of bending the edges of sheet metals, usually to 90° (see also Section 16.7). In **shrink flanging** (Fig. 16.25a), the flange is subjected to compressive hoop stresses; if excessive, however, the stresses can cause the flange periphery to wrinkle. The wrinkling tendency increases with decreasing radius of curvature of the flange. In **stretch flanging**, the flange periphery is subjected to tensile stresses; if excessive, however, they can lead to cracking along the periphery of the flange.

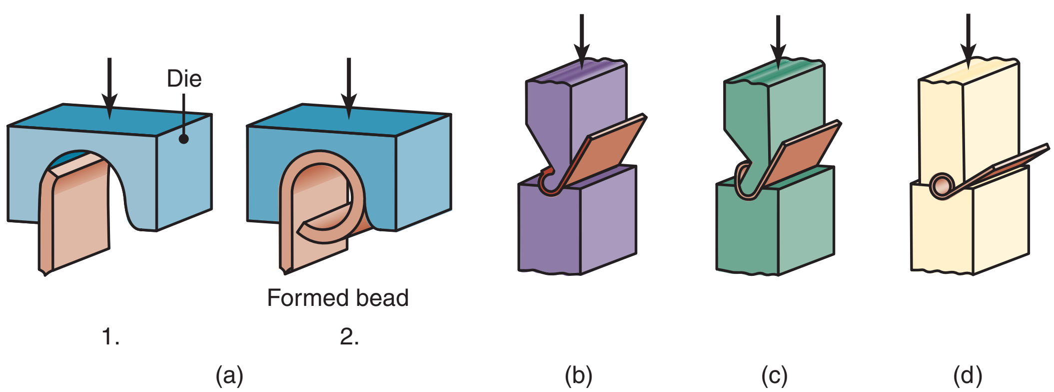


Figure 16.24: (a) Bead forming with a single die. (b) through (d) Bead forming with two dies in a press brake.

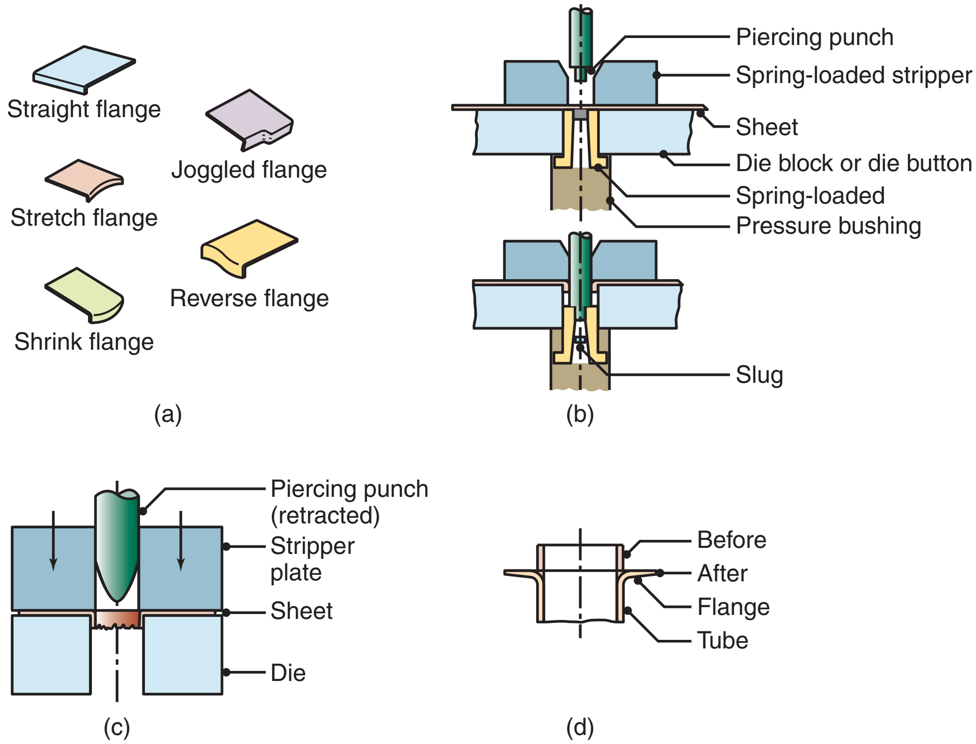


Figure 16.25: Various flanging operations. (a) Flanges on flat sheet. (b) Dimpling. (c) The piercing of sheet metal to form a flange. In this operation, a hole does not have to be pre-punched before the punch descends. Note, however, the rough edges along the circumference of the flange. (d) Flanging of a tube; note the thinning of the edges of the flange.

Roll Forming. Also called *contour-roll forming* or *cold-roll forming*, this process is used for forming continuous lengths of sheet metal and for large production runs. As it passes through a set of driven rolls, the metal strip is bent in consecutive stages (Fig. 16.26). The roll-formed strip is then sheared into specific lengths and stacked.

Typical roll-formed products are door and picture frames, panels, channels, gutters, siding, pipes, and tubing with lock seams (Section 32.5). The length of the part is limited only by the amount of sheet metal supplied to the rolls from a coiled stock. Sheet thickness typically ranges from about 0.125 to 20 mm. Forming speeds are generally below 1.5 m/s, although they can be much higher for specialized applications.

In designing the rolls and their sequence, dimensional tolerances, springback, tearing, and buckling of the strip have to be considered. The rolls generally are made of carbon steel or gray iron; they may be chromium plated, to reduce wear of the rolls and for improved surface finish of the formed product. Lubricants may be used to reduce wear, improve surface finish, and to cool the rolls and the sheet being formed.

Tube Bending and Forming. Bending and forming tubes and of other hollow sections requires special tooling because of the tendency for buckling and folding, as can be demonstrated by bending copper tubing or plastic soda straw. The oldest method of bending a tube or pipe is to first pack it with loose particles (commonly sand), and then bend it in a suitable fixture. The function of the loose filler is to prevent the tube from buckling inward; after bending, the sand is simply shaken out. Tubes also can be plugged with

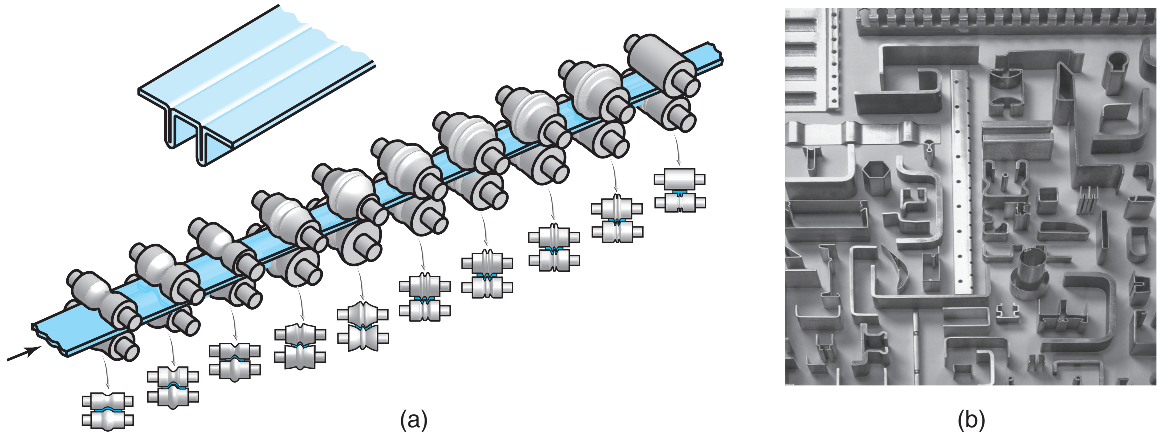


Figure 16.26: (a) Schematic illustration of the roll-forming process. (b) Examples of roll-formed cross sections. *Source:* (b) Courtesy of Voestalpine Roll Forming Corporation.

various flexible *internal mandrels* (Fig. 16.27), serving the same purpose as sand. Because of its lower tendency for buckling, a relatively thick tube can be bent safely without the use of fillers or plugs (see also *tube hydroforming*, Section 16.8).

The beneficial effect of forming metals under high *compressive stresses* is demonstrated in Fig. 16.28 for bending a tube with relatively sharp corners. Note that, in this operation, the tube is subjected to longitudinal compressive stresses, which reduce the stresses in the outer fibers in the bend area, thus improving the bendability of the material (see also Section 2.2.8).

Dimpling, Piercing, and Flaring. In *dimpling* (Fig. 16.25b), a hole is first punched and then expanded into a flange. Flanges may also be made by *piercing* using a shaped punch (Fig. 16.25c); tube ends can be flanged by a similar process (Fig. 16.25d). When the bend angle is less than 90°, as in fittings with conical ends,

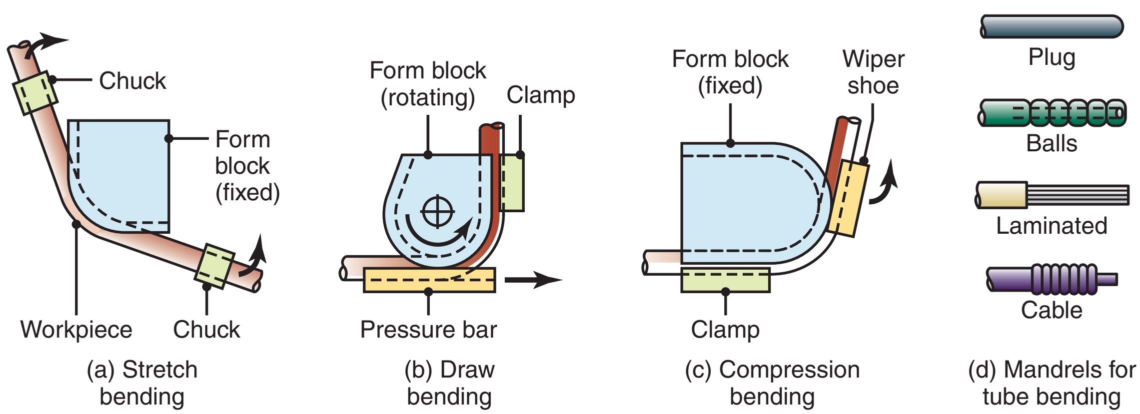


Figure 16.27: Methods of bending tubes. Internal mandrels or filling of tubes with particulate materials, such as sand, are often necessary to prevent collapse of the tubes during bending. Tubes also can be bent by a technique in which a stiff, helical tension spring is slipped over the tube. The clearance between the outer diameter of the tube and the inner diameter of the spring is small; thus, the tube cannot kink and the bend is uniform.

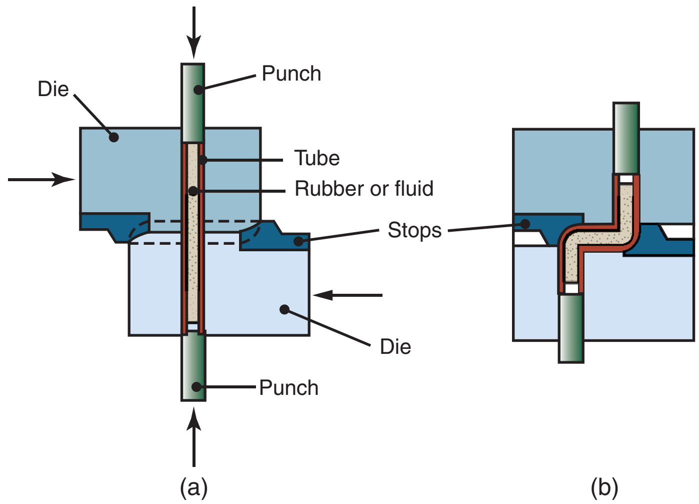


Figure 16.28: A method of forming a tube with sharp angles, using an axial compressive force. Note that the tube is supported internally with rubber or fluid to avoid collapsing during forming. *Source:* After J.L. Remmerswaal and A. Verkaik.

the process is called *flaring*. The condition of the edges (Fig. 16.3) is important in these operations, because stretching the material causes high tensile stresses along the periphery (tensile hoop stresses), which can lead to cracking and tearing of the flange.

As the ratio of flange diameter to hole diameter increases, the strains increase proportionately. Depending on the roughness of the edge, there will therefore be a tendency for cracking along the outer periphery of the flange. To reduce this possibility, sheared or punched edges could be shaved off with a sharp tool (Fig. 16.9) to improve the surface finish of the edge.

Hemming and Seaming. In the *hemming* process, also called *flattening*, the edge of the sheet is folded over itself (Fig. 16.23c). Hemming increases the stiffness of the part, improves its appearance, and eliminates sharp edges. *Seaming* involves joining two edges of sheet metal by hemming (Fig. 16.23d). *Double seams* are made by a similar process using specially shaped rolls for making watertight and airtight joints, such as those in food and beverage containers.

Bulging. This process involves placing a tubular, conical, or curvilinear part into a split-female die, and then expanding the part, usually with a polyurethane plug (Fig. 16.29a). After forming, the punch is retracted, the plug returns to its original shape (by elastic recovery), and the formed part is removed by opening the split dies. Typical products made are coffee and water pitchers, beer barrels, and beads on oil drums. For parts with complex shapes, the plug is shaped, in order to be able to apply higher pressures at critical regions of the part. The major advantages of using polyurethane plugs is that they are highly resistant to abrasion and wear, and do not damage the surface finish of the part being formed (see also Section 16.8).

Segmented Dies. These dies consist of individual segments that are placed inside the part to be formed, and expanded mechanically in a radial direction; the segments are then retracted to remove the formed part. These dies are relatively inexpensive, and they can be used for large production runs.

Stretch Forming. In this process, the sheet metal is clamped along its edges and then stretched over a male die, called a *form block* or *form punch*. The die can move upward, downward, or sideways, depending on the particular design of the machine (Fig. 16.30). Stretch forming is used primarily to make aircraft wing-skin panels, fuselages, and boat hulls. Aluminum skins for the Boeing 767 and 757 aircraft, for example, are

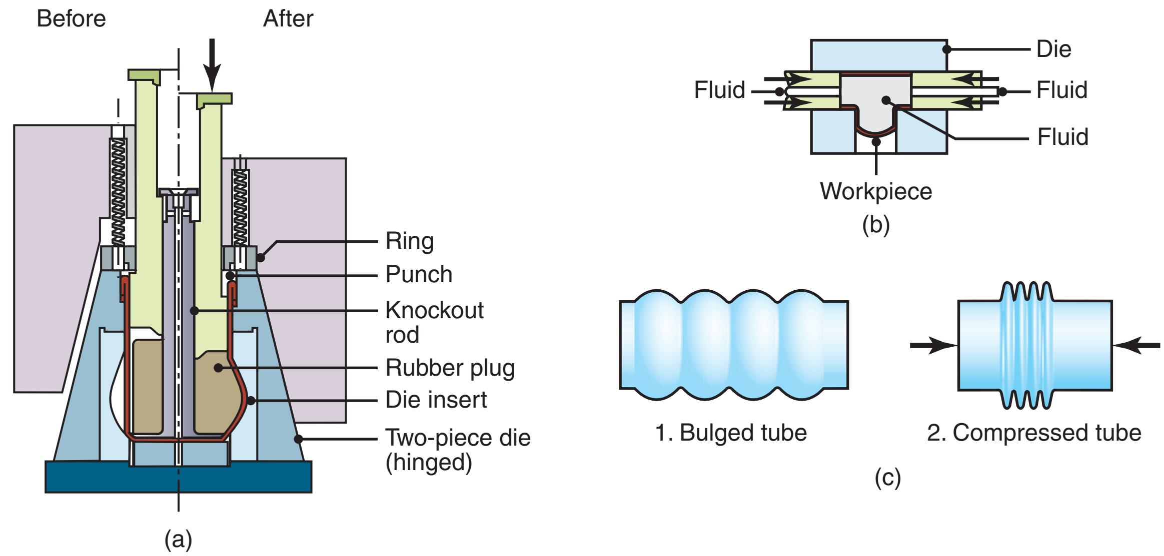


Figure 16.29: (a) The bulging of a tubular part with a flexible plug. Water pitchers can be made by this method. (b) Production of fittings for plumbing by expanding tubular blanks under internal pressure. The bottom of the piece is then punched out to produce a T. (c) Steps in manufacturing bellows.

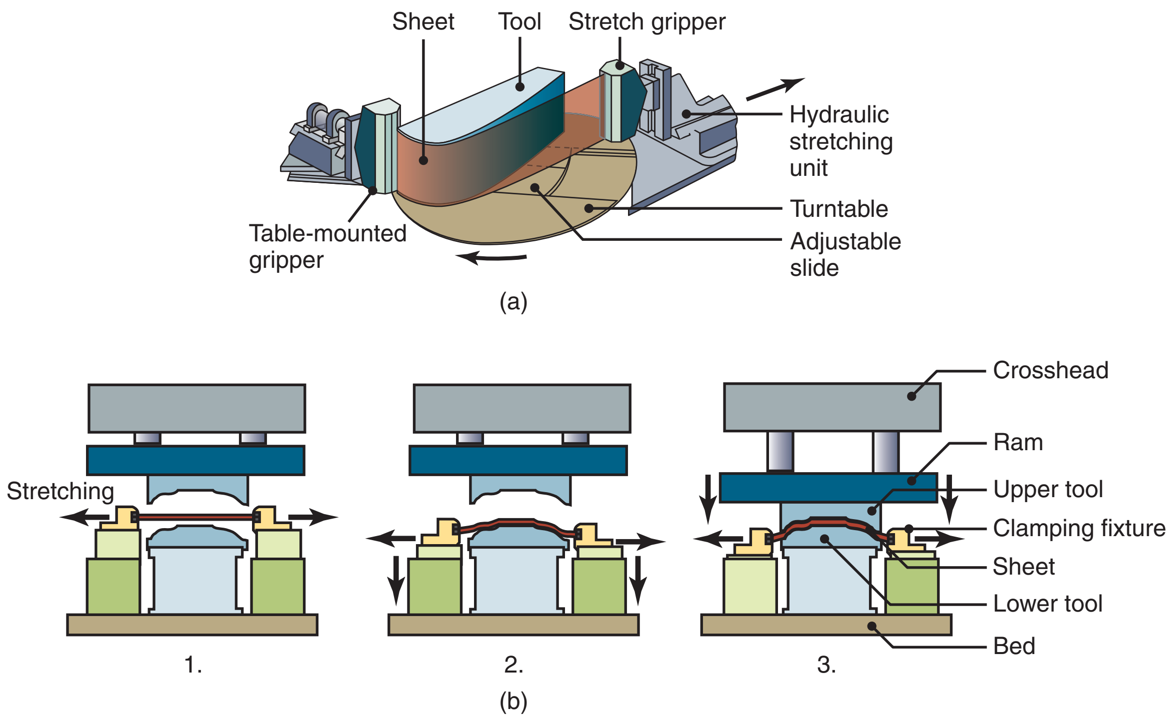


Figure 16.30: Schematic illustration of a stretch-forming process. Aluminum skins for aircraft can be made by this method. *Source:* (a) Courtesy of Cyril Bath Co.

made by stretch forming, with a tensile force of 9 MN. The rectangular sheets are 12 m × 2.5 m × 6.4 mm. Although this process is generally used for low-volume production, it is versatile and economical, particularly for applications in the aerospace industry.

In most operations, the blank is a rectangular sheet, clamped along its narrower edges and stretched lengthwise, thus allowing the material to shrink in its width direction. Controlling the amount of stretching is important in order to prevent tearing. Stretch forming cannot produce parts with sharp contours or with reentrant corners. Accessory equipment can be used in conjunction with stretch forming, including further forming with both male and female dies while the part is under tension. Dies for stretch forming are generally made of zinc alloys, steel, plastics, or hard wood. Most applications require little or no lubrication.

16.7 Deep Drawing

Numerous sheet-metal parts are cylindrical or box shaped, such as pots and pans, all types of containers for food and beverages (Fig. 16.31), stainless-steel kitchen sinks, canisters, and automotive fuel tanks. Such parts usually are made by deep drawing, a process in which a punch forces a flat sheet-metal blank into a die cavity, as shown in Fig. 16.32a. Deep drawing is one of the most important and widely used sheet metalworking processes.

In deep drawing, a round sheet-metal blank is placed over a circular die opening, and is held in place with a **blankholder**, or *hold-down ring* (Fig. 16.32b). The punch travels downward, forcing the blank into the die cavity, thus forming a cup. The major variables in this process are (a) properties of the sheet metal; (b) ratio of blank diameter, D_o to punch diameter, D_p; (c) clearance, c, between punch and die; (d) punch radius, R_p; (e) die-corner radius, R_d; (f) blankholder force; and (g) friction and lubrication between all contacting interfaces.

During the drawing operation, the movement of the blank into the die cavity induces compressive circumferential (hoop) stresses in the flange, which tend to cause the flange to wrinkle during drawing. This phenomenon can be demonstrated simply by trying to force a circular piece of paper into a round cavity. Wrinkling can be reduced or eliminated if a blankholder is pressed downward with a certain force. In order to improve performance, the magnitude of this force can be controlled as a function of punch travel or its location in the blankholder.

Because of the number of variables involved, the *punch force*, F, is difficult to calculate directly. It has been shown, however, that the *maximum punch force*, $F_{\max}$ can be estimated from the formula

$$F_{\max} = \pi D_p T S_{\text{ut}} \left[\left(\frac{D_o}{D_p} \right) - 0.7 \right], \tag{16.9}$$

where the nomenclature is the same as that in Fig. 16.32b. It can be seen that the force increases with increasing blank diameter, thickness, strength, and the ratio (D_o/D_p). The wall of the cup being drawn is subjected principally to a longitudinal (vertical) tensile stress, due to the punch force. Elongation under this stress causes the cup wall to become thinner and, if excessive, it can cause *tearing* of the cup.

16.7.1 Deep Drawability

In a deep-drawing operation, failure generally is a result of *thinning* of the cup wall under the high longitudinal tensile stresses due to the action of the punch. Following the material movement as it flows into the die cavity, it can be seen that the sheet metal (a) must be capable of undergoing a reduction in its width, due to a reduction in diameter, and (b) must also resist thinning under the longitudinal tensile stresses in the cup wall.

Deep drawability is generally expressed by the **limiting drawing ratio** (LDR) as

$$\text{LDR} = \frac{\text{Maximum blank diameter}}{\text{Punch diameter}} = \frac{D_o}{D_p}. \tag{16.10}$$

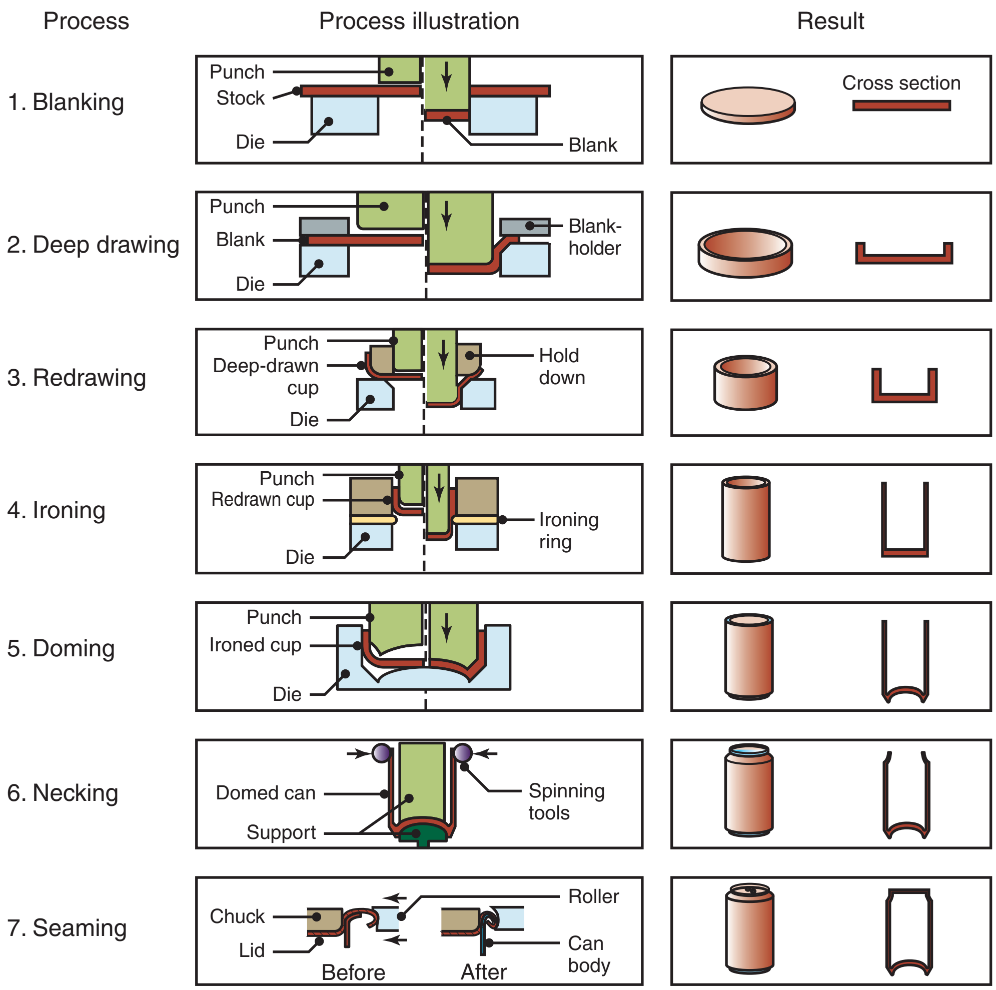


Figure 16.31: The metal-forming processes involved in manufacturing two-piece aluminum beverage cans.

Whether a particular sheet metal can be deep drawn successfully into a round cup has been found to be a function of the **normal anisotropy**, R (also called *plastic anisotropy*), of the sheet metal. Normal anisotropy is defined in terms of the true strains that a tensile test specimen undergoes (Fig. 16.33):

$$R = \frac{\text{Width strain}}{\text{Thickness strain}} = \frac{\epsilon_w}{\epsilon_t}. \tag{16.11}$$

In order to determine the magnitude of R, a specimen is first prepared and subjected to an elongation of 15% to 20%. The true strains that the specimen undergoes are then calculated, in the manner described in Section 2.2. Because cold-rolled sheets are anisotropic in their *planar* direction, the R value of a specimen cut from a rolled sheet will depend on its *orientation* with respect to the rolling direction of the sheet.

An average value, R_{avg} is calculated from the equation

$$R_{\text{avg}} = \frac{R_0 + 2R_{45} + R_{90}}{4} \tag{16.12}$$

where the subscripts are the angles with respect to the rolling direction of the sheet. Some typical R_{avg} values are given in Table 16.4.

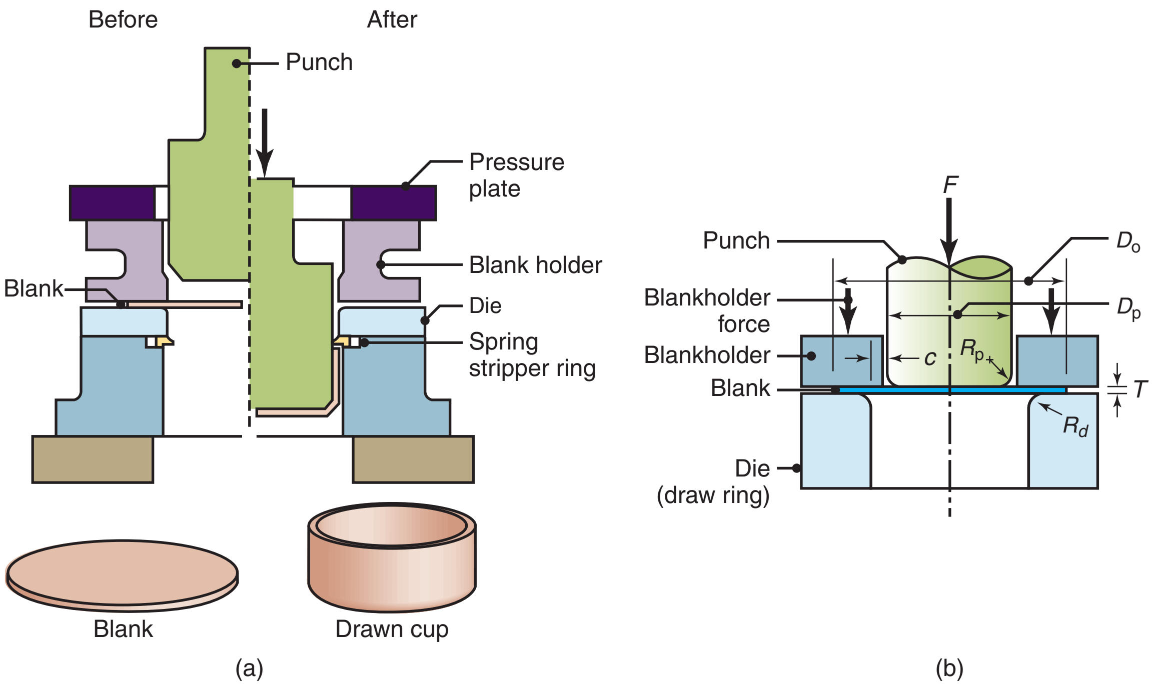


Figure 16.32: (a) Schematic illustration of the deep-drawing process on a circular sheet-metal blank. The stripper ring facilitates the removal of the formed cup from the punch. (b) Process variables in deep drawing. Except for the punch force, F, all the parameters indicated in the figure are independent variables.

The experimentally determined relationship between R_{avg} and the limiting drawing ratio, LDR, is shown in Fig. 16.34. It has been established that no other mechanical property of a sheet metal shows a more consistent relationship to its LDR as does R_{avg}. Thus, by using a simple tensile-test result and obtaining the normal anisotropy of the sheet metal, the limiting drawing ratio of a material can be determined.

Earing. In deep drawing, the edges of cups may become wavy, a behavior called *earing* (Fig. 16.35). Ears are objectionable on deep-drawn cups because they have to be trimmed off; ears serve no useful purpose and they interfere with further processing of the cup, resulting in scrap. Earing is caused by the **planar**

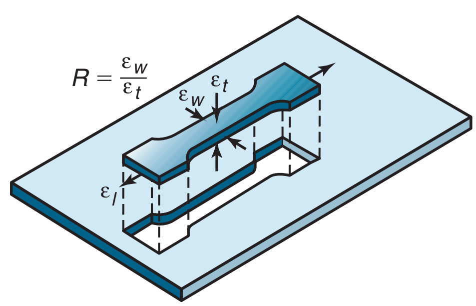


Figure 16.33: Strains on a tensile-test specimen removed from a piece of sheet metal. These strains are then used in determining the normal and planar anisotropy of the sheet metal.

Table 16.4: Typical Ranges of Average Normal Anisotropy, for Various Sheet Metals.

Material	Range of R_{avg}
Zinc alloys	0.4–0.6
Hot-rolled steel	0.8–1.0
Cold-rolled, rimmed steel	1.0–1.4
Cold-rolled, aluminum-killed steel	1.4–1.8
Aluminum alloys	0.6–0.8
Copper and brass	0.6–0.9
Titanium alloys (alpha)	3.0–5.0
Stainless steels	0.9–1.2
High-strength, low-alloy steels	0.9–1.2

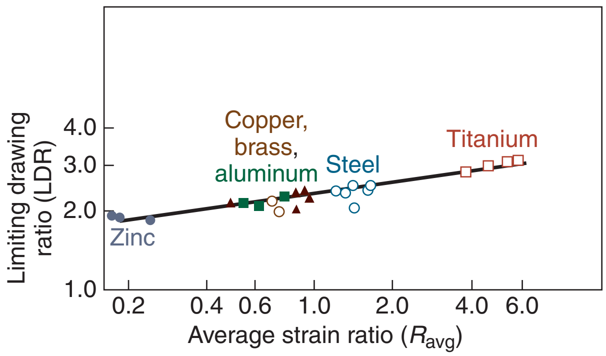


Figure 16.34: The relationship between average normal anisotropy and the limiting drawing ratio for various sheet metals. *Source:* After M. Atkinson.

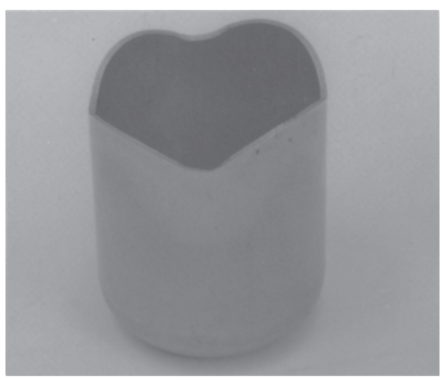


Figure 16.35: Earing in a drawn steel cup, caused by the planar anisotropy of the sheet metal.

anisotropy of the sheet metal, and the number of ears produced may be two, four, or eight, depending on the processing history and microstructure of the material. If the sheet is stronger in its rolling direction than transverse to the rolling direction, and the strength varies uniformly with respect to orientation, then two ears will form. If the sheet has high strength at different orientations, then more ears will form.

The planar anisotropy of the sheet, indicated by ΔR, is defined in terms of directional R values, from the equation

$$\Delta R = \frac{R_0 - 2R_{45} + R_{90}}{2}. \tag{16.13}$$

When $\Delta R = 0$ no ears form, and the height of the ears increases as ΔR increases.

It can be seen that deep drawability is enhanced by a high R_{avg} value and a low ΔR. Generally, however, sheet metals with high R_{avg} also have high ΔR values. Optimum sheet-metal textures can be developed by controlling the type of alloying elements in the material and by adjusting processing parameters during cold rolling of the sheet.

16.7.2 Deep-drawing Practice

Certain guidelines have been established over the years for successful deep-drawing practice. The blankholder pressure is chosen generally as 0.7% to 1.0% of the sum of the yield strength and the ultimate tensile strength of the sheet metal. Too high a blankholder force increases the punch force and causes the cup wall to tear; if the blankholder force is too low, wrinkling of the cup flange will occur.

Clearances are usually 7% to 14% greater than sheet thickness; if they are too small, the blank may be pierced or sheared by the punch. The corner radii of the punch and of the die are also important parameters. If they are too small, they can cause fracture of the cup at its corners; if they are too large, the cup wall may wrinkle (*puckering*).

Draw beads (Fig. 16.36) are often necessary to control the flow of the blank into the die cavity. They restrict the free flow of the sheet metal by bending and unbending it during the drawing cycle, thereby increasing the force required to push the sheet into the die cavity. Draw beads also help reduce the necessary blankholder force, because the beaded sheet has a higher stiffness (due to its higher moment of inertia) and, thus, lowering the tendency to wrinkle. Draw-bead diameters may range from 13 to 20 mm, the latter are applicable to large stampings, such as automotive panels.

Draw beads also are useful in drawing *box-shaped* and *nonsymmetric* parts (Fig. 16.36b and c). Note in Fig. 16.36c, for example, that various regions of the part being drawn undergo different types of deformation during drawing. Recall also the fundamental principle that the material flows in the direction of least resistance.

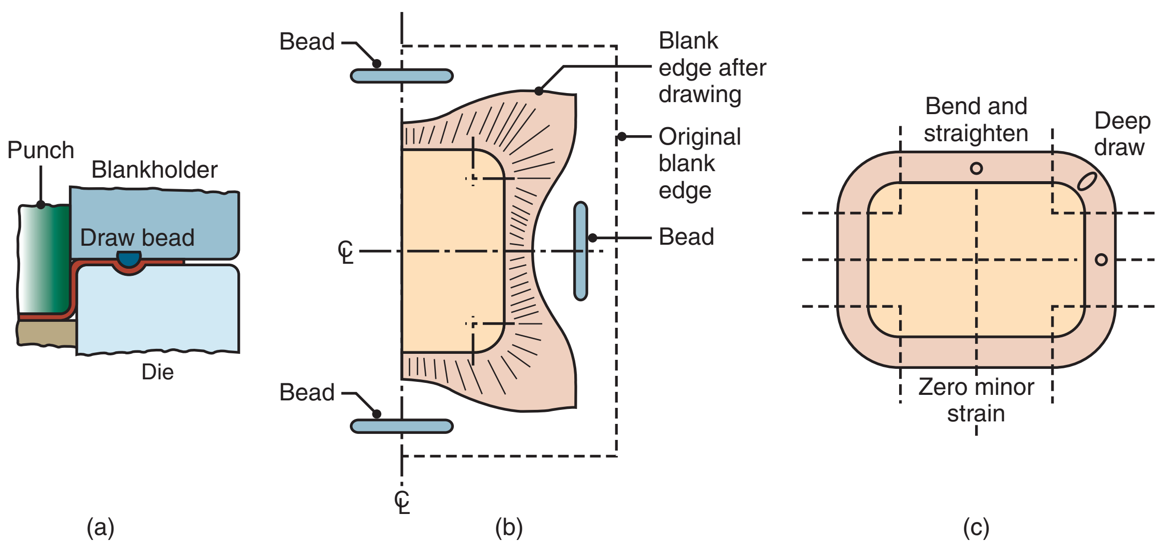


Figure 16.36: (a) Schematic illustration of a draw bead. (b) Metal flow during the drawing of a box-shaped part while using beads to control the movement of the material. (c) Deformation of circular grids in the flange in deep drawing.

In order to avoid tearing of sheet metal during forming, it often is necessary to incorporate the following:

- Proper design and location of draw beads
- Large die radii
- Effective lubrication
- Proper blank size and shape
- Cutting off the corners of square or rectangular blanks, at 45°, to reduce tensile stresses that develop during drawing
- Using blanks free of internal and external defects, including burrs.

Ironing. If the clearance between the punch and the die is sufficiently large, the drawn cup will have thicker walls at its rim than at its base (Fig. 16.32). The reason is that the cup rim consists of material from the outer diameter of the blank, hence it has undergone a larger diameter reduction; consequently, it becomes thicker, than the rest of the cup wall. As a result, the cup will have nonuniform wall thickness.

Wall thickness can be controlled by *ironing*, a process in which a drawn cup is pushed through one or more *ironing rings* (Fig. 16.31). The clearance between the punch and the ironing rings is less than the cup wall thickness, thus the drawn cup has an essentially constant wall thickness. Aluminum beverage cans, for example, are pushed through a set of two or three ironing rings, in one stroke and at very high speeds.

Redrawing. Containers that are difficult to draw in one operation generally undergo *redrawing* (Fig. 16.37). Because of volume constancy of the metal, the cup becomes longer as it is redrawn to a smaller cup diameter. In *reverse redrawing*, the cup is placed upside down in the die, and thus it undergoes bending in a direction opposite to its original configuration.

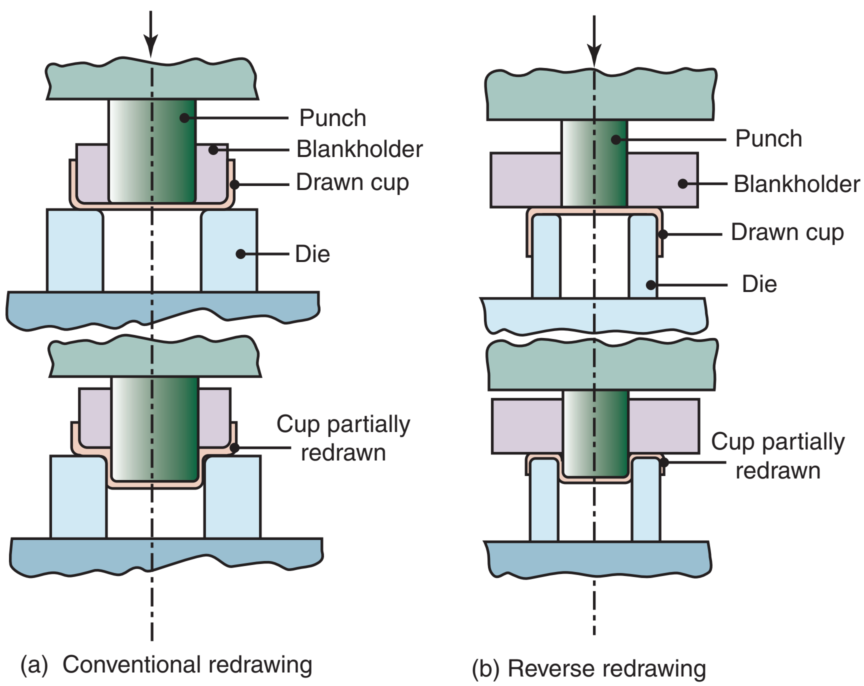


Figure 16.37: Reducing the diameter of drawn cups by redrawing operations: (a) conventional redrawing and (b) reverse redrawing. Small-diameter deep containers may undergo several redrawing operations.

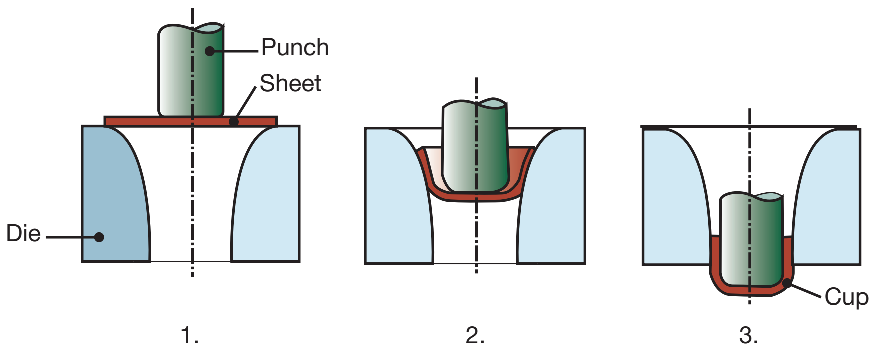


Figure 16.38: Stages in deep drawing without a blankholder, using a *tractrix* die profile.

Drawing without a Blankholder. Deep drawing also may be carried out without a blankholder. The dies are specially contoured for this operation to prevent wrinkling; one example is shown in Fig. 16.38. The sheet metal must be sufficiently thick to prevent wrinkling. The following formula is a general guide:

$$D_o - D_p < 5T, \tag{16.14}$$

where T is the sheet thickness. Thus, the thicker the sheet, the larger the blank diameter, and the deeper the cup, without wrinkling.

Embossing. This is an operation consisting of a shallow or moderate drawing, made with male and female matching shallow dies (Fig. 16.39). Embossing is widely used, principally for stiffening flat sheet-metal panels (thus increasing their moment of inertia) and for decorating, numbering, and lettering.

Tooling and Equipment for Drawing. The most common tool and die materials for deep drawing are tool steels, cast irons, and carbides (Table 5.7). Die-manufacturing methods are described in detail in Section 14.7. Because of the generally axisymmetric shape of the punch and die components, such as for making cylindrical cans and containers, they can be made on computer-controlled machine tools (Section 25.2).

The equipment for deep drawing is usually a *double-action hydraulic press* or a *mechanical press*, the latter being favored because of its higher operating speed. In a double-action hydraulic press, the punch and the blankholder are controlled independently. Punch speeds generally range between 0.1 and 0.3 m/s.

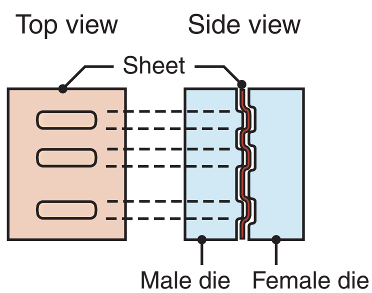


Figure 16.39: An embossing operation with two dies. Letters, numbers, and designs on sheet-metal parts can be produced by this process.

Case Study 16.2 Manufacture of Food and Beverage Cans

Can manufacturing is a major industry, with approximately 100 billion beverage cans and 30 billion food cans produced each year in the United States alone. These containers are strong and lightweight, typically weighing less than 14 g, and they are under an internal pressure of 620 kPa, reliably and without leakage of their contents. There are stringent requirements for their surface finish, since brightly decorated and shiny cans are preferred over dull-looking containers. Considering all of these features, metal cans are very inexpensive; can makers charge approximately $40 per 1000 cans, or about 4 cents per can.

Food and beverage cans may be produced in several styles, the most common being two-piece and three-piece cans. A two-piece can consist of the body and the lid (Fig. 16.40a). The body is made of one piece, drawn and ironed, thus the industry refers to this style as D&I (drawn and ironed) cans. Three-piece cans are produced by attaching a lid and a bottom to a sheet-metal cylindrical body, which is typically made by forming a seam on a sheet metal blank.

Drawn and ironed can bodies are produced from a number of aluminum alloys, but the most common is 3004-H19 (Section 6.2); electrolytic tin-plated ASTM A623 steel is also used for cans. Aluminum lids are made for both steel and aluminum cans, and are produced from 5182-H19 or 5182-H48 aluminum alloy. The lid has a demanding set of design requirements, as can be appreciated by reviewing Fig. 16.40b. Not only must the lid be *scored* easily (the curved grooves around the tab), but an integral rivet is formed and headed (Section 14.4) in the lid, to hold the tab in place.

Aluminum alloy 5182 has the unique characteristics of having sufficient formability to enable forming of the integral rivet without cracking, and also has the ability to be scored. The lids basically are stamped from 5182 aluminum sheet; the pop-top is scored, and a plastic seal is then placed around the periphery of the lid. The polymer layer seals the can's contents after the lid is seamed to the can body.

The traditional method of making the can bodies is shown in Fig. 16.31. The process starts with 140-mm diameter blanks, produced from rolled sheet stock. The blanks are (a) *deep drawn* to a diameter of about 90 mm; (b) *redrawn* to the final diameter of around 65 mm; (c) *ironed* through two or three ironing rings, in one pass; and (d) *domed*, for shaping the can bottom. The deep-drawing and ironing operations are performed in a special press, typically producing cans at speeds over 400 strokes per minute. Following this series of operations, a number of additional processes take place.

Necking of the can body is performed either by *spinning* (Section 16.9) or by *die necking*, which is a forming operation similar to that shown in Fig. 15.21a, where a thin-walled tubular part is pushed into the die, and then spin flanged. The reason for necking the can top is that the 5182 aluminum for the lid is relatively expensive; thus, by tapering the top of the can, a smaller volume of material is required. It should also be noted that the cost of a can often is calculated to millionths of one dollar, hence any design feature that reduces cost will be exploited by this competitive industry.

Source: Courtesy of J.E. Wang, Texas A&M University.

16.8 Rubber Forming and Hydroforming

The processes described in the preceding sections use dies that are made of solid materials, such as cast iron, steel, and carbides. In *rubber forming*, also known as the *Guerin process*, one of the die halves is made of a flexible material, typically a polyurethane membrane. Polyurethanes (Section 7.9) are used widely because of their abrasion resistance, fatigue life, and resistance to cutting or tearing.

In bending and embossing of sheet metal by this process, the female die is replaced with a rubber pad (Fig. 16.41). Note that the outer surface of the sheet is now protected from damage or scratches, because it is not in direct contact with a hard metal surface during forming. Pressures in rubber forming are typically on the order of 10 MPa.

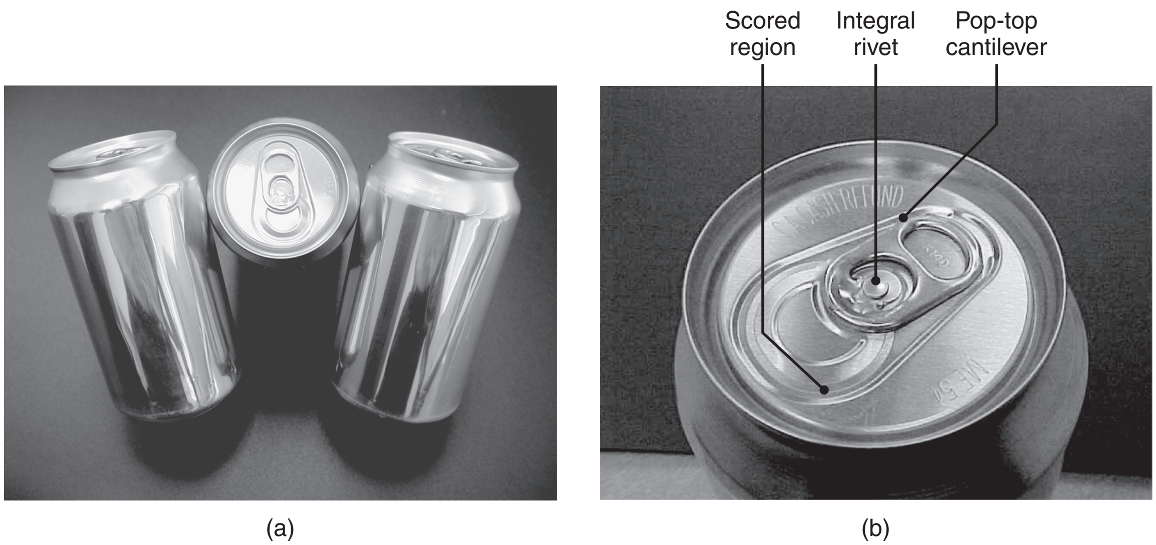


Figure 16.40: (a) Aluminum beverage cans; note the smooth surface. (b) Detail of the can lid, showing the integral rivet and scored periphery for the pop-top.

In the **hydroform** or *fluid-forming process* (Fig. 16.42), the pressure over the rubber membrane is controlled throughout the forming cycle, with a maximum pressure of up to 100 MPa. This method allows close control of the sheet during forming, and prevents its wrinkling or tearing. Deeper draws can be obtained as compared to conventional deep drawing, because the pressure around the rubber membrane forces the cup against the punch. As a result, the friction at the punch–cup interface increases, in turn reducing the longitudinal tensile stresses in the cup, thus delaying fracture.

Control of frictional conditions in rubber forming, as well as in other sheet-forming operations, can be a critical factor in making successful parts. Using proper lubricants and their method of application are also important.

In **tube hydroforming** (Fig. 16.43), metal tubing is shaped in a die, by pressurizing it internally by a fluid, usually water. This process can shape either simple tubes or various intricate hollow shapes (Fig. 16.43b). Parts made include automotive-exhaust and tubular structural components.

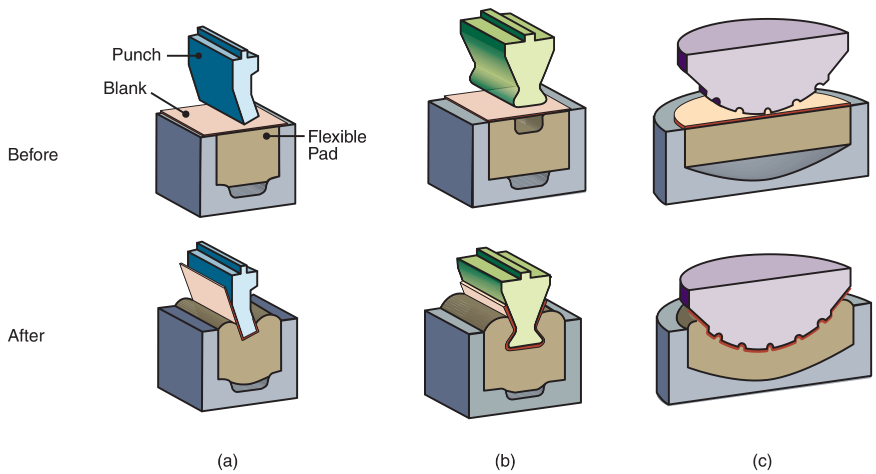


Figure 16.41: Examples of the bending and embossing of sheet metal with a metal punch and with a flexible pad serving as the female die.

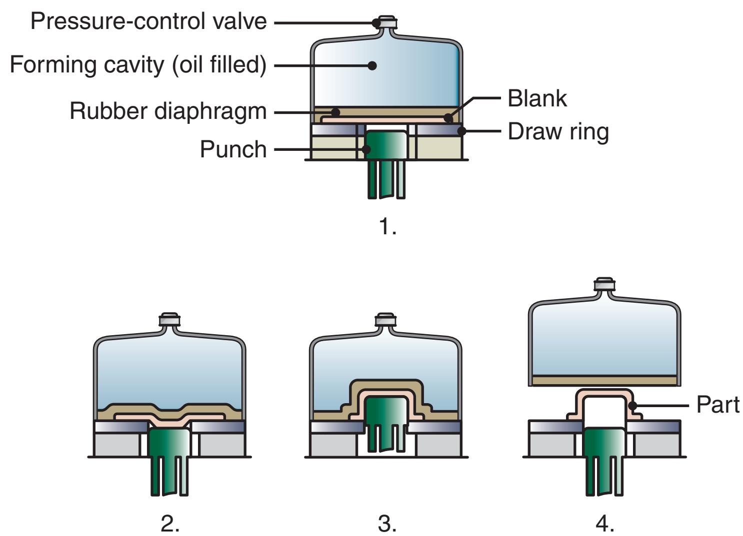


Figure 16.42: The hydroform (or fluid-forming) process. Note that, in contrast to the ordinary deep-drawing process, the pressure in the dome forces the cup walls against the punch. The cup travels with the punch; in this way, deep drawability is improved.

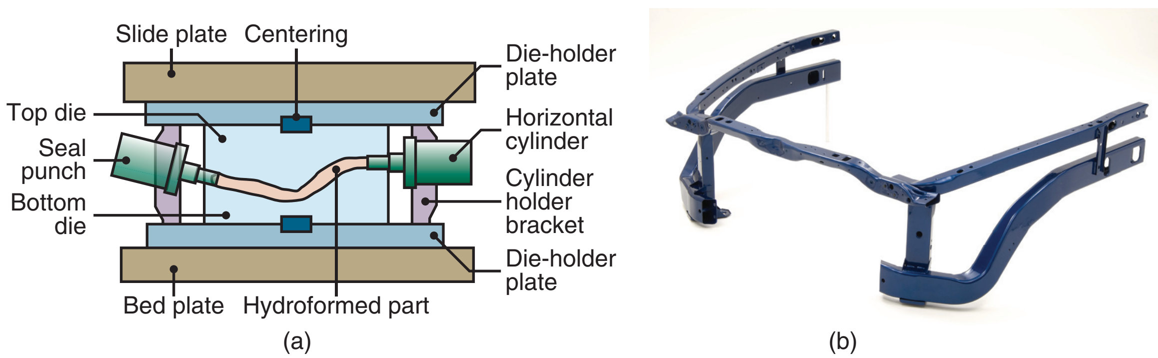


Figure 16.43: (a) Schematic illustration of the tube-hydroforming process. (b) A vehicle body assembly module made up of seven hydroformed tubes. *Source:* Courtesy Vari-Form, Inc.

When selected properly, rubber-forming and hydroforming processes have the advantages of (a) capability to form complex shapes, (b) forming parts with laminated sheets made of various materials and coatings, (c) flexibility and ease of operation, (d) avoiding damage to the surfaces of the sheet, (e) low die wear, and (f) low tooling cost.

Case Study 16.3 Tube Hydroforming of an Automotive Radiator Closure

The conventional assembly used to support an automotive radiator, or a radiator closure, is constructed through stamping of the components, then welding them together. To simplify the design and to achieve weight savings, a hydroformed assembly was designed, as shown in Fig. 16.44. Note that this design uses varying cross sections, an important feature to reduce weight and provide surfaces to facilitate assembly and mounting of the radiator.

A typical tube hydroforming processing sequence consists of the following steps:

1. Bending of tube to desired configuration
2. Tube hydroforming to achieve desired shape
3. Finishing operations, such as shearing of the ends and inspection
4. Assembly, including welding of components.

The operations performed on one of the tube components of the closure is shown in Fig. 16.45. The tube, constructed of steel, with a 300 MPa yield strength, is first bent to shape (Fig. 16.27). The bent tube is then placed in a hydroforming press and the end caps are attached.

Conventional hydroforming involves closing the die onto the tube, followed by internal pressurization to force the tube to the desired shape. Figure 16.46a shows a typical cross section. Note that as the tube is expanded, there is significant wall thinning, especially at the corners, because of friction at the tube–die interface. A sequence of pressures that optimize corner formation is thus followed, as shown in Fig. 16.46b.

In this approach, a first pressure stage (prepressure stage) is applied as the die is closing, causing the tube to partially fill the die cavity and shape the cross-section's corners. After the die is completely closed, the internal pressure is increased to lock-in the shape and provide the support needed for hole piercing. This sequence has the benefit of forming the sharp corners in the cross section by bending, as opposed to pure stretching as in conventional hydroforming. The final wall thickness is much more uniform, producing a more structurally sound component. Figure 16.47 shows a part being hydroformed.

The assembly shown in Fig. 16.44 has 76 holes that are pierced inside the hydroforming die; the ends are then sheared to length. The 10 components in the hydroformed closure are then assembled through robotic gas-metal arc welding (Section 30.4.3), using threaded fasteners to aid in the part's serviceability.

Compared to the original stamped design, the hydroformed design has four fewer components, uses only 20 welds as opposed to 174 for the stamped design, and weighs 10.5 kg versus 14.1 kg. Furthermore, the stiffness of the enclosure and of the water cooling areas are both significantly increased.

Source: Courtesy of B. Longhouse, Vari-Form, Inc.

Figure 16.44: Hydroformed automotive radiator closure, which serves as a mounting frame for the radiator. *Source:* Courtesy of B. Longhouse, Vari-Form, Inc.

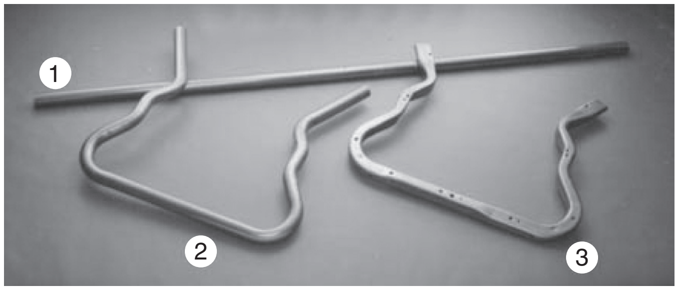


Figure 16.45: Sequence of operations in producing a tube-hydroformed component: (1) tube as cut to length; (2) after bending; (3) after hydroforming. *Source:* Courtesy of B. Longhouse, Vari-Form, Inc.

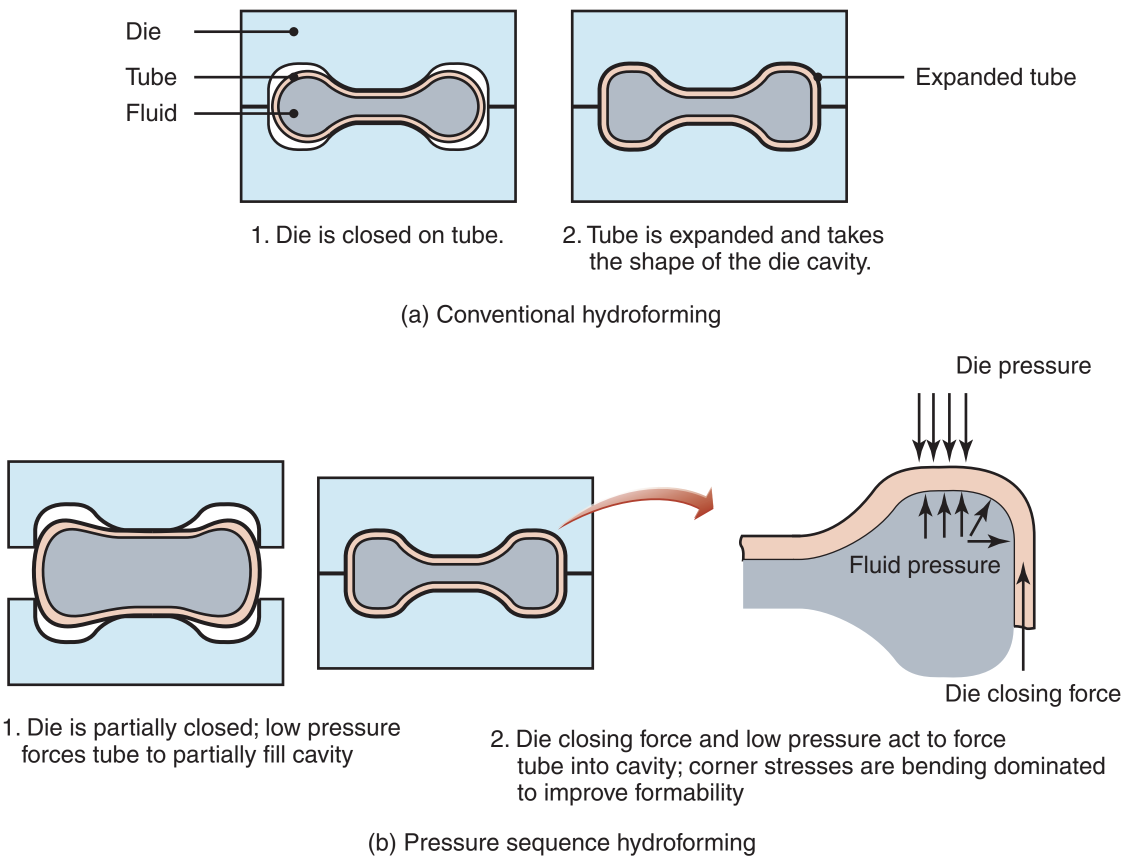


Figure 16.46: Schematic illustration of expansion of a tube to a desired cross section through (a) conventional hydroforming and (b) pressure sequence hydroforming.

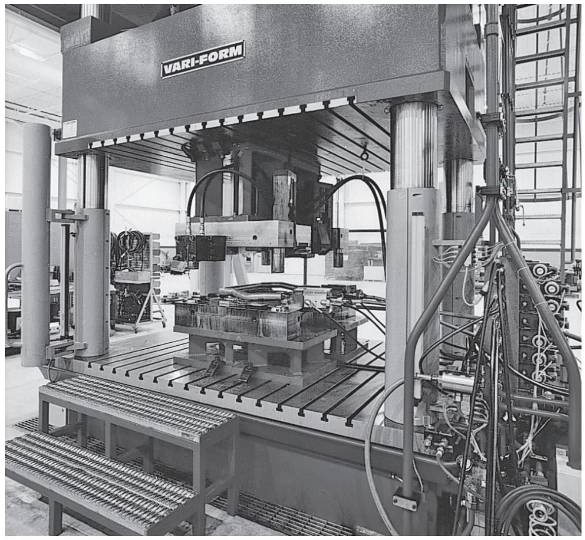


Figure 16.47: View of the tube-hydroforming press, with bent tube in place in the forming die. *Source:* Courtesy of B. Longhouse, Vari-Form, Inc.

16.9 Spinning

Spinning is a process that involves forming of axisymmetric parts over a mandrel, using a variety of tools and rolls. It is a process similar to that of shaping clay on a potter's wheel.

Conventional Spinning. In this process, a circular blank of flat or preformed sheet metal is placed and held against a mandrel, and rotated while a rigid tool shapes the material over the mandrel (Fig. 16.48a). The tool may be activated either manually or, for higher production rates, through computer numerical control. The process typically involves a sequence of passes, requiring considerable skill. Conventional spinning

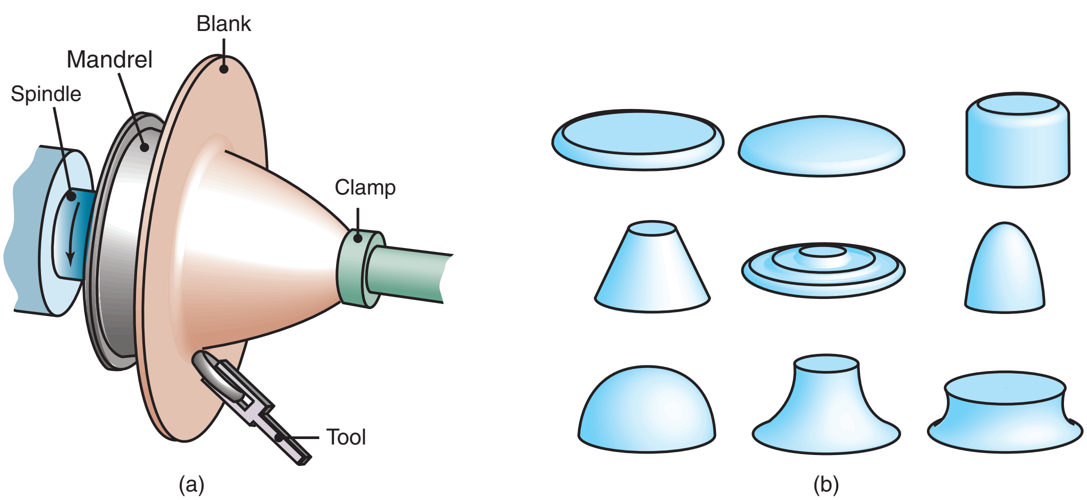


Figure 16.48: (a) Schematic illustration of the conventional spinning process. (b) Types of parts conventionally spun. All parts are axisymmetric.

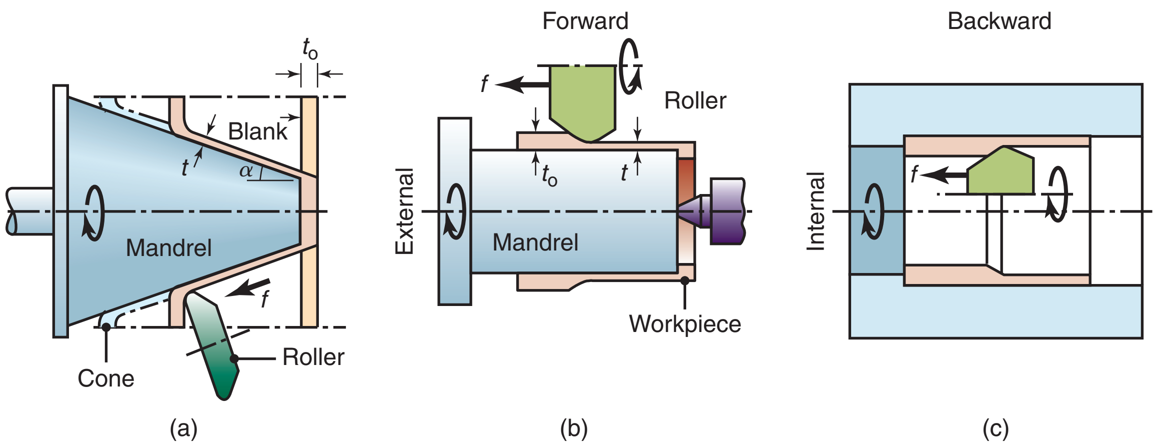


Figure 16.49: (a) Schematic illustration of the shear-spinning process for making conical parts. The mandrel can be shaped so that curvilinear parts can be spun. (b) and (c) Schematic illustrations of the tube-spinning process.

is particularly suitable for making conical and curvilinear shapes (Fig. 16.48b), with part diameters that can range up to 6 m, which otherwise would be difficult or uneconomical to produce. Although most spinning takes place at room temperature, thick parts and metals with high strength or low ductility require spinning at elevated temperatures.

Shear Spinning. Also called *power spinning*, *flow turning*, *hydrospinning*, and *spin forging*, this operation produces axisymmetric conical or curvilinear shapes, reducing the sheet's thickness while maintaining its blank diameter (Fig. 16.49a). A single forming roll can be used, but two rolls are preferable to balance the radial forces acting on the mandrel. Typical parts made are rocket motor casings and missile nose cones. Parts up to 3 m in diameter can be formed by shear spinning. Shear spinning produces little material waste, and it can be completed in a relatively short time, in some cases in as little as a few seconds. Various shapes can be spun with fairly simple tooling, which generally is made of tool steel.

The *spinnability* of a metal is generally defined as the maximum reduction in thickness to which a part can be subjected without fracture. Spinnability is found to be related to the tensile reduction of area of the material, just as is bendability (see Fig. 16.18). Thus, if a metal has a tensile reduction of area of 50% or higher, its thickness can be reduced by as much as 80% in one pass. For metals with low ductility, the operation is carried out at elevated temperatures, by heating the blank in a furnace and transferring it to the mandrel of the machine.

Tube Spinning. In this process, the thickness of hollow, cylindrical blanks is reduced or shaped by spinning them on a round mandrel, using rolls (Fig. 16.49). The parts may be spun *forward* or *backward*. This operation is capable of producing a variety of external and internal profiles, using cylindrical blanks with constant wall thickness. The maximum reduction in thickness per pass is related to the tensile reduction in area of the material, as in shear spinning. Tube spinning can be used to make axisymmetric parts, such as rocket, missile, and jet-engine parts, pressure vessels, and car and truck wheels.

Incremental Forming. *Incremental forming*, also called *incremental sheet forming* (ISF), is a term applied to a class of processes that are used to produce sheet-metal parts without dies. The simplest version is *incremental stretch expanding* (Fig. 16.50), wherein a blank is shaped by a rotating steel rod with a smooth hemispherical tip to produce axisymmetric parts. No special tooling or mandrel is required; the motion of the rod determines the final shape of the part, using one or more passes. Proper lubrication is essential.

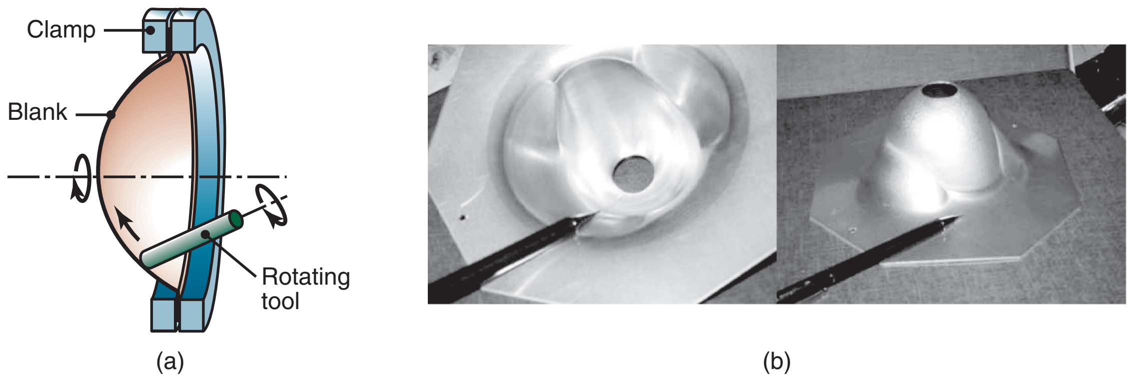


Figure 16.50: (a) Illustration of an incremental-forming operation. Note that no mandrel is used and that the final part shape depends on the path of the rotating tool. (b) An automotive headlight reflector produced through CNC incremental forming. Note that the part does not have to be axisymmetric. *Source:* (b) Courtesy of J. Jeswiet, Queen's University, Ontario.

CNC incremental forming uses a computer numerical control machine tool (see Section 37.3), programmed to follow contours at different depths across the sheet surface. The blank is clamped and is stationary, and the forming tool rotates. Tool paths are calculated in a manner similar to machining (Part IV), using a CAD model of the desired shape as the starting point (see Fig. 20.3). Figure 16.50b depicts an example of a part produced by this method; note that the part does not have to be axisymmetric.

The main advantages of CNC incremental forming are high flexibility in the shapes that can be produced and low tooling costs. This process has been used for rapid prototyping of sheet-metal parts (Chapter 20). The main drawbacks include low production rates and limitations on materials that can be shaped.

Case Study 16.4 Computer-aided Incremental Sheet Forming

A trend in modern automotive manufacture is the more frequent product changes and updates, as well as demand for customized parts and product personalization (see also Section 39.4). Accordingly, fast, high-quality, low-cost prototyping and low volume production processes are needed. Additive manufacturing approaches (Chapter 20) are useful for many parts, but sheet metal sections in the sizes needed are difficult or impossible to produce. For thin sheet metal products, incremental sheet forming (ISF) is a new, promising alternative to additive manufacturing (Fig. 16.50). For small batch volume production (500 units per year, see Table 37.2), it offers lower manufacturing cost compared to conventional forming processes and also faster design time.

During ISF, the sheet periphery is clamped in a blank holder. A generic tool, usually with a hemispherical end, is applied to locally deform the sheet with a predefined toolpath and the final desired geometry is achieved progressively. A typical spiral tool path is also shown in the figure. The die-less nature of SPIF combined with its universal tooling and high flexibility makes this approach highly favorable. A major advantage of ISF is that the process can be performed using standard CNC milling machines; for larger parts, robots have also been used to move the tool (Section 37.6).

The final material properties are close to those produced by conventional forming, because ISF avoids melting and other metallurgical issues associated with additive manufacturing; it also cold works the sheet metal. In addition, the complex stress state developed during ISF suppresses necking, which significantly enhances the material formability as compared to conventional forming processes such as stamping and hydroforming.

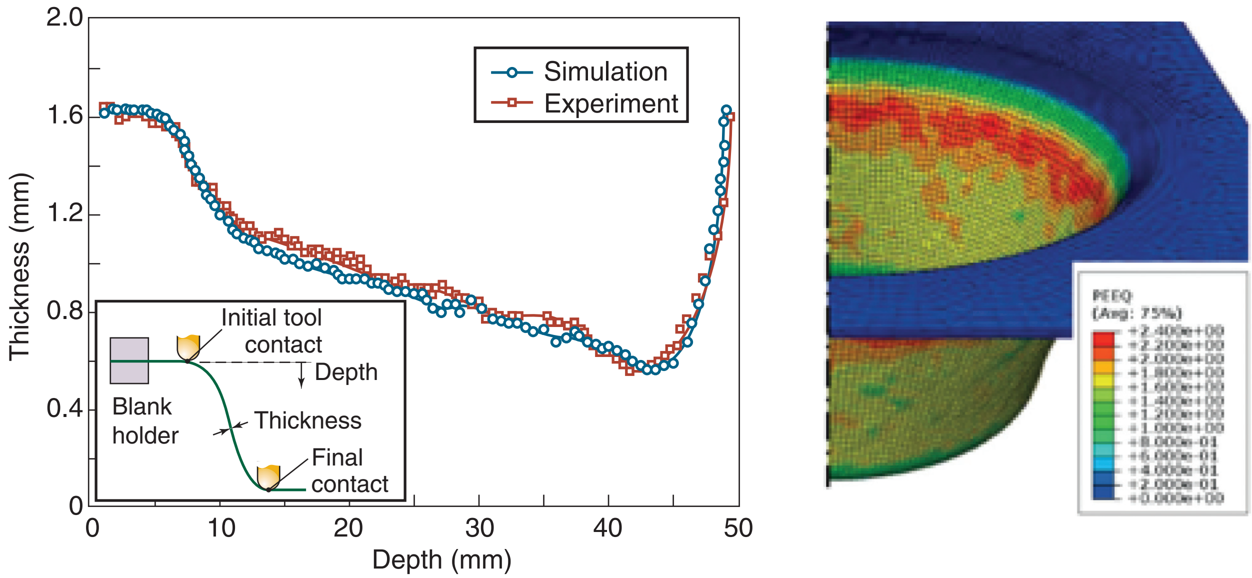
Rapid design cycles require associated computer aided engineering models capable of predicting the part geometry (Section 38.4), especially since ISF has a fairly high failure rate unless tool paths are carefully planned. There are now commercial software packages for stamping operations that can predict final part geometry and thereby allow an engineer to accurately design the required stamping dies. Similar tools are under development for ISF to design the appropriate tool path. Modern models can predict part thickness (Fig. 16.51); current research efforts are underway to incorporate springback.

Ford Motor Company has reported using ISF for producing body panels for prototype vehicle production. This technology offers the opportunity for evaluating design changes quickly, which can speed up the product development process. ISF is also being explored for aircraft frame construction where the annual volumes are consistent with the capability of this process. Boeing is leading a project being executed by the Manufacturing USA Lightweight Innovations for Tomorrow (LIFT) Institute to evaluate the potential of ISF. The target component, an airframe fuel cover, was shown to be economically competitive with the baseline hydroforming process up to volumes of around 500 units per year. Further advances are being made to enable faster tool speeds and larger spiral step sizes while still delivering the desired part geometry and mechanical properties, offering the potential for cost advantage at even higher production volumes.

Source: Courtesy of A. Taub and M. Banu, University of Michigan.

16.10 Superplastic Forming

The superplastic behavior of some metals (Section 2.2.7) within certain temperature ranges involves tensile elongations on the order of up to 2000%. Examples of such materials are zinc–aluminum and titanium alloys, with very fine grains, typically less than 10 to 15 μm (see Table 1.1). Superplastic alloys can be formed



(a) (b)

Figure 16.51: Finite element model of single point incremental forming. (a) Validation of the thickness profile for a funnel shape; (b) Strain distribution of a 67° cone.

into complex shapes by *superplastic forming*, a process that employs common metalworking techniques, as well as by polymer-processing techniques, such as thermoforming, vacuum forming, and blow molding (Chapter 19). The behavior of the material in superplastic forming is similar to that of bubble gum or hot glass, which, when blown, expands several times its original diameter before it bursts.

Superplastic alloys, particularly Zn-22Al and Ti-6Al-4V, can also be formed by bulk-deformation processes, including closed-die forging, coining, hubbing, and extrusion (Chapters 14 and 15). Common die materials in superplastic forming are low-alloy steels, cast tool steels, ceramics, graphite, and plaster of paris. Their selection depends on the forming temperature and the strength of the superplastic alloy.

The very high ductility and relatively low strength of superplastic alloys offer the following advantages:

- Complex shapes can be formed from one piece, with fine detail, close tolerances, and elimination of secondary operations
- Weight and material savings can be significant, because of the high formability of the materials
- Little or no residual stresses are present in the formed parts
- Because of the low strength of the material at forming temperatures, tooling can be made of materials that have lower strength than those in other metalworking processes, thus tooling costs are lower.

On the other hand, superplastic forming has the following limitations:

- The material must not be superplastic at service temperatures, as otherwise the part will undergo shape changes during its use
- Because of the high strain-rate sensitivity of the superplastic material (Section 2.2.7), it must be formed at sufficiently low strain rates, typically 10^{-4} to 10^{-2}/s. Forming times range anywhere from a few seconds to several hours; cycle times are thus much longer than those of conventional forming processes.

Diffusion Bonding/Superplastic Forming. Fabricating complex sheet-metal structures by combining *diffusion bonding* with *superplastic forming* (DB/SPF) is an important manufacturing strategy, particularly in the aerospace industry. Typical structures produced are shown in Fig. 16.52, in which flat sheets are first diffusion bonded (Section 31.7) and then formed. In this process, selected locations of the sheets are first diffusion bonded while the rest of the interfaces remains unbonded, using a layer of material (*stop-off*) to prevent bonding. The structure is then expanded in a mold, thus taking the shape of the mold, typically by using pressurized neutral (argon) gas. These structures have high stiffness-to-weight ratios, because they are thin and, by design, have high section moduli; an important feature that makes this process particularly attractive in aerospace and aircraft applications.

The DB/SPFS process improves productivity by eliminating mechanical fasteners and produces parts with good dimensional accuracy and low residual stresses. The technology is well advanced for titanium structures for aerospace applications. In addition to various aluminum alloys being developed using this technique, other metals for superplastic forming include various nickel alloys.

16.11 Hot Stamping

Increasing fuel economy in automobiles has received considerable attention in recent years for both environmental and economic reasons. To achieve fuel economy without compromising performance or safety, manufacturers have increasingly applied advanced materials in automobiles. Die-cast magnesium or extruded aluminum components are examples, but these materials are not sufficiently stiff or as well suited as steel for occupant safety. Thus, there has been a recent trend to consider hot stamping of advanced high-strength steels.

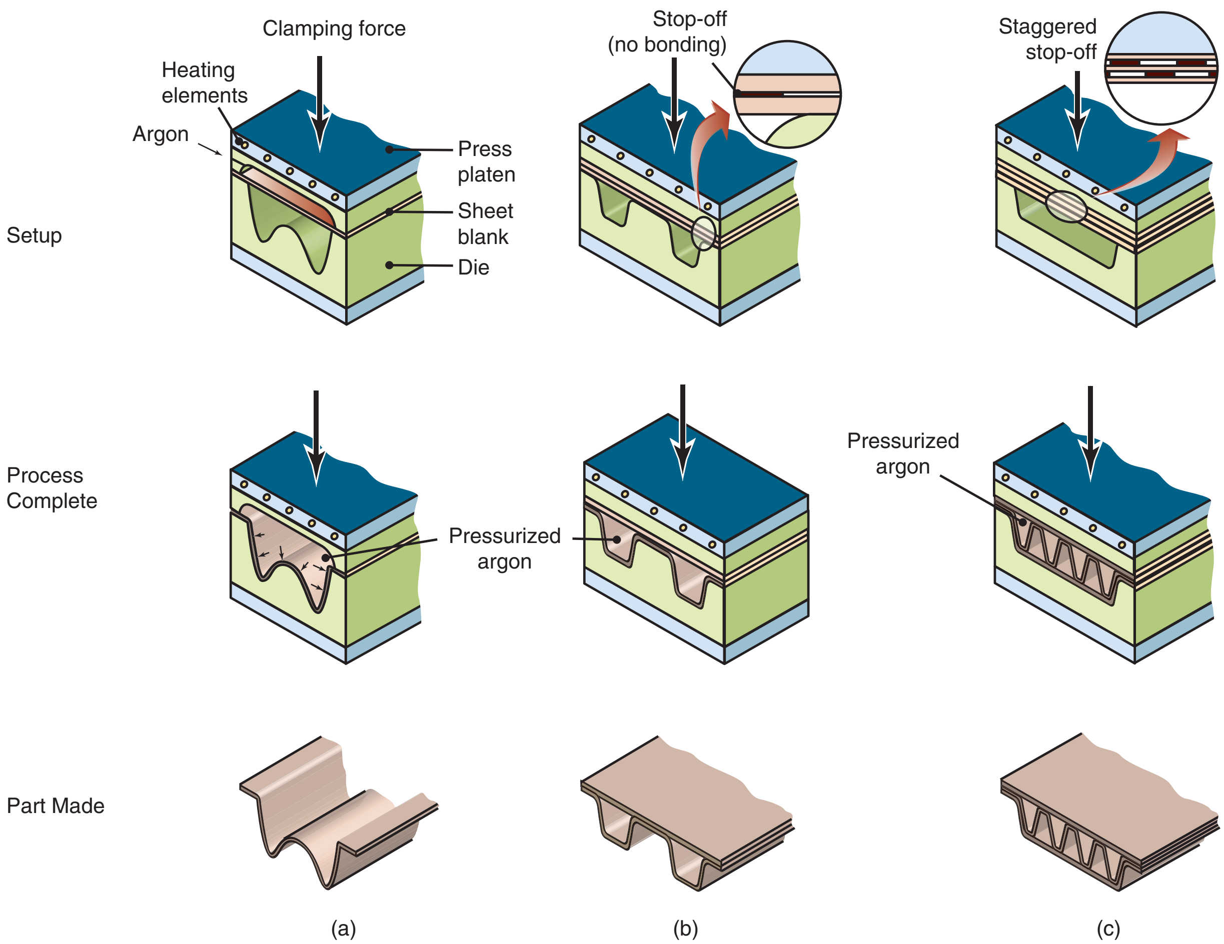


Figure 16.52: Production of structures by superplastic forming and diffusion bonding of sheet metals. Such structures have high stiffness-to-weight ratio. The process involves heating the sheet, followed by pressurization of the working gas, usually argon. Cycle times are relatively long, allowing both superplastic forming and diffusion bonding in the same die. (a) Single sheet superplastic forming. (b) Two sheet superplastic forming with diffusion bonding. A stop-off, commonly yttrium for titanium structures, is placed where bonding is not desired, and the pressurized argon is introduced between the sheets to form them. (c) Three sheet superplastic forming with diffusion bonding. The inner and outer sheets conform to the die profile, while the middle sheet produces a structure with support struts.

As described in Section 5.5.5, high-strength TRIP and TWIP steels are now available, with yield strengths and ultimate strengths exceeding 1300 MPa and 2000 MPa, respectively (Table 5.4). Conventional forming of these materials would be difficult or impossible, because of the high forces required and the excessive springback after forming. For these reasons, the sheet metal is preheated to above 900°C (usually 1000°-1200°C) and hot stamped. To extend die life and to quench the material within the die, as described below, the tooling is maintained at a much lower temperature, typically 400°–500°C.

Hot stamping allows exploitation of steel phases to facilitate forming and maximize part strength. Basically, the steel is maintained at elevated temperatures to form austenite (see Section 4.4), which has a ductile fcc structure at elevated temperatures. When shaped and brought into contact with the much cooler tooling, the steel is rapidly quenched, forming martensite, which is a very hard and strong but brittle form of steel (Section 4.7).

A typical hot-stamping sequence involves the following steps:

1. The material is heated up to the austenization temperature, and allowed to *dwell* or *soak* for a sufficiently long time to ensure that quenching will be quick when it contacts the die, but not before. Three basic methods are used to heat blanks prior to stamping: roller hearth furnaces, induction heating, and resistive heating. The last two methods have the advantage of shorter soak times, but may not lead to uniform temperature distributions throughout the part. The soak time must be optimized in order to ensure proper quenching while minimizing the cycle time.

2. In order to avoid cooling of the part before shaping it, the blank must be transferred to the dies as quickly as possible. Forming must be performed quickly, before the beginning of transformation of austenite into martensite.

3. Once the part is formed, the dies remain closed while the part is quenched, which takes from 2 to 10 seconds, depending on sheet thickness, temperature of sheet and die, and workpiece material. The cooling rate must be higher than 27°C/s to develop martensite. Forming is done with steel tools that have cooling channels incorporated in them, in order to maintain proper tooling temperature. A complete transformation into martensite results in the high strengths given in Table 5.4. It should be noted that quenching from austenite to martensite results in an increase in volume, thus influencing the residual stress distribution and workpiece distortion in forming.

Pressurized hot gas (air or nitrogen) can also be used as a working media to form the material, similar to hydroforming (Section 16.8). This method improves formability and, with proper process control, allows for more uniform blank and tooling temperatures, and thus lower residual stresses and warping.

Because the workpiece is hot and quenching must be done very rapidly, hot stamping is usually performed without a lubricant; also, shot blasting (Section 26.8) is often required after forming to remove scale from part surfaces. The steel may also be later coated with an aluminum-silicon layer, to prevent oxidation and eliminate the grit blasting step. In such a case, the coating requires a slightly longer soak time, in order to properly bond to the steel substrate.

Hot stamping is not restricted to steels; magnesium alloys ZEK100, AZ31, and ZE10 are also attractive because of their light weight, but these materials have limited formability at room temperature. Therefore, they are stamped at temperatures up to 300°C. Also, some advanced aluminum-alloy sheets are formed at elevated temperatures for improved ductility, and even develop superplastic behavior.

Electrically Assisted Forming. A recent technology has involved the application of high current through a metallic workpiece during the forming operation, known as *electrically assisted forming* (EAF). This procedure has also been applied to bulk forming operations, and is believed to increase the formability of materials. The mechanisms involved may be associated with increased temperature, although it has been suggested that the heating is higher in the vicinity of dislocations (Section 1.4.1), and that the malleability associated with electron mobility is increased.

16.12 Specialized Forming Processes

Although not as commonly used as the other processes described thus far, especially for high-rate and high-volume production, several other sheet-forming processes are used for specialized applications.

Explosive Forming. By controlling their quantity and shape, explosives also are a source of energy for sheet-metal forming. In *explosive forming*, first utilized to shape metals in the early 1900s, the sheet blank is clamped over a die, and then the entire assembly is lowered into a tank filled with water (Fig. 16.53a). The air in the die cavity is evacuated, an explosive charge is placed at a certain height, and the charge is then detonated.

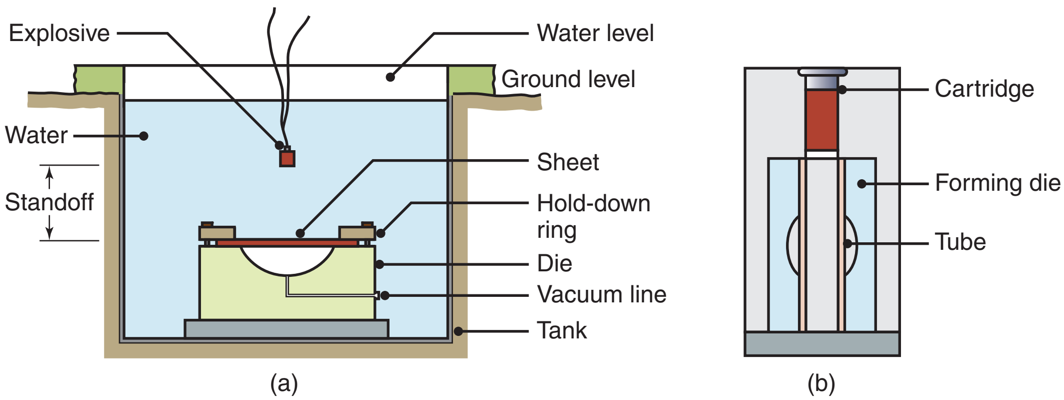


Figure 16.53: (a) Schematic illustration of the explosive-forming process. (b) Illustration of the confined method of the explosive bulging of tubes.

The explosive generates a shock wave, developing a pressure that is sufficient to form the blank. The *peak pressure*, p, generated in water is given by the expression

$$p = K\left(\frac{\sqrt[3]{W}}{R}\right)^{a}, \tag{16.15}$$

where p is in MPa, K is a constant that depends on the type of explosive, such as 21,600 for TNT (trinitrotoluene), W is the weight of the explosive in pounds, R is the distance of the explosive from the sheet-metal surface (called the *standoff*), in feet, and a is a constant, generally taken as 1.15.

A variety of shapes can be formed by explosive forming, provided that the material is sufficiently ductile at the very high rates of deformation that is characteristic of this process (see Table 2.4). The process is versatile, as there is virtually no limit to the size of the sheet or the plate that can be formed. It is suitable particularly for low-quantity production runs of large parts, such as those used in aerospace applications. Steel plates 25-mm thick and 3.6 m in diameter have been formed by this method, as have tubes, with wall thicknesses as much as 25 mm.

The explosive-forming method also can be used at a much smaller scale, as shown in Fig. 16.53b. In this case, a *cartridge* (canned explosive) is used as the source of energy. The process can be useful in bulging and expanding of thin-walled tubes for specialized applications.

The mechanical properties of parts made by explosive forming have been found to be similar to those made by conventional forming techniques. Depending on the number of parts to be produced, dies for this method can be made of aluminum alloys, steel, ductile iron, zinc alloys, reinforced concrete, wood, plastics, or composite materials.

Electromagnetically Assisted Forming. In *electromagnetically assisted forming*, also called *magnetic-pulse forming*, the energy stored in a capacitor bank is discharged rapidly through a magnetic coil. In a typical example, a ring-shaped coil is placed over a tubular workpiece; the workpiece is then collapsed by magnetic forces over the inner part (mandrel), thus making an integral assembly (Fig. 16.54).

The mechanics of this process is based on the fact that a magnetic field produced by the coil (Fig. 16.54a), crosses the metal tube (an electrical conductor) and generates *eddy currents* in the tube. These currents in turn, produce their own magnetic field. The forces produced by the two magnetic fields oppose each other, repelling the coil and the tube from each other, and collapsing the tube over the inner piece. The higher the electrical conductivity of the workpiece, the higher are the magnetic forces. Note that it is not necessary for the workpiece material to have magnetic properties, but it must be electrically conducting.

Under very high strain rates, the formability of the material is increased, dimensional accuracy is improved, and springback and wrinkling are reduced. The design of the magnetic coil is an important factor

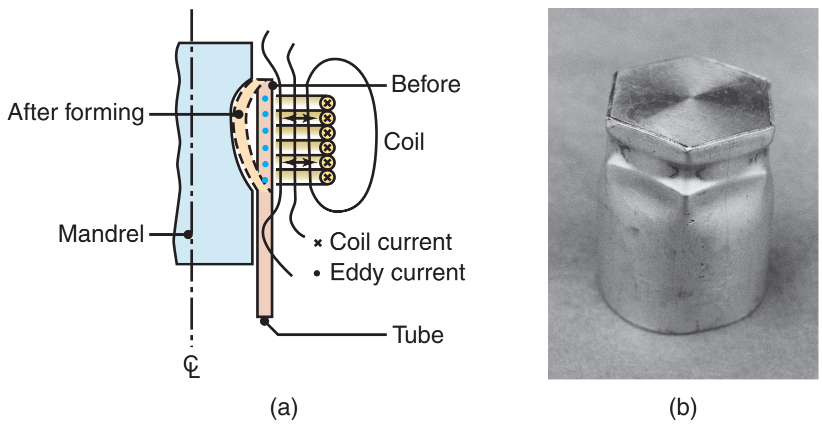


Figure 16.54: (a) Schematic illustration of the magnetic-pulse-forming process used to form a tube over a plug. (b) Aluminum tube collapsed over a hexagonal plug by the magnetic-pulse-forming process.

in the success of this operation. Flat magnetic coils can be made for use in such operations as embossing and shallow drawing of sheet metals.

First used in the 1960s, electromagnetically assisted forming has been demonstrated to be particularly effective for aluminum alloys. Electromagnetically assisted forming has been applied to (a) collapsing thin-walled tubes over rods, cables, and plugs; (b) compression crimp sealing of automotive oil filter canisters; (c) specialized sheet-forming operations; (d) bulging and flaring operations; and (e) swaging end fittings onto torque tubes for the Boeing 777 aircraft.

Peen Forming. As shown in Fig. 16.55, peen forming is used to produce curvatures on thin sheet metals by shot peening (Section 34.2) one surface of the sheet, subjecting the surface of the sheet to compressive stresses, which tend to laterally expand the surface layer (see also Section 2.11). Because the material below the peened surface remains rigid, the surface expansion causes the sheet to develop a curvature. Compressive surface residual stresses are also induced, improving fatigue strength. The shots are made of cast-iron or steel, and are discharged either from a rotating wheel or by an air blast from a nozzle. Peen forming is used by the aircraft industry to generate smooth and complex curvatures on aircraft wing skins. Cast-steel

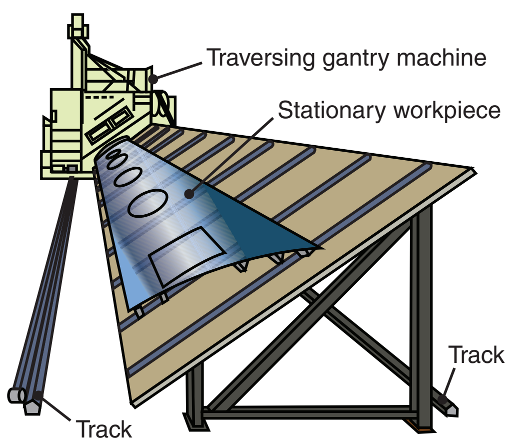


Figure 16.55: Schematic illustration of a peen-forming machine to shape a large sheet-metal part, such as an aircraft-skin panel. Note that the sheet is stationary and the peening head travels along its length. *Source:* Metal Improvement Company.

shot about 2.5 mm in diameter, traveling at speeds of 60 m/s, have been used to form wing panels 25 m long. For heavy structural sections, shot diameters as large as 6 mm may be used. The peen-forming process also is used for *straightening* twisted or bent parts, including out-of-round rings to make them round.

Laser Beam Forming. This process involves the application of laser beams as a localized heat source over specific regions of sheet metals. The steep thermal gradients developed through the thickness of the sheet produce thermal stresses, which are sufficiently high to cause localized plastic deformation of the sheet. With this method, a sheet can be bent permanently without using any dies.

In **laser-assisted forming**, the laser acts as a localized heat source, reducing the strength of the sheet metal at specific locations, thus improving formability and increasing process flexibility. Applications include straightening, bending, embossing, and forming of complex tubular or flat components.

Microforming. Microfilming is a more recent development and includes a family of processes that are used to produce very small metallic parts and components. Examples of *miniaturized products* include a wristwatch with an integrated digital camera and a multiple-gigabyte computer storage components. Typical parts made by microforming include springs, screws, small shafts for micromotors, and a variety of cold-headed, extruded, bent, embossed, coined, punched, or deep-drawn parts. Dimensions are typically in the submillimeter range, and part weights are on the order of milligrams.

Electrohydraulic Forming. Also called *underwater spark* or *electric-discharge forming*, the source of energy is a spark between two electrodes connected to each other with a short, thin wire. The rapid discharge of the energy from a capacitor bank through the wire generates a shock wave in the water, similar to those developed in explosive forming. The pressure in the water medium is sufficiently high to form the part. The energy levels are lower than those in explosive forming, being typically a few kJ. Electrohydraulic forming is a batch process and can be used to produce various small parts.

Case Study 16.5 Cymbal Manufacture

Cymbals (Fig. 16.56a) are an essential percussion instrument for all forms of music. Modern drum-set cymbals cover a wide variety of sounds, from deep, dark, and warm to bright, high-pitched, and cutting. Some cymbals sound musical, while others are trashy. A wide variety of sizes, shapes, weights, hammerings, and surface finishes (Fig. 16.56b) is available to achieve the desired performance.

Cymbals are produced from such metals as B20 bronze (80% Cu–20% Sn with a trace of silver), B8 bronze (92% Cu–8% Sn), nickel–silver alloy, and brass (see also Section 6.4). The manufacturing sequence for producing a bronze cymbal is shown in Fig. 16.57. The B20 metal is first cast into mushroom-shaped ingots. The ingot is then rolled successively up to 14 times, with water cooling the metal with each pass through the rolling mill. Special care is taken to roll the bronze at a different angle with each pass, to minimize anisotropy and develop an even, round shape.

The as-rolled blanks are then reheated and stretch formed (pressed) into the cup or bell shape, which determines the cymbal's overtones. The cymbals are then center drilled or punched, to make hang holes, and trimmed on a rotary shear to approximate final diameters. This operation is followed by another stretch-forming step, to achieve the characteristic Turkish dish form that controls the cymbal's pitch.

Automatic peen forming is done on machinery (Fig. 16.58a) and without using templates, since the cymbals have already been pressed into shape, but the peening pattern is controllable and uniform. The size and pattern of the peening operations depend on the desired response, such as tone, sound, response, and pitch of the cymbal. The cymbals are then hammered to impart a distinctive character to each instrument. Hammering can be done by hand, which involves placing the bronze blank on a steel anvil, where the cymbals then are struck manually by hand hammers.

Several finishing operations are performed on the cymbals, which can involve merely cleaning and printing of identifying information, as some musicians prefer the natural surface appearance and sound of shaped, hot-rolled bronze. More commonly, the cymbals are turned on a lathe, and without using any machining fluid in order to remove the oxide surface and reduce the thickness of the cymbal for the desired weight and sound. As a result, the surface finish becomes lustrous and, in some cases, also develops a favorable microstructure.

Some cymbals are polished to a glossy brilliant finish. In many cases, the surface indentations from peening persist after finishing; this is recognized as an essential performance feature of a cymbal, and it is also an aesthetic feature appreciated by musicians. Various surface finishes associated with modern cymbals are shown in Fig. 16.56b.

Source: Courtesy of W. Blanchard, Sabian Ltd.

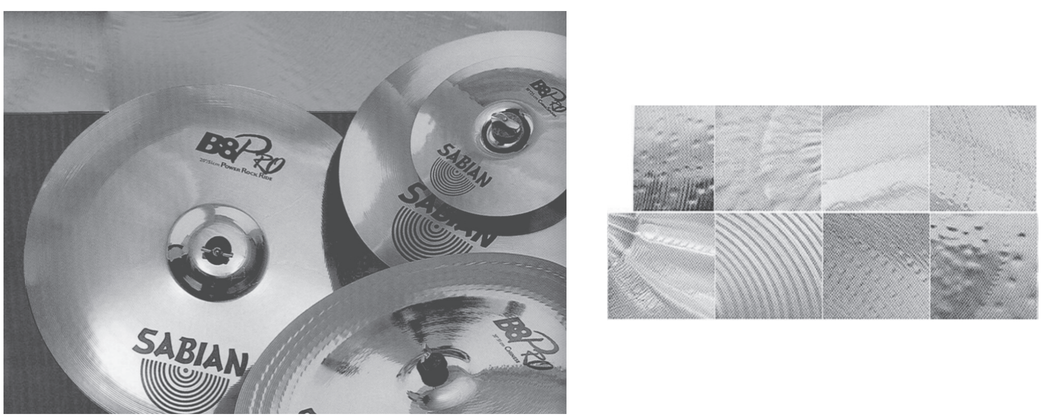
16.13 Manufacturing of Metal Honeycomb Structures

A *honeycomb structure* basically consists of a core of honeycomb, or other corrugated shapes, bonded to two thin outer skins (Fig. 16.59). The most common example of such a structure is *corrugated cardboard*, which has a high stiffness-to-weight ratio and is used extensively in packaging for shipping consumer and industrial goods. Because of their light weight and high resistance to bending, metal honeycomb structures are used for aircraft and aerospace components, buildings, and transportation equipment. The chassis of the Koenigsegg, a Swedish sports car, for example, is made partly of aluminum honeycomb with an integrated fuel tank. Honeycomb structures also may be made of nonmetallic materials, such as polymers and a variety of composite materials.

Honeycomb structures are made most commonly of 3000-series aluminum, but for specialized applications and corrosion resistance, they may also be made of titanium, stainless steels, and nickel alloys. Reinforced plastics, such as aramid-epoxy, also are used to make these structures.

There are two basic methods of manufacturing honeycomb materials. In the **expansion process**, which is the more common (Fig. 16.59a), sheets are first cut from a coil, and an *adhesive* (Section 32.4) is applied at intervals (node lines) on their surfaces. The sheets are then stacked and cured in an oven, developing strong bonds at their adhesive surfaces. The block is then sliced to required dimensions and stretched to develop a honeycomb structure.

In the **corrugation process** (Fig. 16.59b) the sheet first passes through a pair of specially designed rolls, thus making a corrugated sheet; it is then cut into specific lengths. An adhesive is then applied to the



(a) (b)

Figure 16.56: (a) Selected common cymbals. (b) Detailed view of different surface textures and finishes of cymbals. *Source:* Courtesy of W. Blanchard, Sabian Ltd.

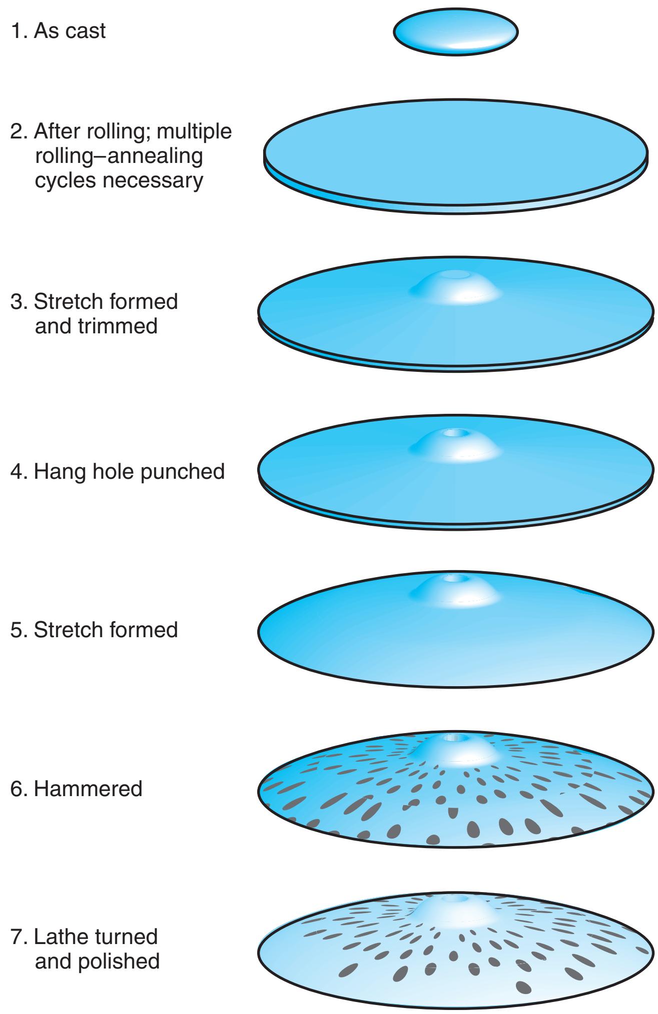


Figure 16.57: Manufacturing sequence for the production of cymbals. *Source:* Courtesy of W. Blanchard, Sabian Ltd.

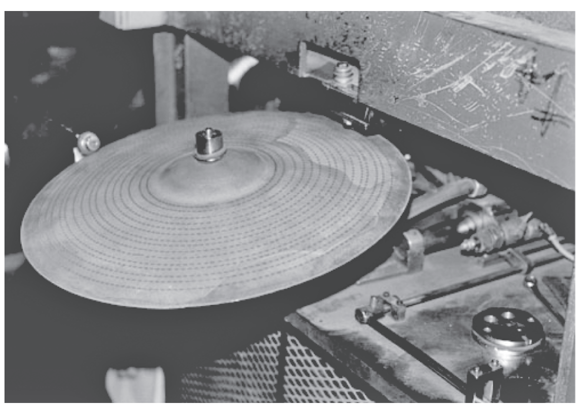


Figure 16.58: Automated hammering of a cymbal on a peening machine. *Source:* Courtesy of W. Blanchard, Sabian Ltd.

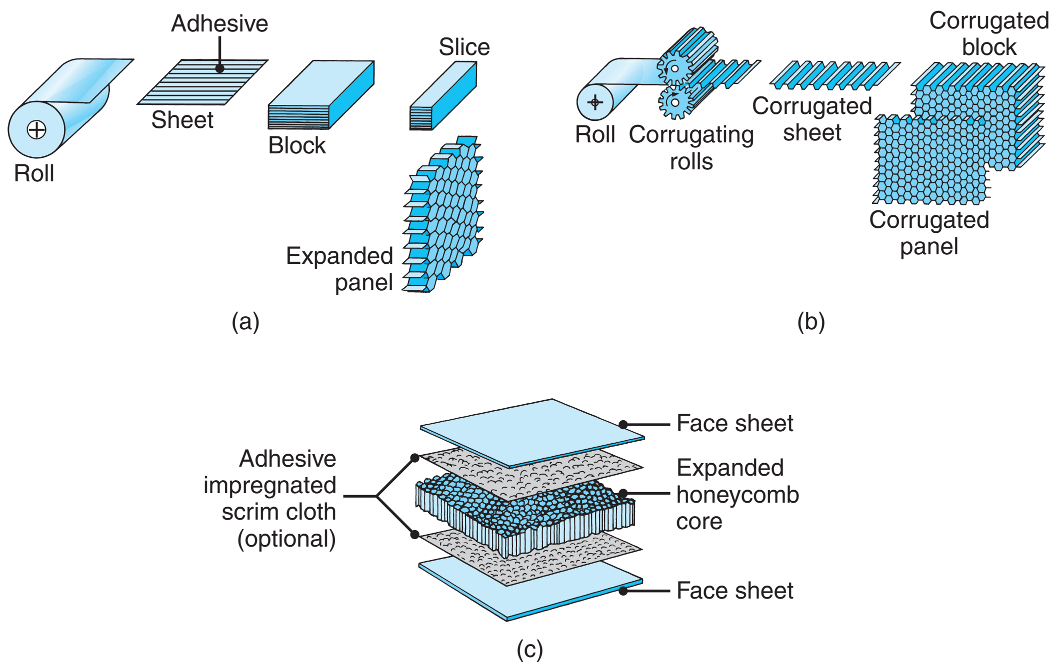


Figure 16.59: Methods of manufacturing honeycomb structures: (a) expansion process; (b) corrugation process; and (c) assembling a honeycomb structure into a laminate.

node lines; the corrugated sheets are then stacked into a block, and the block is cured. Because the sheets are already preformed, no expansion stage is involved. The honeycomb is finally made into a sandwich structure (Fig. 16.59c), using face sheets that are joined by adhesives (or they are *brazed*; Section 32.2) to the top and bottom surfaces.

16.14 Design Considerations in Sheet-metal Forming

The following design guidelines apply to sheet-metal forming operations, identifying the most significant design considerations.

Blank Design. Material scrap is the primary concern in blanking operations (see also Table 40.4). Poorly designed parts will not *nest* properly, and there can be considerable scrap produced (Fig. 16.60).

Bending. The main concerns in bending operations are fracture, wrinkling, and inability to properly make the bend. As shown in Fig. 16.61, a sheet-metal part with a flange will force the flange to undergo compression, which may cause buckling (see also *flanging*, Section 16.6). Buckling can be controlled with a relief notch, cut to limit the stresses developed in bending, or else a design modification as shown in Fig. 16.62 can be made. Right-angle bends have similar difficulties; relief notches can be used to avoid tearing as shown in the figure.

Because the bend radius is a highly stressed area, all stress concentrations should be removed from the bend-radius location, such as holes near bends. It is advantageous to move the hole away from the bend area, but when this is not possible, a crescent slot or ear can be used (Fig. 16.63a). Similarly, in bending flanges, tabs and notches should be avoided, because their stress concentrations will greatly reduce formability. When tabs are necessary, large radii should be specified to reduce stress concentration (Fig. 16.63b).

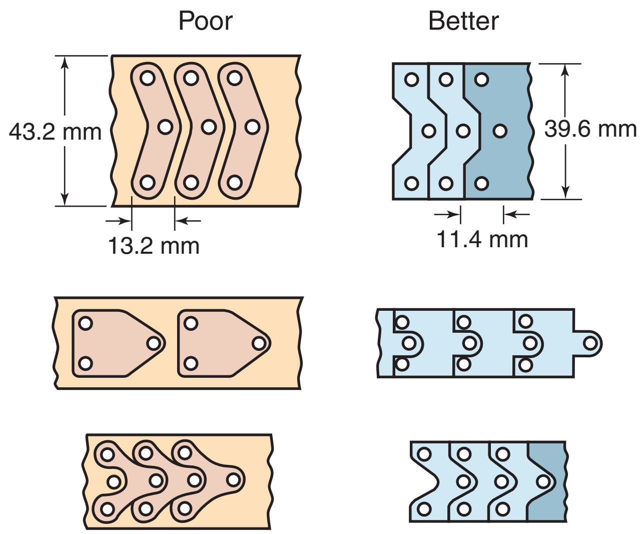


Figure 16.60: Efficient nesting of parts for optimum material utilization in blanking. *Source:* Courtesy of Society of Manufacturing Engineers.

If notches are necessary, it is important to orient them properly with respect to the grain direction of the sheet metal. As shown in Fig. 16.17, bends ideally should be perpendicular to the rolling direction of the sheet, or oblique if this is not possible, in order to avoid cracking. Bending with very small radii can be accomplished by scoring or embossing (Fig. 16.64), but this operation can cause fracture. Burrs should not be present in a bend allowance (see Fig. 16.2), because they are less ductile and can lead to crack initiation and propagation into the rest of the sheet.

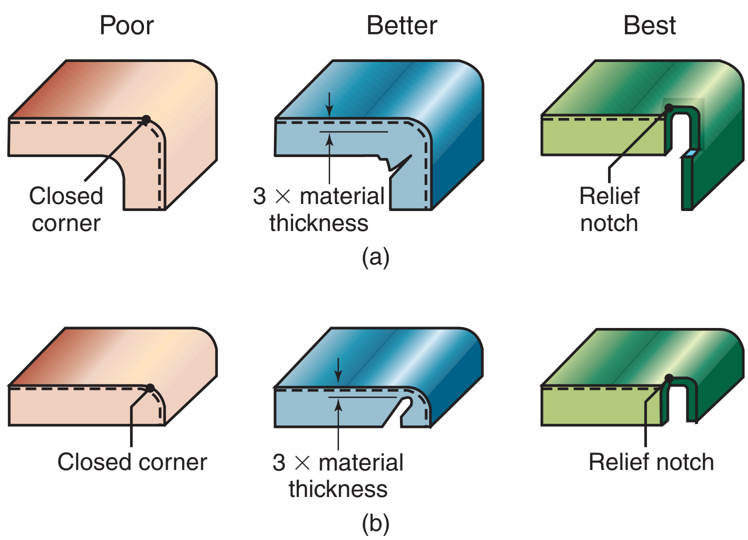


Figure 16.61: Control of tearing and buckling of a flange in a right-angle bend. *Source:* Courtesy of Society of Manufacturing Engineers.

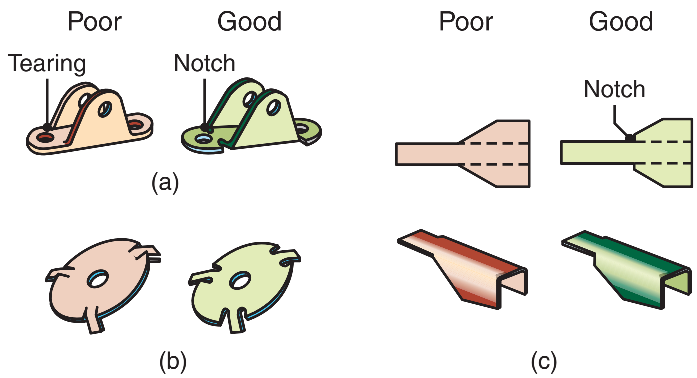


Figure 16.62: Application of notches to avoid tearing and wrinkling in right-angle bending operations. *Source:* Courtesy of Society of Manufacturing Engineers.

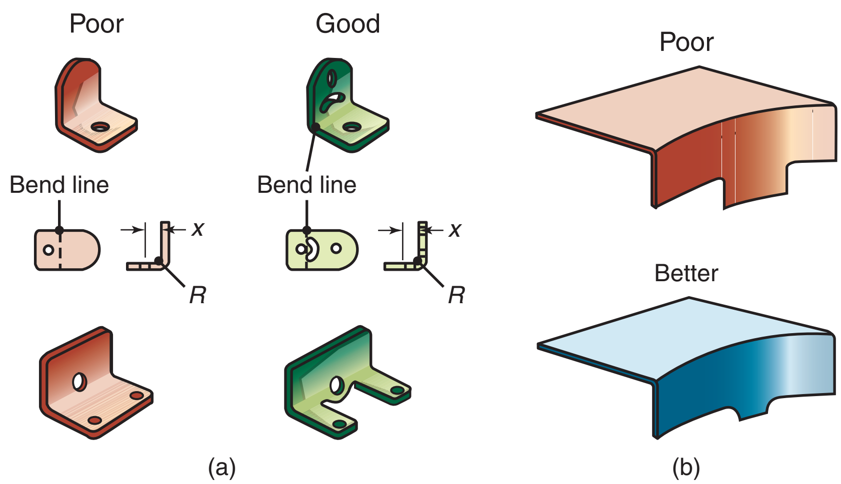


Figure 16.63: Stress concentrations near bends. (a) Using a crescent or ear for a hole near a bend. (b) Reducing the severity of tab in a flange. *Source:* Courtesy of Society of Manufacturing Engineers.

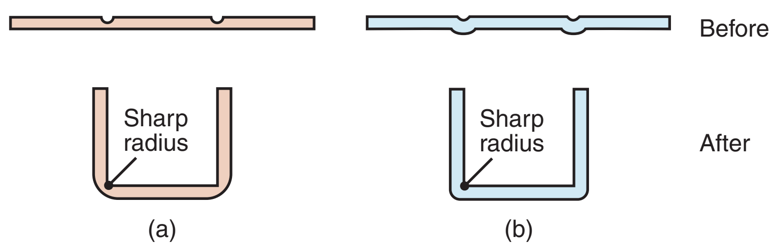


Figure 16.64: Application of (a) scoring or (b) embossing to obtain a sharp inner radius in bending. Unless properly designed, these features can lead to fracture. *Source:* Courtesy of Society of Manufacturing Engineers.

Roll Forming. The process should, in general, be designed so as to control springback. Also, it is not difficult to include perforating rolls in the forming line, so that periodic holes, notches, or embossed areas can be located on the roll-formed shape.

Stamping and Progressive-die Operations. In progressive dies (Section 16.2.3), the cost of tooling and the number of stations are determined by the number and spacing of the features on a part. Thus, it is advantageous to keep the number of features to a minimum. Closely spaced features may provide insufficient clearance for punches, and may require two punches. Narrow cuts and protrusions may present difficulties in forming with a single punch and die set.

Deep Drawing. After a cup is deep drawn, it invariably will spring back slightly towards its original shape. For this reason, designs requiring a vertical wall may be difficult to draw. Relief angles, at least 3° on each wall, make it easier to draw. Cups with sharp internal radii are difficult to draw, and deep cups will often require one or more subsequent ironing operations.

16.15 Equipment for Sheet-metal Forming

For most general pressworking operations, the basic equipment consists of mechanical, hydraulic, pneumatic, or pneumatic–hydraulic presses; they are available with a wide variety of designs, features, capacities, and computer controls. Servo presses (Section 14.8) are now being used for forming sheet metals, because of their ability to vary speed and forces in a controlled manner during forming. Typical designs for press frames are shown in Fig. 16.65 (see also Figs. 14.19 and 16.23f). The proper design, stiffness, and construction of such equipment is essential to the efficient operation of the system, and for achieving high production rate, good dimensional accuracy, and high product quality.

The traditional **C-frame** structure (Fig. 16.65a) has been used widely for ease of tool and workpiece accessibility; however, it is not as stiff as the **box-type pillar** (Fig. 16.65e) or the **double-column frame** structure (Fig. 16.65f). Accessibility to working areas in presses has become less important, due to advances in automation and in the use of industrial robots and computer controls.

Press selection for sheet-metal forming operations depends on several factors:

1. Type of forming operation, size and shape of dies, and tooling required

2. Size and shape of parts

3. Length of stroke of the slide, number of strokes per minute, operating speed, and shut height (the distance from the top of the bed to the bottom of the slide, with the stroke down)

4. Number of slides: single-action presses have one reciprocating slide; double-action presses have two slides, reciprocating in the same direction; they typically are used for deep drawing, one slide for the punch and the other for the blankholder. Triple-action presses have three slides; they generally are used for reverse redrawing and other complicated forming operations.

5. Maximum force required (press capacity and tonnage rating)

6. Type and level of mechanical, hydraulic, and computer controls

7. Features for changing the dies; because the time required for die changes can be significant (as much as a few hours), they affect productivity, for which rapid die-changing systems have been developed. In a system called SMED (*single-minute exchange of dies*), die setups can be changed in less than 10 minutes, by using computer-controlled hydraulic or pneumatic systems.

8. Safety features.

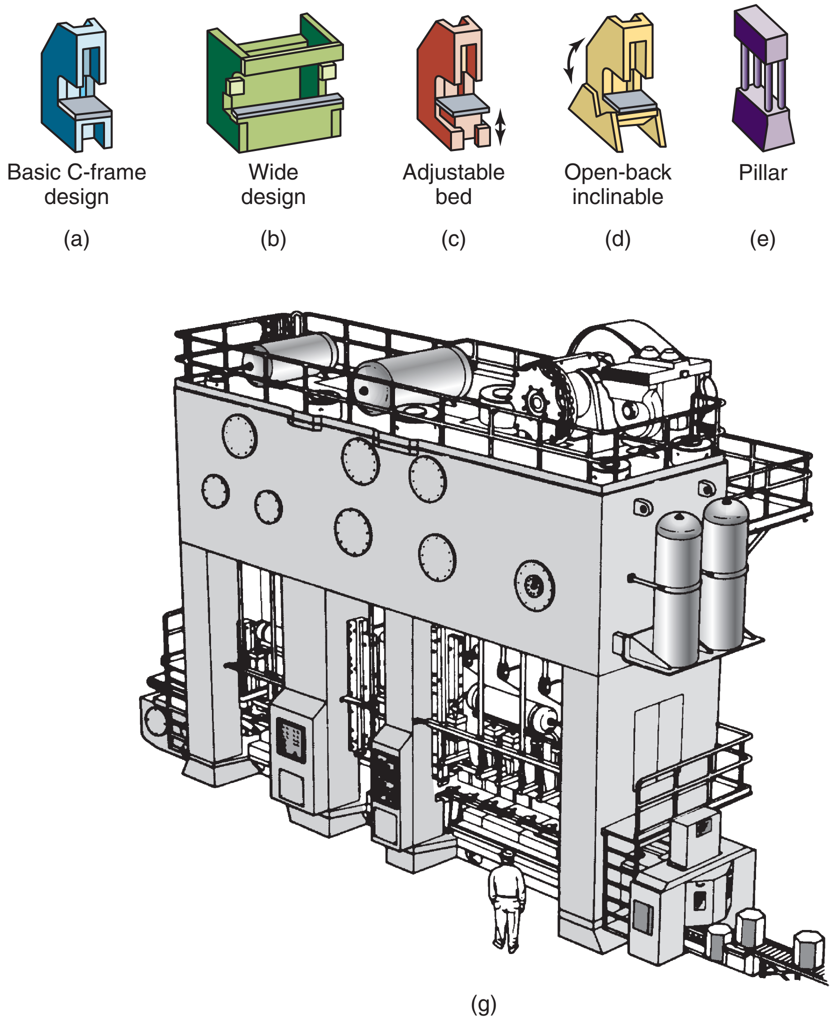


Figure 16.65: (a) through (f) Schematic illustrations of types of press frames for sheet-metal forming operations. Each type has its own characteristics of stiffness, capacity, and accessibility. (g) A large stamping press. *Source:* (g) Verson Allsteel Company.

Because a press can be a major capital investment (see Table 40.6), its present and its future use for making a broad variety of parts and applications for a long period of time must be investigated. Versatility and multiple different uses are important factors in press selection, particularly for product modifications and for making new products to respond to continually changing market demands.

16.16 Economics of Sheet-forming Operations

Sheet-metal forming involves economic considerations similar to those for the other metalworking operations. Sheet-forming operations are very versatile, and a number of different processes can be considered to produce the same part economically. The costs involved (Section 40.10) depend on the particular operations, such as tools, dies, molds, and equipment costs and labor. For small and simple parts, costs and lead times to make the dies are relatively low. On the other hand, for large-scale operations, such as stretch

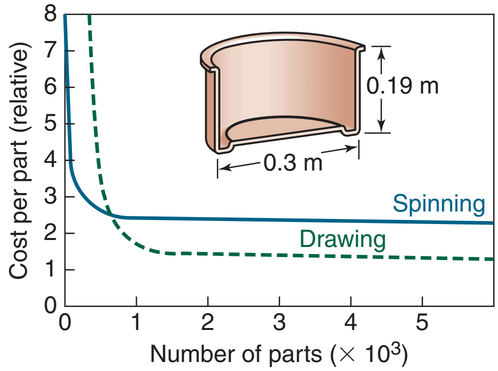


Figure 16.66: Cost comparison for manufacturing a round sheet-metal container either by conventional spinning or by deep drawing. Note that for small quantities, spinning is more economical.

forming of aircraft panels and making boat hulls, these costs are very high. Moreover, because the volume of such products made is low, the cost per piece can be very high (see also Fig. 14.21).

Deep drawing, for example, requires expensive dies and tooling, but a very large number of parts can be produced with the same setup, such as beverage cans, metal containers, and similar household products. These costs for other processes, such as punching, blanking, bending, and spinning, vary considerably.

Equipment costs can vary widely, depending largely on the complexity of the forming operation, part loading and unloading features, part size and shape, and level of automation and computer controls involved. Automation, in turn, directly affects the labor and the skill level required; note that the higher the extent of automation, the lower the skill level required. Furthermore, sheet-metal parts generally require some finishing operations, one of the most common being deburring of the part edges, which generally is labor intensive. Even though significant advances have been made in automated deburring, it still requires costly computer-controlled equipment.

As an example of the versatility of sheet-forming operations and the costs involved, recall that a cup-shaped part can be made by (a) deep drawing, (b) spinning, (c) rubber forming, (d) explosive forming, impact extrusion, (e) casting, or (f) fabrication by assembling or welding together different pieces. Each of these methods involves different processes, and different costs. The part shown in Fig. 16.66, for example, can be made either by deep drawing or by conventional spinning, but the tooling costs for the two processes are significantly different.

Deep-drawing dies have several components, and making them costs much more than the relatively simple mandrels and tools required in such a process as spinning; note also that surface finish, dimensional accuracy, and the properties of the two products will be different Consequently, the tooling cost per part in drawing will be high, especially if only a few parts are required. This part also can be shaped by deep drawing and in much shorter time than by spinning, even if the latter operation is automated and computer controlled; also, spinning generally requires more skilled labor. Considering these factors, the break-even point for this part is around 700 parts, and for quantities higher than that, deep drawing is more economical.

Summary

- Sheet-metal-forming processes are among the most versatile of all metalworking operations. They generally are used on workpieces having high ratios of surface area to thickness. Unlike bulk deformation processes, such as forging and extrusion, sheet-metal forming operations do not undergo much change in their thickness.

- Several test methods have been developed for predicting the formability of sheet metals.

- Important material parameters are the quality of the sheared edge of the blank, the capability of the sheet to stretch uniformly, the material's resistance to thinning, its normal and planar anisotropy, grain size, and for low-carbon steels, yield-point elongation.
- The forces and energy required in forming processes are transmitted to the sheet through either solid tools and dies or by flexible rubber or polyurethane members, or by electrical, chemical, magnetic, and gaseous means.
- Because sheet metals are thin, springback, buckling, wrinkling, and tearing are significant factors in shaping them. These tendencies can be reduced or eliminated by proper tool and die design, and minimizing the unsupported length of the sheet during processing.
- Superplastic forming of diffusion-bonded sheets is an important process for making complex sheet-metal structures, particularly for applications where high stiffness-to-weight ratios are important.
- For general sheet-forming operations, forming-limit diagrams are very useful, because they establish quantitative relationships among the major and minor principal strains, indicating safe regions of forming.

Key Terms

Beading
Bendability
Bend allowance
Bending
Blankholder
Blanking
Bulging
Burnished surface
Burr
Clearance
Compound dies
Deburring
Deep drawing
Dent resistance
Dimpling
Drawbead
Drawing
Earing
Electrohydraulic forming
Embossing
Explosive forming
Fine blanking
Flanging
Formability
Forming-limit diagram
Hemming
Honeycomb structures
Hot stamping
Hydroform process
Incremental forming
Ironing
Laser forming
Limiting drawing ratio
Lüder's bands
Magnetic-pulse forming
Microforming
Minimum bend radius
Nesting
Nibbling
Normal anisotropy
Peen forming
Planar anisotropy
Plastic anisotropy
Press brake
Progressive dies
Punching
Redrawing
Roll forming

Rubber forming
Shaving
Shearing
Slitting
Spinning
Springback
Steel rule
Stretch forming
Superplastic forming
Tailor-welded blanks
Transfer dies

Bibliography

Altan, T., and Tekkaya, E. (eds.), **Sheet Metal Forming: Processes and Applications**, ASM International, 2012.

ASM Handbook, Vol. 14B: **Metalworking: Sheet Forming,** ASM International, 2006.

Boljanovic, V., **Sheet Metal Forming Process and Die Design**, 2nd ed., Industrial Press, 2014.

Davies, G., **Materials for Automobile Bodies**, 2nd ed., Butterworth-Heinemann, 2012.

Fundamentals of Tool Design, 6th ed., Society of Manufacturing Engineers, 2010.

Hosford, W.F., and Caddell, R.M., **Metal Forming: Mechanics and Metallurgy**, 4th ed., Cambridge, 2014.

Hu, P., Ying, L., and He, B., **Hot Stamping Advanced Manufacturing Technology of Lightweight Car Body**, Elsevier, 2016.

Jie, L.Y., **Sheet Metal Bending Technology**, Mechanical Industry, 2000.

Kumar, S, and Hussein, H.M.A., **AI Applications in Sheet Metal Forming**, Springer, 2016.

Marciniak, Z., Duncan, J.L., and Hu, S.J., **Mechanics of Sheet Metal Forming**, 2nd ed., Butterworth-Heinemann, 2002.

Rapien, B.L., **Fundamentals of Press Brake Tooling**, Hanser, 2010.

Suchy, I., **Handbook of Die Design**, 2nd ed., McGraw-Hill, 2005.

Tschaetch, H., **Metal Forming Practise: Processes, Machines, Tools**, Springer, 2006.

Review Questions

16.1. How does sheet-metal forming differ from rolling, forging, and extrusion?

16.2. What causes burrs? How can they be reduced or eliminated?

16.3. Explain the difference between punching and blanking.

16.4. Describe the difference between compound, progressive, and transfer dies.

16.5. Describe the characteristics of sheet metals that are important in sheet-forming operations. Explain why they are important.

16.6. Describe the features of forming-limit diagrams (FLDs).

16.7. List the properties of materials that influence springback. Explain why and how they do so.

16.8. Give one specific application for each of the common bending operations described in this chapter.

16.9. Why do tubes buckle when bent? What is the effect of the tube thickness-to-diameter ratio?

16.10. Define normal anisotropy, and explain why it is important in determining the deep drawability of a material.

16.11. Describe earing and why it occurs.

16.12. What are the advantages of rubber forming? Which processes does it compete with?

16.13. Explain the difference between deep drawing and redrawing.

16.14. How is roll forming fundamentally different from rolling?

16.15. What is nesting? What is its significance?

16.16. Describe the differences between compound, progressive, and transfer dies.

16.17. What is microforming?

16.18. Explain the advantages of superplastic forming.

16.19. What is hot stamping? For what materials is it used?

16.20. What is springback? What is negative springback?

Qualitative Problems

16.21. Explain the differences that you have observed between products made of sheet metals and those made by casting and forging.

16.22. Take any three topics from Chapter 2, and, with specific examples for each, show their relevance to the topics covered in this chapter.

16.23. Do the same as for Problem 16.22, but for Chapter 3.

16.24. Identify the material and process variables that influence the punch force in shearing, and explain how each of them affects this force.

16.25. Explain why springback in bending depends on yield stress, elastic modulus, sheet thickness, and bend radius.

16.26. Explain why cupping tests may not predict well the formability of sheet metals in actual forming processes.

16.27. Identify the factors that influence the deep-drawing force, F, in Fig. 16.32b, and explain why they do so.

16.28. Why are the beads in Fig. 16.36b placed in those particular locations?

16.29. A general rule for dimensional relationships for successful drawing without a blankholder is given by Eq. (16.14). Explain what would happen if this limit were exceeded.

16.30. Section 16.2.1 stated that the punch stripping force is difficult to estimate because of the many factors involved. Make a list of these factors with brief explanations about why they would affect the stripping force.

16.31. Is it possible to have ironing take place in an ordinary deep-drawing operation? What is the most important factor?

16.32. Note the roughness of the periphery of the flanged hole in Fig. 16.25c, and comment on its possible effects when the part is used in a product.

16.33. What recommendations would you make in order to eliminate the cracking of the bent piece shown in Fig. 16.17c? Explain your reasons.

16.34. It has been stated that the quality of the sheared edges can influence the formability of sheet metals. Explain why.

16.35. Give several specific examples from this chapter in which friction is desirable and several in which it is not desirable.

16.36. As you can see, some of the operations described in this chapter produce considerable scrap. Describe your thoughts regarding the reuse, recycling, or disposal of this scrap. Consider its size, its shape, and its contamination by metalworking fluids during processing.

16.37. Through changes in clamping or die design, it is possible for a sheet metal to undergo a negative minor strain. Explain how this effect can be advantageous.

16.38. How would you produce the part shown in Fig. 16.43b other than by tube hydroforming?

16.39. It has been stated that the thicker the sheet metal, the higher is the curve in the forming-limit diagram. Explain why.

16.40. If a cupping test (see Fig. 16.13) were to be performed using a pressurized lubricant instead of a spherical die, would you expect the forming limit diagram to change? Why or why not?

16.41. What are the advantages of rubber forming? Which processes does it compete with?

16.42. Which of the processes described in this chapter use only one die? What are the advantages of using only one die?

16.43. It has been suggested that deep drawability can be increased by (a) heating the flange and/or (b) chilling the punch by some suitable means. Comment on how these methods could improve drawability.

16.44. Offer designs whereby the suggestions given in Problem 16.43 can be implemented. Would the required production rate affect your designs? Explain.

Quantitative Problems

16.45. Calculate R_{avg} for a metal where the R values for the 0°, 45°, and 90° directions are 0.9, 1.6, and 1.9, respectively. What is the limiting drawing ratio (LDR) for this material?

16.46. Calculate the value of ΔR in Problem 16.45. Will any ears form when this material is deep drawn? Explain.

16.47. Estimate the limiting drawing ratio for the materials listed in Table 16.4.

16.48. Using Eq. (16.15) and the K value for TNT, plot the pressure as a function of weight (W) and R, respectively. Describe your observations.

16.49. Section 16.5 states that the k values in bend allowance depend on the relative magnitudes of R and T. Explain why this relationship exists.

16.50. For explosive forming, calculate the peak pressure in water for 0.11 kg of TNT at a standoff distance of 1.5 m. Comment on whether or not the magnitude of this pressure is sufficiently high to form sheet metals.

16.51. Measure the respective areas of the solid outlines in Fig. 16.14a, and compare them with the areas of the original circles. Calculate the final thicknesses of the sheets, assuming that the original sheet is 1-mm thick.

16.52. Plot Eq. (16.6) in terms of the elastic modulus, E, and the yield strength, S_y, of the material, and describe your observations.

16.53. What is the minimum bend radius for a 2.0-mm-thick sheet metal with a tensile reduction of area of 35%? Does the bend angle affect your answer? Explain.

16.54. Survey the technical literature and explain the mechanism by which negative springback can occur in V-die bending. Show that negative springback does not occur in air bending.

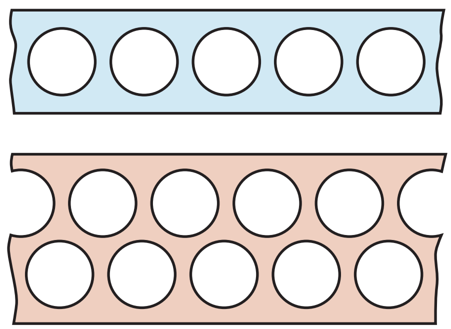
16.55. Using the data in Table 16.3 and referring to Eq. (16.5), calculate the tensile reduction of area for the materials and the conditions listed in the table.

16.56. What is the force required to punch a square hole 25 mm on each side in a 0.15-mm-thick 5052-O aluminum sheet by using flat dies? What would be your answer if beveled dies are used?

16.57. In Case Study 16.2, it was stated that the reason for reducing the tops of cans (necking) is to save material for making the lid. How much material will be saved if the lid diameter is reduced by 10%? By 20%?

16.58. A cup is being drawn from a sheet metal that has a normal anisotropy of 3. Estimate the maximum ratio of cup height to cup diameter that can be drawn successfully in a single draw. Assume that the thickness of the sheet throughout the cup remains the same as the original blank thickness.

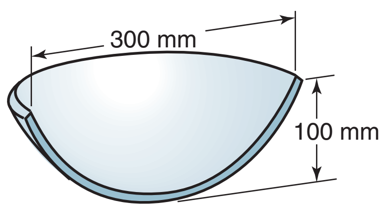
16.59. Estimate the percent scrap in producing round blanks if the clearance between blanks is one-tenth of the radius of the blank. Consider single and multiple-row blanking, as sketched below.



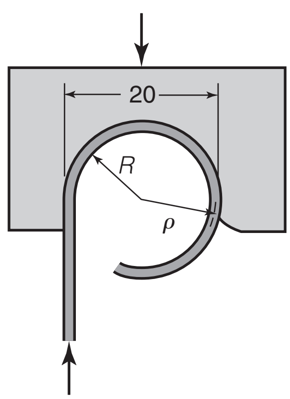
16.60. Estimate the maximum bending force required for a 3-mm thick and 300-mm wide Ti-6Al-4V titanium alloy, annealed and quenched at 25°C, in a V-die with a width of 180 mm.

16.61. Plot the final bend radius as a function of initial bend radius in bending for (a) 5052-O aluminum; (b) 5052-H34 Aluminum; (c) C24000 brass; and (d) AISI 304 stainless steel.

16.62. The figure below shows a parabolic profile that will define the mandrel shape in a spinning operation. Determine the equation of the parabolic surface. If a spun part will be produced from a 15-mm thick blank, determine the minimum required blank diameter.



16.63. A straight bead is being formed on a 2-mm-thick aluminum sheet in a 20-mm-diameter die, as shown in the accompanying figure. Let $S_y = 90$ MPa. Considering springback, calculate the outside diameter of the bead after it is formed and unloaded from the die.



16.64. Calculate and plot the springback in bending 1-mm thick sheet metal around radii from 0.25 to 250 mm for (a) 303 stainless steel; (b) 1100-O aluminum; (c) HK31A magnesium; (d) Ti-6Al-4V.

16.65. Circular blanks of 5052-O aluminum, with a diameter of 30 mm and a thickness of 4 mm are to be mass-produced as the starting material for a tube for a paintball gun. The available press has an 800 kN capacity and can take a maximum of 300 mm wide strip. Material utilization improves if more rows are cut from a strip. (a) Determine the force required to blank a single slug. (b) Determine the maximum number of slugs that can be blanked simultaneously by the press. (c) Determine the material utilization if the space around a blanked part needs to be the same as the thickness, or 4 mm.

16.66. Assume that you are an instructor covering the topics described in this chapter and you are giving a quiz on the numerical aspects to test the understanding of the students. Prepare two quantitative problems and supply the answers.

Synthesis, Design, and Projects

16.67. Examine some of the products in your home or in an automobile that are made of sheet metal, and discuss the process or combination of processes by which you think they were made.

16.68. Consider several shapes to be blanked from a large sheet (such as oval, triangular, and L-shaped) by laser-beam cutting, and sketch a nesting layout to minimize scrap generation.

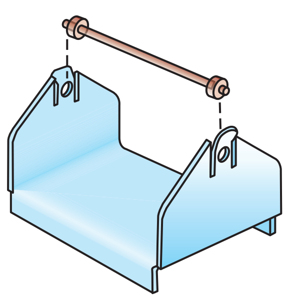
16.69. Give several specific product applications for (a) hemming and (b) seaming.

16.70. Many axisymmetric missile bodies are made by spinning. What other methods could you use if spinning processes were not available?

16.71. Give several structural designs and applications in which diffusion bonding and superplastic forming can be used jointly. Comment on whether this combination is capable of producing parts at high volume.

16.72. Metal cans are either two-piece (in which the bottom and sides are integral) or three-piece (in which the sides, the bottom, and the top are each separate pieces). For a three-piece can, should the vertical seam in the can body be (a) in the rolling direction, (b) normal to the rolling direction, or (c) oblique to the rolling direction? Prove your answer.

16.73. The design shown is proposed for a metal tray, the main body of which is made from cold-rolled sheet steel. Noting its features and that the sheet is bent in two different directions, comment on various manufacturing considerations. Include factors such as anisotropy of the rolled sheet, its surface texture, the bend directions, the nature of the sheared edges, and the way the handle is snapped in for assembly.



16.74. Suggest consumer-product designs that could utilize honeycomb structures. For example, an elevator can use a honeycomb laminate as a stiff and lightweight floor material.

16.75. Using a ball-peen hammer, strike the surface of aluminum sheets of various thicknesses until they develop a curvature. Describe your observations about the shapes produced.

16.76. Inspect a common paper punch and observe the shape of the punch tip. Compare it with those shown in Fig. 16.10 and comment on your observations.

16.77. Obtain an aluminum beverage can and slit it in half lengthwise with a pair of tin snips. Using a micrometer, measure the thickness of the can bottom and the wall. Estimate the thickness reductions in ironing and the diameter of the original blank.

16.78. In order to improve its ductility, a coil of sheet metal is placed in a furnace and annealed. However, it is observed that the sheet has a lower limiting drawing ratio than it had before being annealed. Explain the reasons for this behavior.

16.79. With automotive parts, it is often advantageous to have a part with tailored properties. For example, a pillar that provides structural support for the operator’s compartment may be strong but less ductile at the center, but more ductile and less strong where the pillar attaches to the remainder of the car structure. List ways of producing such tailored properties in hot stampings.

16.80. Give three examples of sheet metal parts that (a) can and (b) cannot be produced by incremental forming.

16.81. Conduct a literature search and obtain the equation for a tractrix curve, as used in Fig. 16.38.

16.82. On the basis of experiments, it has been suggested that concrete, either plain or reinforced, can be a suitable material for dies in sheet-metal forming operations. Describe your thoughts regarding this suggestion, considering die geometry and any other factors that may be relevant.

16.83. Investigate methods for determining optimum shapes of blanks for deep-drawing operations. Sketch the optimally shaped blanks for drawing rectangular cups, and optimize their layout on a large sheet of metal.

16.84. Design a box that will contain a 100-mm $\times$ 150-mm $\times$ 75-mm volume. The box should be produced from two pieces of sheet metal and require no tools or fasteners for assembly.

16.85. Repeat Problem 16.84, but design the box from a single piece of sheet metal.

16.86. Obtain a few pieces of cardboard and carefully cut the profiles to produce bends as shown in Fig. 16.61. Demonstrate that the designs labeled as “best” are actually the best designs. Comment on the difference in strain states between the designs.

Chapter 17

Powder-metal Processes and Equipment

- This chapter describes the principles of powder metallurgy processes for producing net-shape parts from metal powders.
- The chapter begins by reviewing methods of producing and blending of metal powders and investigates the shapes that powders will develop based on the particular process employed to make them.
- Operations such as compaction to consolidate the powder into desired shapes and sintering to fuse the particles to achieve the required strength are then described in detail.
- Additional processes particular to powder metallurgy products are then presented, and general design rules are reviewed.
- The chapter ends with process capabilities and economics of powder metallurgy as compared with other competing manufacturing operations.

Typical products made: Connecting rods, piston rings, gears, cams, bushings, bearings, cutting tools, surgical implants, magnets, metal filters, and surgical implants.

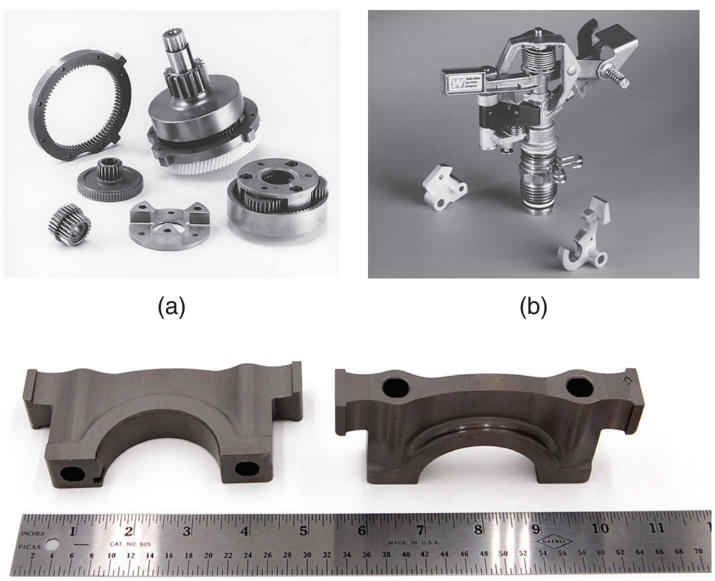
Alternative processes: Casting, forging, additive manufacturing, and machining.

17.1 Introduction

In the manufacturing processes described thus far, the raw materials have been metals and alloys that are either in a molten state (casting) or in solid form (metalworking). In **powder metallurgy** (PM), metal powders are compacted into desired and often complex shapes and then sintered (heated without melting) to form a solid piece. This process first was used in Egypt in about 3000 B.C. to make iron tools. One of its first modern uses was in the early 1900s to make the tungsten filaments for incandescent light bulbs. The availability of a wide range of metal powder compositions, the ability to produce parts to net dimensions (**net-shape forming**), and the unique mechanical properties that result from PM give this process its numerous attractive applications.

A wide range of complex parts and components can be made by powder-metallurgy techniques (Fig. 17.1): balls for ballpoint pens, piston rings, connecting rods, brake pads, gears, cams, and bushings, tool steels, tungsten carbides, and cermets as tool and die materials, graphite brushes impregnated with copper for electric motors; magnetic materials; metal filters and oil-impregnated bearings with controlled porosity; metal foams; and surgical implants; as well as *structural* parts for aircraft, such as landing gear components, engine-mount supports, engine disks, impellers, and engine nacelle frames.

Powder metallurgy has become competitive with processes such as casting, forging, and machining, particularly for relatively complex parts made of hard and high-strength alloys. Although most parts made typically weigh less than 2.5 kg, they can weigh as much as 50 kg. It has been shown that PM parts can be produced economically in quantities as small as 5000 per year, and as much as 100 million per year for vibrator weights for cell phones.



(a) (b)

(c)

Figure 17.1: (a) Examples of typical parts made by powder-metallurgy (PM) processes. (b) Upper trip lever for a commercial irrigation sprinkler made by PM. This part is made of an unleaded brass alloy; it replaces a die-cast part with a 60% cost savings. (c) Main-bearing powder-metal caps for 3.8- and 3.1-liter General Motors automotive engines. *Source:* (a) and (b) Reproduced with permission from *Success Stories on PM Parts*, Metal Powder Industries Federation, Princeton, New Jersey, 1998. (c) Courtesy of the Metal Powder Industries Federation.

The most commonly used metals in PM are iron, copper, aluminum, tin, nickel, titanium, and refractory metals. For parts made of brass, bronze, steels, and stainless steels, *prealloyed powders* are used, where each powder particle itself is an alloy. The sources for metals are generally bulk metals and alloys, ores, salts, and various other compounds.

17.2 Production of Metal Powders

The powder-metallurgy process basically consists of the following operations, in sequence (Fig. 17.2):

1. **Powder production**
2. **Blending**
3. **Compaction**
4. **Sintering**
5. **Finishing.**

17.2.1 Methods of Powder Production

There are several methods of producing metal powders, and most powders can be produced by more than one method; the choice depends greatly on the requirements of the end product. The microstructure, bulk and surface properties, chemical purity, porosity, shape, and size distribution of the particles depend on the particular process used (Figs. 17.3 and 17.4). These characteristics are important because they significantly affect flow and permeability during compaction and in subsequent sintering operations. Particle sizes produced typically range from 0.1 to 1000 μm.

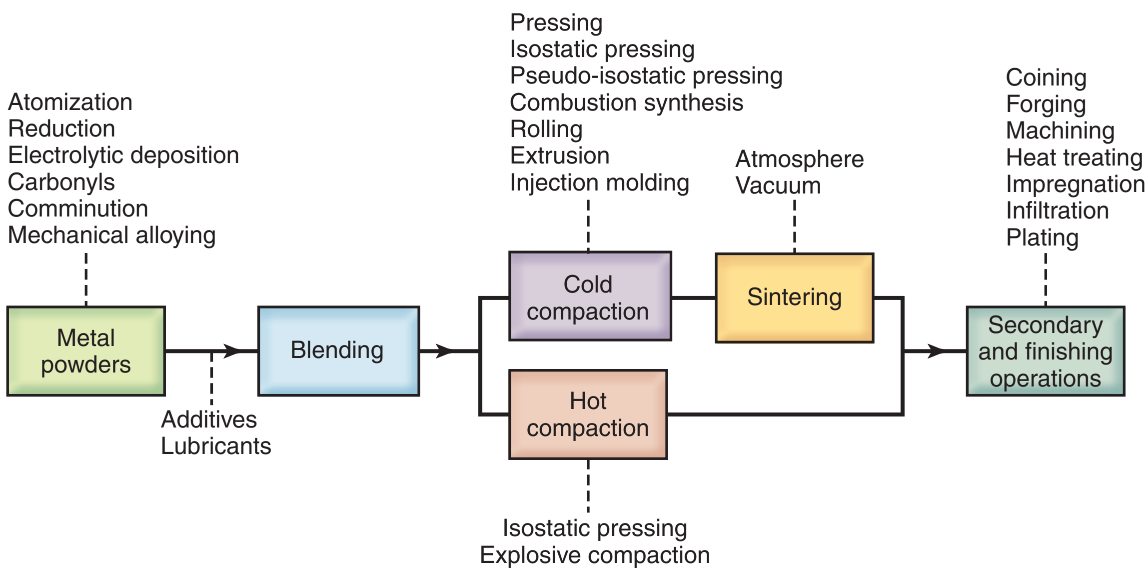


Figure 17.2: Outline of processes and operations involved in producing powder-metallurgy parts.

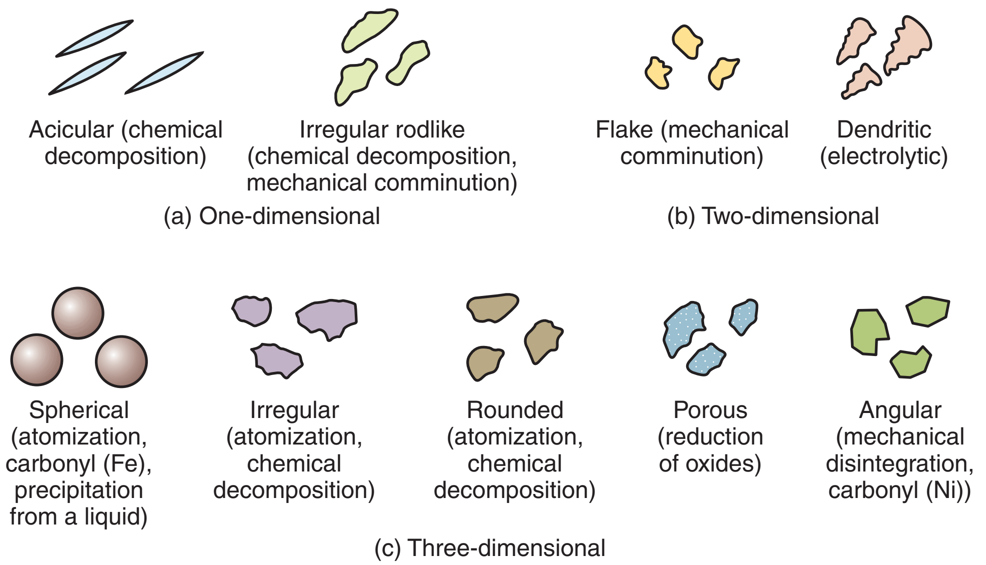


Figure 17.3: Particle shapes in metal powders and the processes by which they are produced. Iron powders are produced by many of these processes (see also Fig. 17.4).

Atomization. Atomization involves a liquid-metal stream, produced by injecting molten metal through a small orifice, whereby the stream is broken up by jets of inert gas or air (Fig. 17.5a) or water (Fig. 17.5b), known as *gas* or *water atomization*, respectively. The size and shape of the particles formed depend on such factors as the temperature of the molten metal, rate of flow, nozzle size, and jet characteristics. The use of water results in a slurry of metal powder and liquid at the bottom of the atomization chamber. Although the powders must be dried before they can be used, the water allows for more rapid cooling of the particles, resulting in higher production rates. Gas atomization usually produces more spherical particles (see Fig. 17.3c).

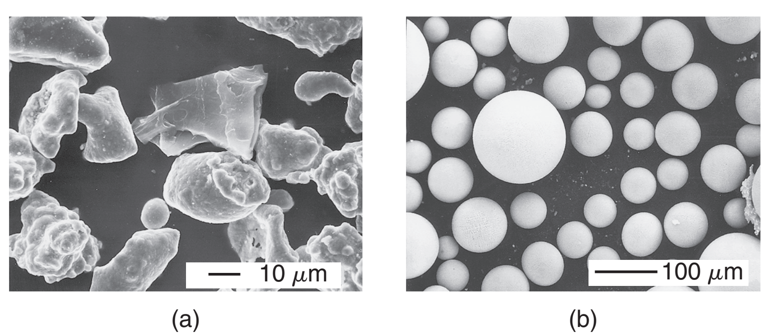


Figure 17.4: (a) Scanning-electron microscopy image of iron-powder particles made by atomization. (b) Nickel-based superalloy (Udimet 700) powder particles made by the rotating electrode process; see Fig. 17.5d. *Source:* After P.G. Nash, Illinois Institute of Technology, Chicago.

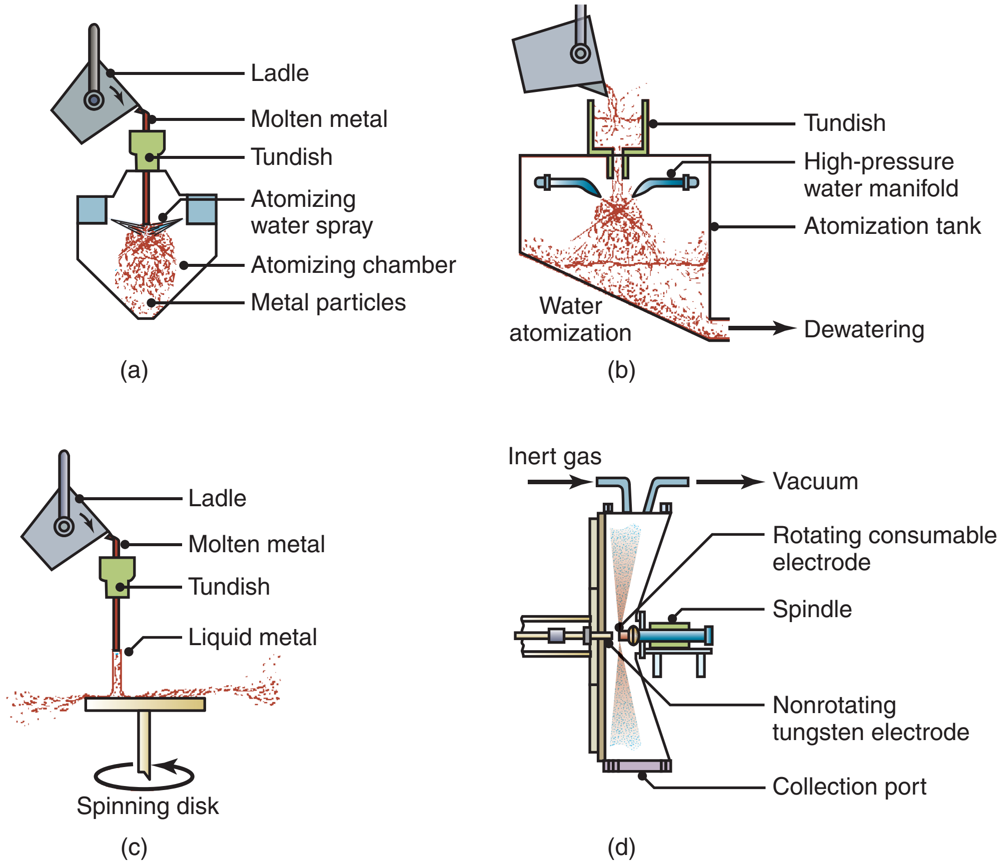


Figure 17.5: Methods of metal-powder production by atomization: (a) gas atomization; (b) water atomization; (c) centrifugal atomization with a spinning disk or cup; and (d) atomization with a rotating consumable electrode.

In *centrifugal atomization*, the molten-metal stream drops onto a rapidly rotating disk or cup; the centrifugal forces then break up the stream and generate particles (Fig. 17.5c). In a variation of this method, a consumable electrode is rotated rapidly (at about 15,000 rev/min) in a helium-filled chamber (Fig. 17.5d); the centrifugal force then breaks up the molten tip of the electrode into metal particles.

Reduction. The *reduction* of metal oxides (removing oxygen) uses gases such as hydrogen and carbon monoxide as reducing agents. By this means, very fine metallic oxides are reduced to the metallic state. The powders produced are spongy and porous, with uniformly sized spherical or angular shapes.

Electrolytic Deposition. *Electrolytic deposition* utilizes aqueous solutions or fused salts. The powders produced are among the purest made.

Carbonyls. *Metal carbonyls*, such as iron carbonyl $[Fe(CO)_5]$ and nickel carbonyl $[Ni(CO)_4]$, are formed by allowing iron or nickel to react with carbon monoxide. The reaction products are then decomposed to iron and nickel, and they turn into small, dense, and uniformly spherical particles, with high purity.

Comminution. *Mechanical comminution* (*pulverization*) involves either crushing (Fig. 17.6), milling in a ball mill, or grinding brittle or less ductile metals into small particles. A *ball mill* (Fig. 17.6b) is a machine with a rotating hollow cylinder partly filled with steel or white cast-iron balls. The powder or particles placed

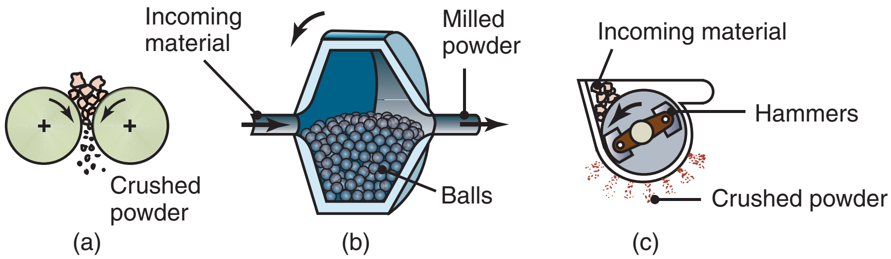


Figure 17.6: Methods of mechanical comminution to obtain fine particles: (a) roll crushing, (b) ball mill, and (c) hammer milling.

into a ball mill are thus impacted by the balls as the cylinder is rotated, or its contents may be agitated. This action has two effects: (a) the particles are periodically fractured, resulting in smaller particles; and (b) the shape of the particles is affected, such that with brittle materials, the particles produced have angular shapes. With ductile metals, they are flaky and not particularly suitable for powder-metallurgy applications.

Mechanical Alloying. In *mechanical alloying*, powders of two or more materials are mixed in a ball mill, as illustrated in Fig. 17.7. Under the impact of the hard balls, the powders fracture and bond together by diffusion, entrapping the second phase and forming alloy powders or composites. The dispersed phase can result in the strengthening of the particles or it can impart special electrical or magnetic properties to the powder.

Miscellaneous Methods. Less commonly used methods for making powders are:

- **Precipitation** from a chemical solution
- Production of fine metal chips by **machining**
- **Vapor condensation.**

Nanopowders. More recent developments include the production of *nanopowders* of copper, aluminum, iron, titanium, and various other metals (see also *nanomaterials* in Section 8.8). When the material is subjected to large plastic deformation, by compression and shear and at stress levels of 5500 MPa during processing, the particle size is reduced and the material becomes pore free, thus possessing

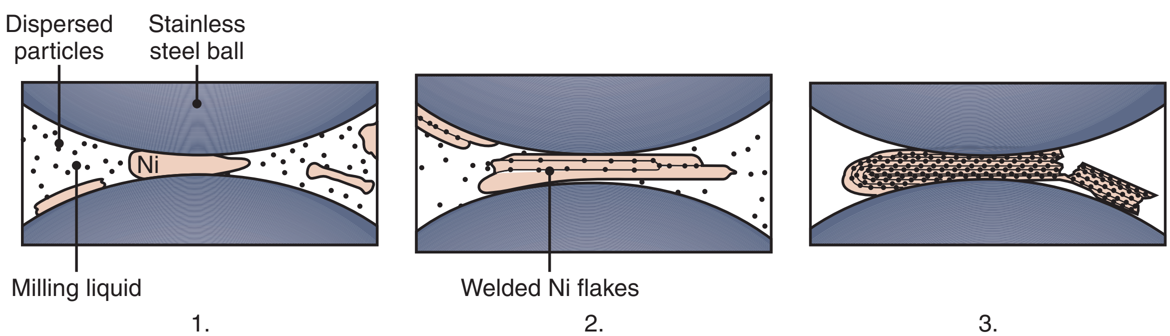


Figure 17.7: Mechanical alloying of nickel particles with dispersed smaller particles. As the nickel particles are flattened by the balls, the second, smaller phase is impressed into the nickel surface and eventually is dispersed throughout the particle due to successive flattening, fracture, and welding that occur.

enhanced properties. Because these powders are *pyrophoric* (they ignite spontaneously), or are contaminated readily when exposed to air, they are shipped as thick slurries under hexane gas (which itself is highly volatile and combustible).

Microencapsulated Powders. In micro encapsulation, metal powders are coated completely with a binder. For electrical applications, such as magnetic components of ignition coils and in other pulsed AC and DC applications, the binder acts as an insulator, preventing electricity from flowing between the particles and thus reducing eddy-current losses. The powders are compacted by warm pressing, and are used with the binder still in place (see also *powder-injection molding*, Section 17.3.3).

17.2.2 Particle Size, Shape, and Distribution

Particle size is generally controlled by *screening*, that is, passing the metal powder through screens (*sieves*) of various mesh sizes. The horizontal screens are stacked on top of each other, with the mesh size becoming finer as the powder flows downward through the screens. The larger the mesh size, the smaller is the opening in the screen. A mesh size of 30, for example, has an opening of 600 μm, size 100 has 150 μm, and size 400 has 38 μm. This method is similar to the numbering of abrasive grains; the larger the number, the smaller is the size of the abrasive particle (see Section 26.2).

Several other methods also are available for particle-size analysis:

1. **Sedimentation**, which involves measuring the rate at which particles settle in a fluid.
2. **Microscopic analysis**, which may include the use of transmission and scanning-electron microscopy.
3. **Light scattering**, from a laser that illuminates a sample, consisting of particles suspended in a liquid medium. The particles cause the light to be scattered; a detector then digitizes the signals and computes the particle-size distribution.
4. **Optical methods**, such as particles blocking a beam of light, whereby they are sensed by a photocell.
5. **Suspending particles** in a liquid and detecting particle size and distribution by electrical sensors.

Particle Shape. A major factor in processing characteristics, particle shape is described in terms of aspect ratio or shape factor. *Aspect ratio* is the ratio of the largest dimension to the smallest dimension of the particle, and ranges from unity (for a spherical particle) to 10 or higher (for flakelike or needlelike particles).

Also called the *shape index*, **shape factor** (SF) is the ratio of the surface area of the particle to its volume. It is normalized by reference to a spherical particle of equivalent volume; thus, the shape factor for a flake is higher than that for a sphere.

Size Distribution. The size distribution of particles is an important consideration because it affects the processing characteristics of the powder. The distribution of particle size is given in terms of a *frequency-distribution plot* (see Section 36.7), where the maximum is called the *mode size*. Other properties of metal powders that have an effect on their behavior in processing are (a) *flow properties*, when the powders are being filled into dies, (b) *compressibility*, when they are being compacted, and (c) *density*.

17.2.3 Blending Metal Powders

Blending (mixing) powders is the next step in powder-metallurgy processing; it is carried out for the following purposes:

- Powders of different metals can be blended in order to impart special physical and mechanical properties and characteristics to the product. Mixtures of metals can be produced by alloying the metal before producing its powder, or else the blends themselves can be produced. Proper mixing is essential to ensure the uniformity of mechanical properties throughout the part.

- Even when made of a single metal, powders may vary significantly in their size and shape, hence they must be blended to ensure uniformity from part to part. An ideal mix is one in which all of the particles of each material, and of each size and morphology, are distributed uniformly. Blending powders of two size distributions can result in a higher compacted density. This can be visualized by considering a HCP structure (see Fig. 1.5) and then considering the same structure with small particles filling the gaps between large particles. However, this denser compact may not have the uniformity of makeup as a blend of equally sized powders.

- *Lubricants* can be mixed with the powders to improve their flow characteristics. They reduce friction between the metal particles, improve flow of the powder mix into dies, and improve die life. Common lubricants are stearic acid or zinc stearate, in a proportion of from 0.25 to 5% by weight.

- Other additives, such as *binders* (as in sand molds, Section 11.2.1), are used to impart sufficient *green strength* (see Section 17.3); *additives* also can be used to facilitate sintering.

Powder mixing must be carried out under controlled conditions to avoid contamination and deterioration. *Deterioration* is caused by excessive mixing, which may alter the shape of the particles and work-harden them, thus making the subsequent compaction process more difficult. Powders can be mixed in air, in inert atmospheres (to avoid oxidation), or in liquids, which act as lubricants and make the mix more uniform. Several types of blending equipment are available (Fig. 17.8). These operations are now controlled by microprocessors to improve and maintain quality.

Hazards. Because of their typically high surface area-to-volume ratio, metal powders can become explosive, particularly aluminum, magnesium, titanium, zirconium, and thorium. Great care must therefore be exercised, both during blending and in storage and handling. Precautions include (a) maintaining a humidity

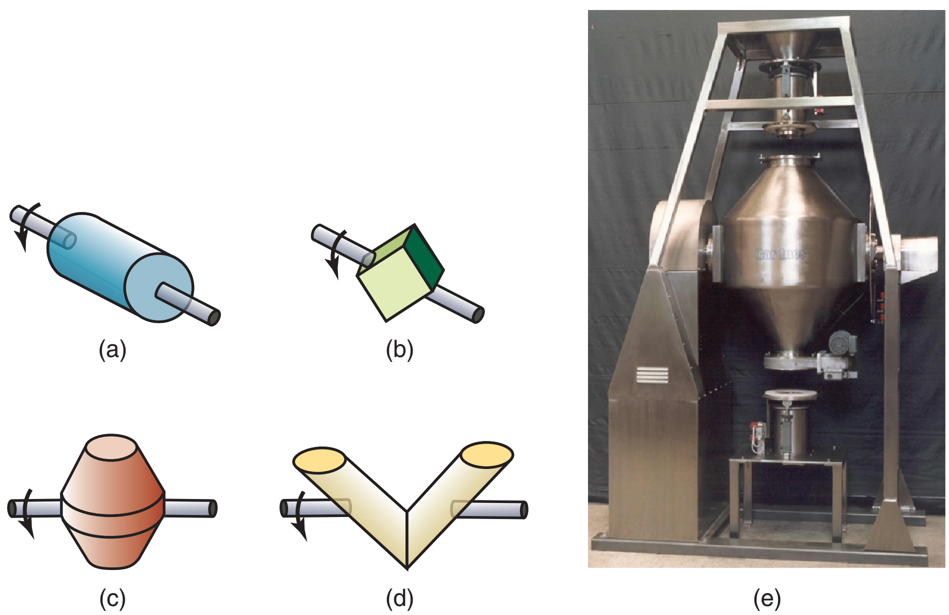


Figure 17.8: Powder mixing and blending. (a) through (d) Some common bowl geometries for mixing or blending powders. (e) A mixer suitable for blending metal powders. Since metal powders are abrasive, mixers rely on the rotation or tumbling of enclosed geometries, as opposed to using aggressive agitators. *Source:* Courtesy of Kemutec Group, Inc.

and grounding equipment and personnel to prevent static charge, (b) preventing sparks by using non-sparking tools, (c) avoiding friction as a source of heat, and (d) avoiding dust clouds and exposed ignition sources, such as open flames.

17.3 Compaction of Metal Powders

Compaction, or **pressing**, is the step in which the blended powders are pressed into dies, as shown in Fig. 17.9. The purposes of compaction are to (a) impart the required shape, density, and particle-to-particle contact and (b) make the part sufficiently strong for further processing. The powder, or *feedstock*, is fed into the die by a *feed shoe*, and the upper punch then descends into the die. The presses used are actuated either hydraulically or mechanically. The process generally is carried out at room temperature, although it can also be done at elevated temperatures for high melting-point metals.

The stages in powder compaction are shown in Fig. 17.10. First, the powder is loosely packed, so there is significant porosity. With low applied pressure, the powder rearranges itself, filling the voids and producing a denser powder, but the stresses at contact points among the powders particles are still low. Continued compaction causes increased contact stress and thus plastic deformation of the powders, resulting in increased powder adhesion and compaction.

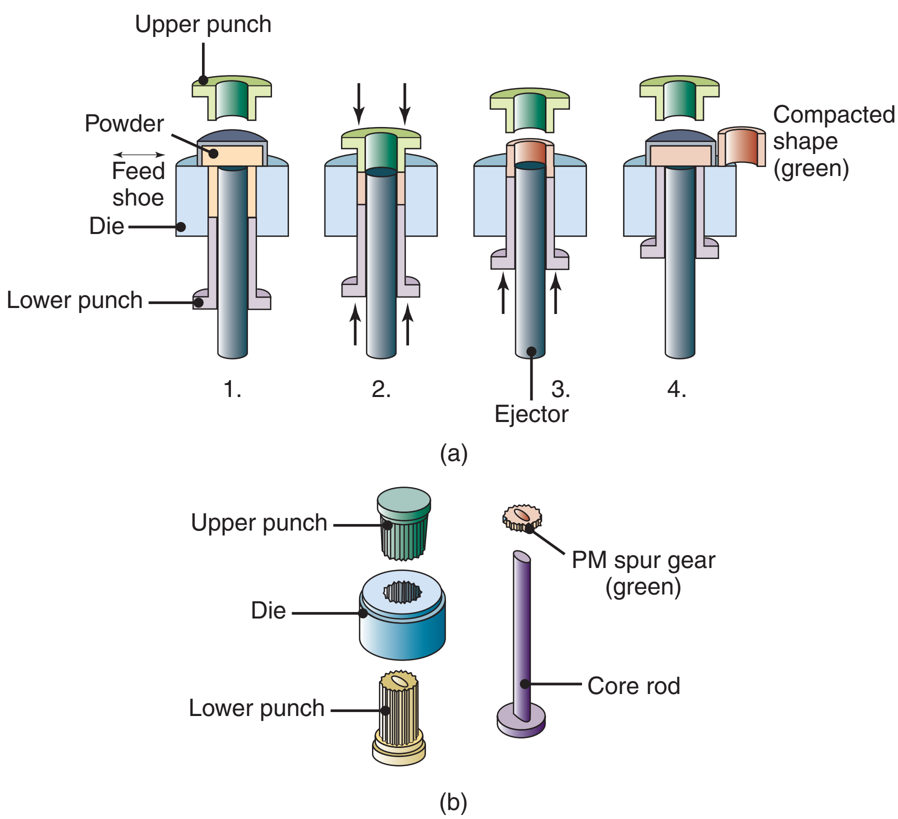


Figure 17.9: (a) Compaction of metal powder to form a bushing. The pressed-powder part is called green compact. (b) Typical tool and die set for compacting a spur gear. *Source:* Courtesy of the Metal Powder Industries Federation.

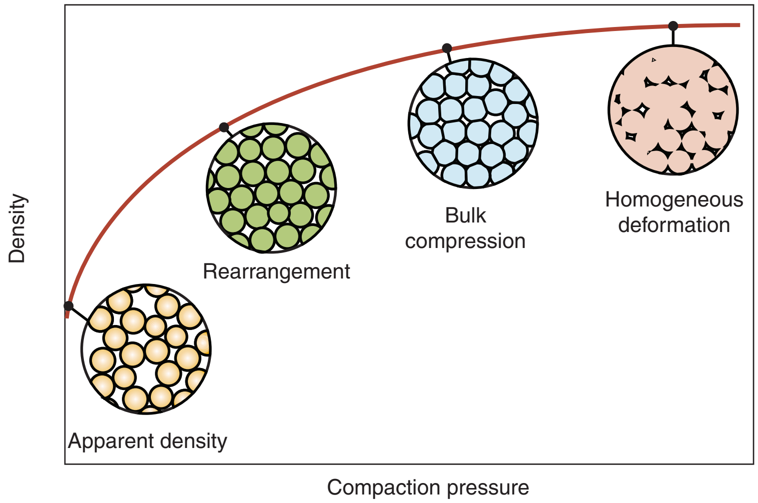


Figure 17.10: Compaction of metal powders. At low compaction pressures, the powder rearranges without being deformed, leading to a high rate of density increase. Once the powders are more closely packed, plastic deformation occurs at their interfaces, leading to further density increases although at lower rates. At very high densities, the powder behaves like a bulk solid.

The pressed powder is known as **green compact**, because the part has low strength, just as in green parts in slip casting (Section 18.2.1). These parts are very fragile and can easily crumble or become damaged—a situation that is exacerbated by poor pressing practices, such as rough handling or insufficient compaction. For higher green strength, the powder must be fed properly into the die cavity and sufficient pressure must be developed throughout the part.

Density is relevant during three different stages in PM processing: (1) as loose powder, (2) as a green compact, and (3) after sintering. The particle shape, average size, and size distribution all affect the packed density of loose powder. An important factor in density is the *size distribution* of the particles. If all of the particles are of the same size, then there will always be some porosity when packed together. Theoretically, the porosity is at least 24% by volume. Observe, for example, a box filled with tennis balls; there are always open spaces between the individual balls. Introducing smaller particles into the powder mix will begin to fill the spaces between the larger powder particles, and thus result in a higher density of the compact.

The density after compaction, called **green density**, depends primarily on the (a) compaction pressure; (b) powder composition; and (c) hardness of the powder (Fig. 17.11a). The higher the compacting pressure and the softer the powder, the higher is the green density. The density and its uniformity within a compact can be improved with the addition of a small quantity of *admixed* (blended-in) lubricant.

The effect of particle shape on green density can best be understood by considering two powder grades with the same chemical composition and hardness: one with a spherical particle and the other with an irregular shape. The spherical grade will have a higher apparent density (*fill density*), but after compaction under higher pressure, compacts from both grades will have similar green densities. When comparing two similar powders that were pressed under some standard conditions, the powder that gives a higher green density is said to have a higher *compressibility*.

The higher the density of the compacted part, the higher are its strength and elastic modulus (Fig. 17.11b). The reason is that with higher density, the compact has fewer and smaller pores in the same volume, and hence higher strength. Because of friction between (a) the metal particles in the powder and (b) the punch surfaces and die walls, the density within the part can vary considerably. This variation can be minimized by proper punch and die design and by control of friction. Thus, it may be necessary to use multiple punches, each with separate movements, in order to ensure that the density is more uniform throughout the part (Fig. 17.12). Recall a similar discussion regarding the compaction of sand in mold

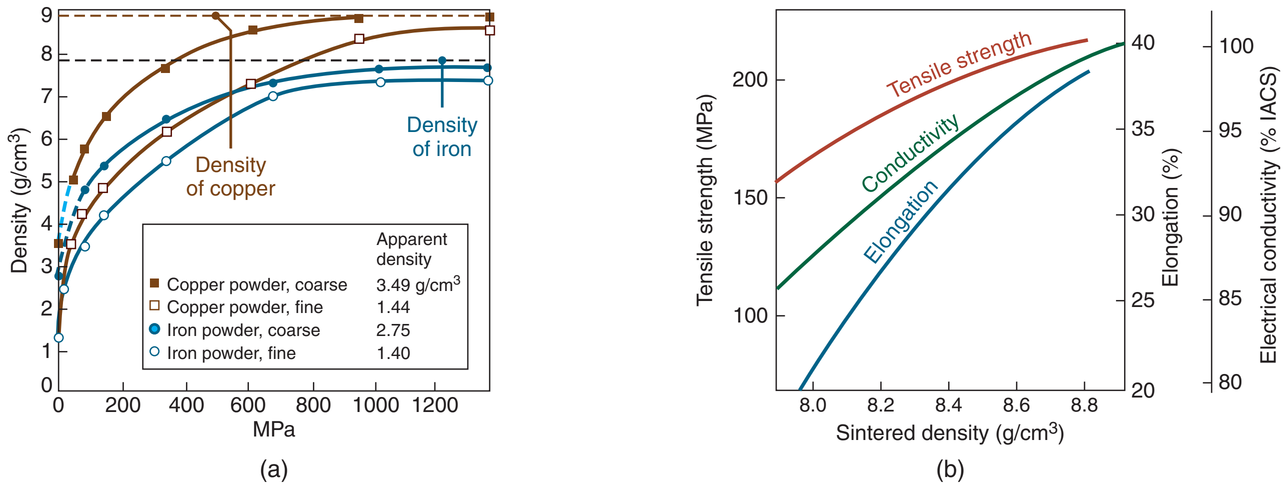


Figure 17.11: (a) Density of copper- and iron-powder compacts as a function of compacting pressure. Density greatly influences the mechanical and physical properties of PM parts. (b) Effect of density on tensile strength, elongation, and electrical conductivity of copper powder. *Source:* (a) After F.V. Lenel, (b) After the International Annealed Copper Standard (IACS) for electrical conductivity.

making (see Fig. 11.7). On the other hand, in some compacted parts, such as gears and cams, density variations may be desirable. For example, densities can be increased in critical locations where high strength and wear resistance are important.

Pressure distribution during compaction. As can be seen in Fig. 17.12, the pressure during compaction decays rapidly away from tooling surfaces. The pressure distribution along the length of the compact in a single action press can be determined to be

$$p_x = p_o e^{-4\mu kx/D}, \tag{17.1}$$

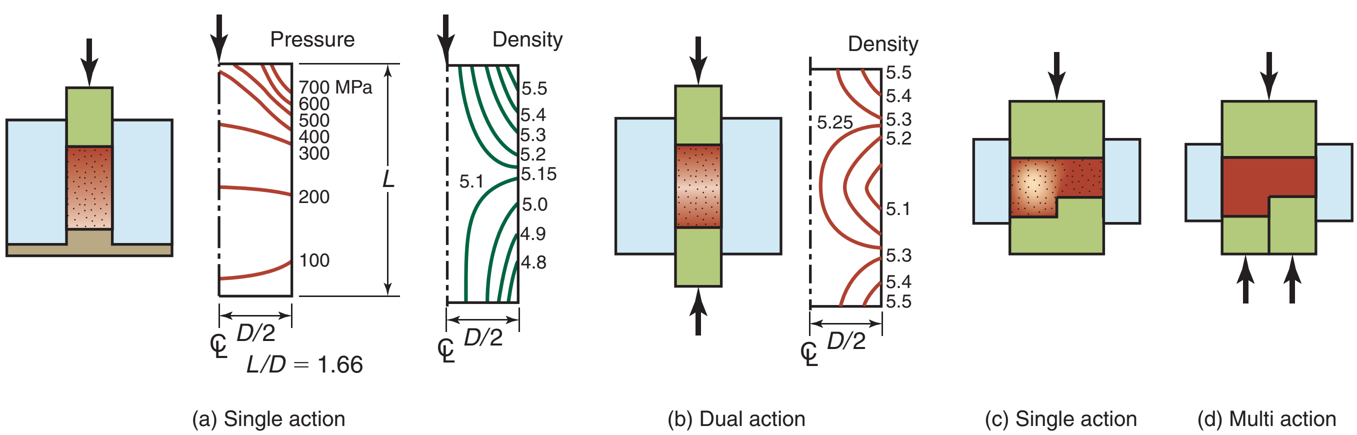


Figure 17.12: Density variation in compacting metal powders in various dies. *Source:* After P. Duwez and L. Zwell.

Table 17.1: Compacting Pressures for Various Powders.

Material	Pressure (MPa)
Metals	
Aluminum	70–275
Brass	400–700
Bronze	200–275
Iron	350–800
Tantalum	70–140
Tungsten	70–140
Other materials	
Aluminum oxide	110–140
Carbon	140–165
Cemented carbides	140–400
Ferrites	110–165

where μ is the coefficient of friction between particles and the container wall, D is the compact diameter, and p is the pressure in the compacting direction, x. Note that the pressure on the bottom of the punch is p_o. Equation (17.1) also includes a variable to account for friction between particles, k, where

$$\sigma_r = kp_x$$

and σ_r is the stress in the radial direction. If there is no friction between the particles, $k = 1$, the powder behaves like a fluid, and thus $\sigma_r = p_x$, signifying a state of hydrostatic pressure. If there is very high friction, $k = 0$, and the pressure will be low near the punch. It can be seen from Eq. (17.1) that the pressure within the compact decays as the coefficient of friction, the parameter k, and the length-to-diameter ratio increase. The pressure required for pressing metal powders typically ranges from 70 MPa for aluminum to 800 MPa for high-density iron parts (see Table 17.1).

17.3.1 Equipment

Press capacities for powder metallurgy are generally around 1.8 to 2.7 MN (180 to 270 metric tons), although presses with much higher capacities are used for special applications. Most applications actually require less than 1 MN (90 metric tons). For small tonnage, crank- or eccentric-type mechanical presses are used; for higher capacities, toggle or knuckle-joint presses are employed (see Fig. 14.19b). Hydraulic presses (Fig. 17.13) with capacities as high as 45 MN (4500 metric tons) can be used for large parts.

Press selection depends on part size and its configuration, density requirements, and production rate. However, the higher the pressing speed, the greater is the tendency for the press to trap air in the die cavity, and thus prevent proper compaction.

17.3.2 Isostatic Pressing

Green compacts may subsequently be subjected to *hydrostatic pressure* in order to achieve more uniform compaction and density. Typical applications include automotive cylinder liners and high-quality parts, such as turbine shafts, oil pipeline component and pump manifolds, valves, and bearings.

In **cold isostatic pressing** (CIP), the metal powder is placed in a flexible rubber mold (Fig. 17.14), typically made of neoprene rubber, urethane, or polyvinyl chloride (Section 7.9). The assembly is then pressurized hydrostatically in a chamber, usually using water. The most common pressure is 400 MPa, although pressures of up to 1000 MPa may be used. The ranges for CIP and other compacting methods in terms of the size and complexity of a part are shown in Fig. 17.15.

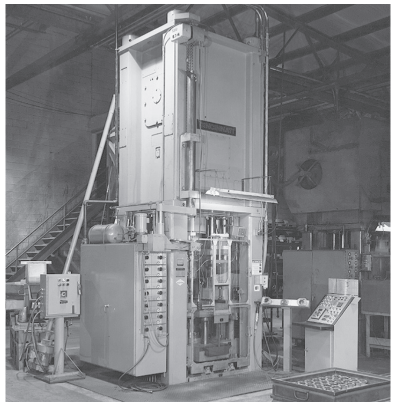


Figure 17.13: A 7.3-MN mechanical press for compacting metal powder. *Source:* Courtesy of Cincinnati Incorporated.

In **hot isostatic pressing** (HIP), the container is typically made of a high-melting-point sheet metal, generally of mild or stainless steel, and the pressurizing medium is high-temperature inert gas or a vitreous (glasslike) fluid (Fig. 17.16). Typical pressures are as high as 100 MPa, although they can be three times higher with temperatures up to 1200°C. The main advantage of HIP is its ability to produce compacts having almost 100% density, good metallurgical bonding of the particles, and good mechanical properties.

The HIP process is used mainly to produce superalloy components for the aircraft and aerospace industries and in military, medical, and chemical applications. It also is used (a) to close internal porosity, especially for parts subjected to fatigue or wear environments; (b) to improve properties in superalloy and titanium-alloy castings for the aerospace industry; and (c) as a final densification step for tungsten-carbide cutting tools and PM tool steels (Chapter 22). Hot isostatic pressing is also applied to ceramics and additive manufactured parts where strength and wear resistance are important.

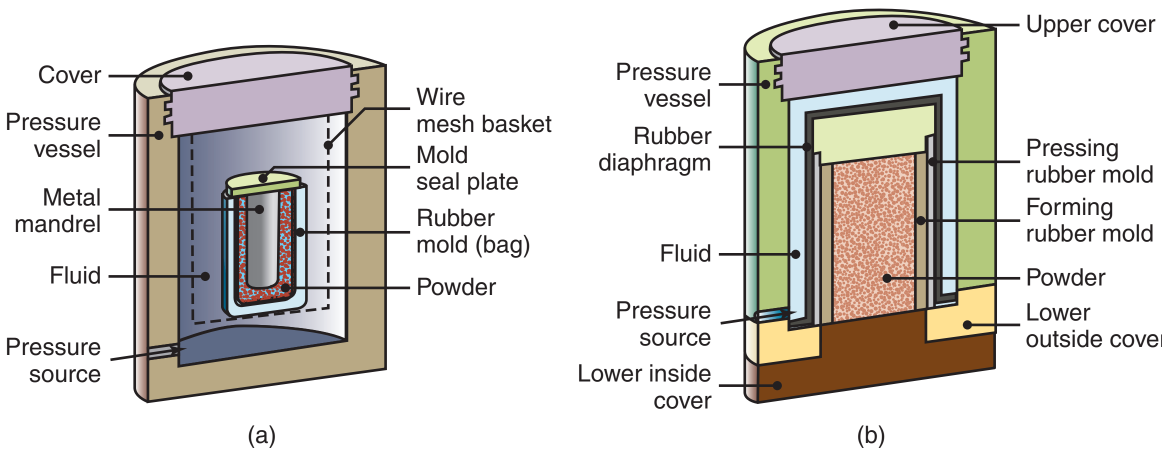


Figure 17.14: Schematic diagram of cold isostatic pressing. Pressure is applied isostatically inside a high-pressure chamber. (a) The wet bag process to form a cup-shaped part; the powder is enclosed in a flexible container around a solid-core mandrel. (b) The dry bag process used to form a PM cylinder.

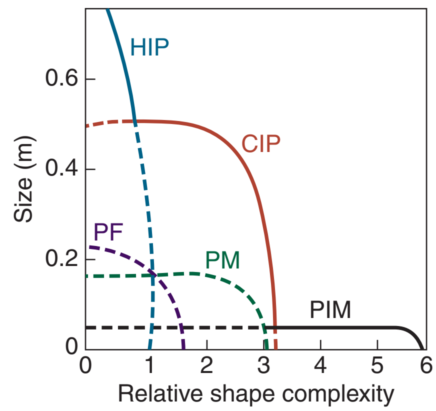


Figure 17.15: Capabilities, with respect to part size and shape complexity, available in various PM operations. PF = powder forging. *Source:* Courtesy of the Metal Powder Industries Federation.

The main advantages of hot isostatic pressing over conventional PM are as follow:

- Because of the uniformity of pressure from all directions, the absence of die-wall friction, and long processing times at elevated temperature (causing metal powders to creep; see Section 2.8), it produces fully-dense compacts of practically uniform grain structure and density, irrespective of part shape; thus the properties are *isotropic*. Parts with high length-to-diameter ratios have been produced, with very uniform density, strength, toughness, and good surface details.
- HIP is capable of handling much larger parts than those in other compacting processes. On the other hand, HIP limitations are:
  - Dimensional tolerances are higher than those in other compacting methods.
  - Equipment costs are higher and production time is longer than those in other processes.
- HIP is applicable only to relatively small production quantities, typically less than 10,000 parts per year.

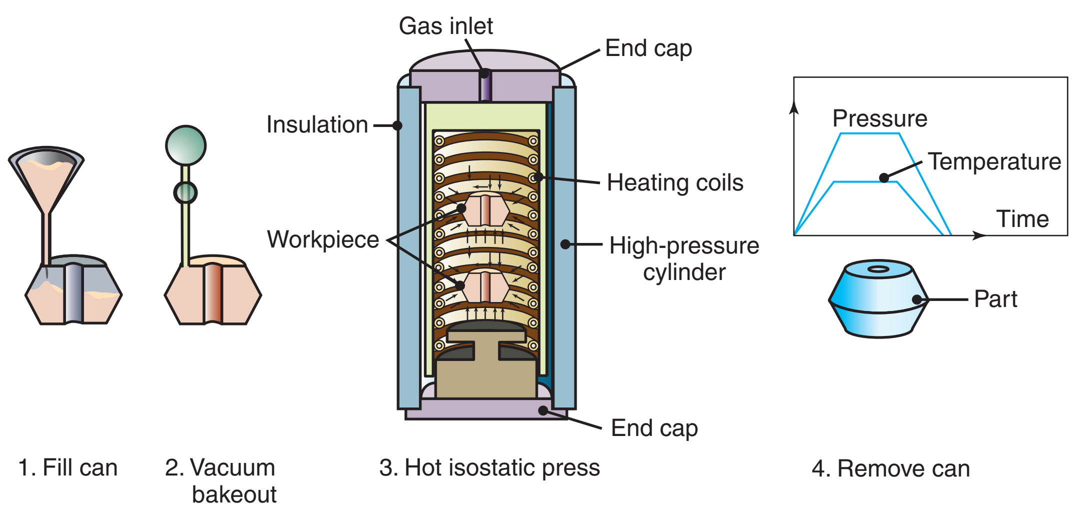


Figure 17.16: Schematic illustration of hot isostatic pressing. The pressure and temperature variation versus time are shown in the diagram.

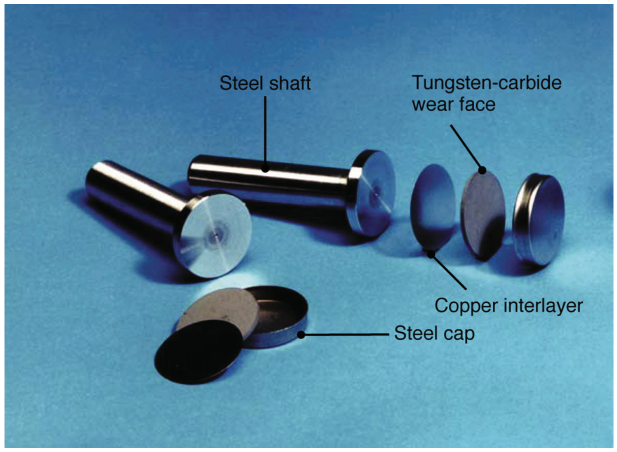


Figure 17.17: A valve lifter for heavy-duty diesel engines produced from a hot-isostatic-pressed carbide cap on a steel shaft. *Source:* Courtesy of the Metal Powder Industries Federation.

Case Study 17.1 Hot Isostatic Pressing of a Valve Lifter

An HIP-clad valve lifter, used in a full range of medium- to heavy-duty truck diesel engines, is shown in Fig. 17.17. The 0.2-kg valve lifter rides on the camshaft and opens and closes the engine valves. Consequently, it is desirable to have a tungsten-carbide (WC) face for wear resistance, and a steel shaft for fatigue resistance. Before the HIP valve lifter was developed, these parts were produced by furnace brazing (Section 32.2), but they resulted in occasional field failures and relatively high scrap rates. Because the required annual production of these parts is over 400,000, high scrap rates are particularly undesirable.

The new part consists of a (a) 9% Co-bonded tungsten-carbide face made from pressed and sintered powder, (b) steel sheet-metal cap fitted over the WC disk, (c) copper-alloy foil interlayer, and (d) steel shaft. The steel cap is electron-beam welded to the steel shaft; then the assembly is hot isostatically pressed to provide a very strong bond. HIP takes place at 1010°C and at a pressure of 100 MPa. The tungsten-carbide surface has a density of 14.52 to 14.72 g/cm^3, a hardness of 90.8 ± 5 HRA, and a minimum transverse rupture strength of 2450 MPa.

Secondary operations are limited to grinding the face to remove any protruding sheet-metal cap and to expose the wear-resistant tungsten-carbide face. The high reliability of the HIP bond greatly reduced scrap rates to under 0.2%. No field failures have been experienced in over four years of full production.

Source: Courtesy of the Metal Powder Industries Federation.

17.3.3 Miscellaneous Compacting and Shaping Processes

Powder-injection Molding. Also called **metal-injection molding** (MIM), in *powder-injection molding* (PIM), very fine metal powders ($<10\ \mu$m) are blended with a 25 to 45% polymer or a wax-based binder. The mixture then undergoes a process similar to die casting (Section 11.4.5; see also *injection molding of plastics* in Section 19.3), where it is injected into the mold at a temperature of 135° to 200°C. Parts generally have sprues and runners, as with injected molded parts (Fig. 17.18), and hence they are carefully separated before additional processing. The molded green parts are placed in a low-temperature oven to burn off the plastic (*debinding*), or the binder is removed by solvent extraction. Often, a small amount of binder may be retained to provide

Figure 17.18: A single shot of metal injection molded components, with sprue, runners, and gates (see also Fig. 19.10). *Source:* Courtesy HARBEC, Inc.

sufficient green strength for transfer of parts to a sintering furnace at temperatures as high as 1375°C. Subsequent operations, such as hole tapping, metal infiltration, and heat treating, also may be performed as required.

Generally, metals suitable for powder-injection molding are those that melt at temperatures above 1000°C, such as carbon and stainless steels, tool steels, copper, bronze, and titanium. Typical parts made are components for small-caliber gun barrels, scope rings for rifles, door hinges, impellers for sprinkler systems, and surgical knives.

The major advantages of powder-injection molding over conventional compaction are:

- Complex shapes, with wall thicknesses as small as 5 mm, can be molded then removed easily from the dies.
- Mechanical properties are nearly the same as those for wrought parts.
- Dimensional tolerances are good.
- High production rates can be achieved by using multicavity dies (see Figs. 11.20 and 19.10).
- Parts produced by the PIM process compete well against small investment-cast parts and forgings, and complex machined parts. However, the PIM process does not compete well with zinc and aluminum die casting (Section 11.4.5), or with screw machining (Section 23.3.4).

The major limitations of PIM are the high cost for small production runs and the need for fine metal powders.

An example where the advantages of metal injection molding are apparent is in the production of light-duty gears, such as office equipment, where load and power is low. An inexpensive gear can be produced directly from metal injection molding, instead of first producing a blank, such as from casting or forging, followed by costly machining and finishing operations (Section 24.7). Avoiding the high machining costs thus results in significant savings; however, this approach may not be suitable for more demanding applications, such as automobile transmissions or in gear pumps.

Forging. In *powder forging* (PF), the part produced from compaction and sintering serves as the *preform* in a hot-forging operation. The forged products are almost fully dense, and have good surface finish, good dimensional tolerances, and uniform and fine grain size. The superior properties obtained make powder forging particularly suitable for highly stressed parts, such as automotive connecting rods and jet-engine components.

Rolling. In *powder rolling*, also called *roll compaction*, the metal powder is fed directly into the roll gap in a two-high rolling mill (Fig. 17.19), and is compacted into a continuous strip at speeds up to 0.5 m/s. The rolling operation can be carried out at room or elevated temperatures. Sheet metal for electrical and electronic components and for coins can be made by this process.

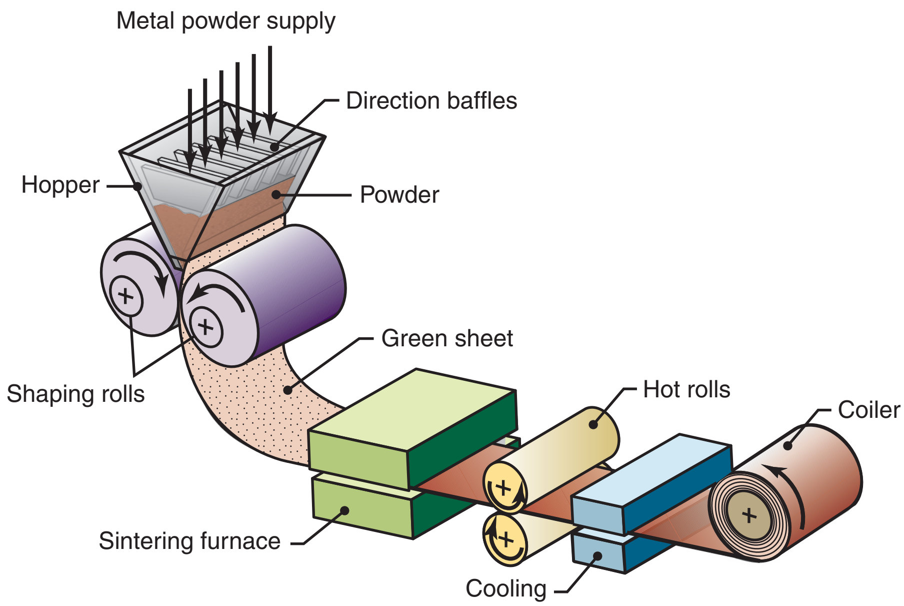


Figure 17.19: An illustration of powder rolling.

Extrusion. Powders can be compacted by *extrusion*, where the powder is encased in a metal container and hot extruded. After sintering, the parts may be reheated and forged in a closed die (Section 14.3) to their final shape. Superalloy powders, for example, are hot extruded for enhanced properties.

Pressureless Compaction. In this operation, the die is gravity filled with metal powder, and the powder is then sintered directly in the die. Because of the resulting low density, pressureless compaction is used principally for porous metal parts such as filters.

Spray Deposition. This is a shape-generation process (Fig. 17.20), involving (a) an atomizer, (b) a spray chamber with an inert atmosphere, and (c) a mold for producing preforms. The mold may be made in various shapes, such as billets, tubes, disks, and cylinders.

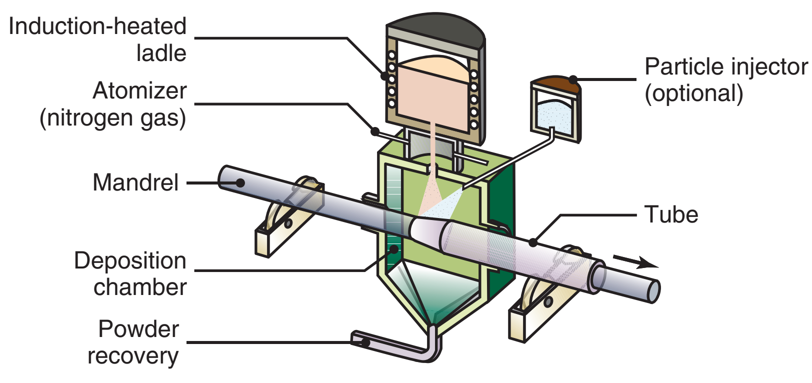


Figure 17.20: Spray deposition (*Osprey process*) in which molten metal is sprayed over a rotating mandrel to produce seamless tubing and pipe.

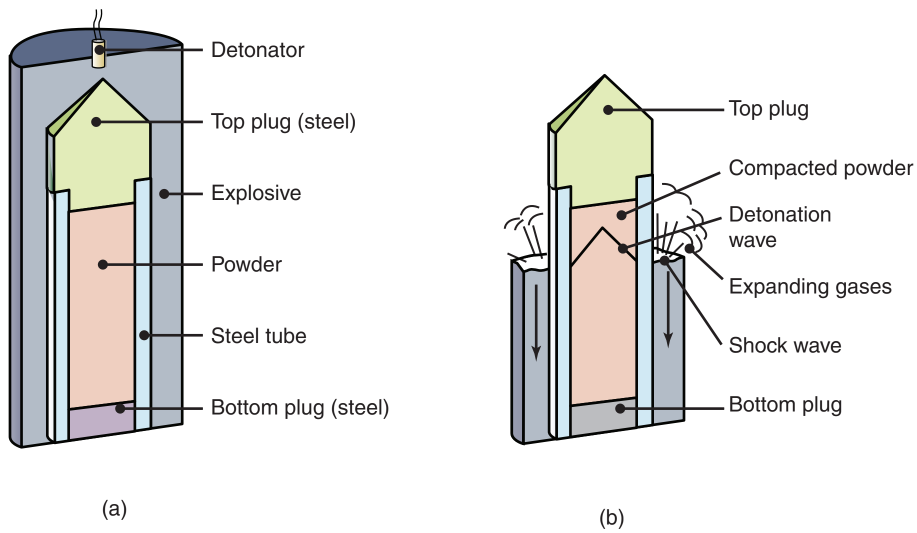


Figure 17.21: Schematic illustration of explosive compaction. (a) A tube filled with powder is surrounded by explosive media inside a container typically made of cardboard or wood. (b) After detonation, a compression wave follows the detonation wave, resulting in a compacted powder material.

Although there are several variations of this process, the best known is the *Osprey process*, shown in Fig. 17.20. After the metal is atomized, it is deposited onto a cooled preform mold, typically made of copper or ceramic, where it solidifies. The metal particles bond together, developing a density that usually is above 99% of the solid-metal density. Spray-deposited forms may subsequently be subjected to additional shaping and consolidation processes, such as forging, rolling, or extrusion. The grain size of the part made is fine, and its mechanical properties are comparable to those of wrought products made of the same alloy.

Dynamic and Explosive Compaction. Some metal powders that are difficult to compact with sufficiently high green strength can be compacted rapidly to near full density, using the setup shown in Fig. 17.21. The explosive drives a mass into green powder at high velocities, generating a shock wave that develops pressures up to 30 GPa. The shock wave traverses across the powder metal part at speeds up to 6 km/s. The powder is often preheated to prevent fracture during compaction.

Combustion Synthesis. *Combustion synthesis* takes advantage of the highly combustible nature of metal powders, by placing a lightly compacted powder into a pressure vessel. An ignition source is then introduced, such as an arc from a tungsten electrode, igniting the powder. The explosion produces a shock wave that travels across the powder, developing heat and pressure that is sufficient for compaction.

Pseudo-isostatic Pressing. In *pseudo-isostatic pressing*, a preform is preheated, surrounded by hot ceramic or graphite granules, and placed in a container. A mechanical press compacts the granules and the preform. Note that the granules are sufficiently large so that they cannot penetrate the pores of the PM part itself. Compaction is uniaxial, but because of the presence of the granules, the loading on the preform is multiaxial. This process has cycle times shorter than HIP, but because the pressure is not strictly hydrostatic, dimensional changes during compaction are not uniform.

Powder Bed Processing. Some PM parts can be produced by powder bed processing, a class of *additive manufacturing* described in detail in Section 20.6.

17.3.4 Punch and Die Materials

The selection of punch and die materials for powder metallurgy depends on the abrasiveness of the powder metal and the number of parts to be produced. Most common die materials are air- or oil-hardening tool steels, such as D2 or D3, with a hardness range from 60 to 64 HRC (Table 5.8). Because of their higher hardness and wear resistance, tungsten-carbide dies are used for more severe applications. Punches generally are made of similar materials.

Close control of die and punch dimensions is essential for die life and proper compaction. Too large a clearance between the punch and the die will allow the metal powder to penetrate the gap, where it will severely interfere with the operation and cause eccentricity in parts made. Diametral clearances generally are less than 25 μm. Die and punch surfaces must be lapped or polished, and in the direction of tool movements in the die, for improved die life and overall performance.

17.4 Sintering

As described in Section 17.3, the green compact is brittle and its *green strength* is low. *Sintering* is the process whereby green compacts are heated, in a controlled-atmosphere furnace, to a temperature below the melting point of the metal, but sufficiently high to allow bonding (fusion) of the individual particles to impart strength to the part. The nature and strength of the bond between the particles involve the complex mechanisms of diffusion, plastic flow, evaporation of volatile materials in the compact, recrystallization, grain growth, and extent of pore shrinkage.

The principal variables in sintering are temperature, time, and furnace atmosphere. Temperatures (Table 17.2) are generally within 70 to 90% of the melting point of the metal or alloy (see Table 3.1). Sintering times (Table 17.2) range from a minimum of about 10 minutes for iron and copper alloys to as much as eight hours for tungsten and tantalum.

Continuous-sintering furnaces, used for most production, have three chambers:

1. *Burn-off chamber*, for volatilizing the lubricants in the green compact, so as to improve bond strength and prevent cracking.

2. *High-temperature chamber*, for sintering.

3. *Cooling chamber*.

For optimum properties, proper control of the furnace atmosphere is essential for successful sintering. An oxygen-free atmosphere is necessary to control the carburization and decarburization of iron and iron-based compacts, and to prevent oxidation of the powders. A vacuum is generally used for sintering refractory-metal alloys and stainless steels. The gases most commonly used for sintering are hydrogen, dissociated or burned ammonia, partially combusted hydrocarbon gases, and nitrogen.

Table 17.2: Sintering Temperature and Time for Various Metals.

Material	Temperature (°C)	Time (min)
Copper, brass, and bronze	760–900	10–45
Iron and iron–graphite	1000–1150	8–45
Nickel	1000–1150	30–45
Stainless steels	1100–1290	30–60
Alnico alloys (for permanent magnets)	1200–1300	120–150
Ferrites	1200–1500	10–600
Tungsten carbide	1430–1500	20–30
Molybdenum	2050	120
Tungsten	2350	480
Tantalum	2400	480

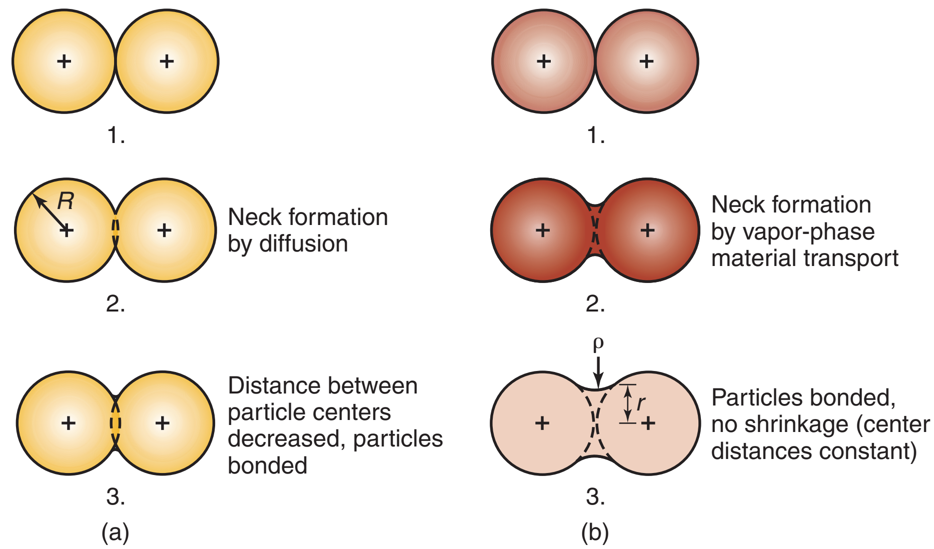


Figure 17.22: Schematic illustration of two mechanisms for sintering metal powders: (a) solid-state material transport; and (b) vapor-phase material transport. R is the particle radius, r is the neck radius, and ρ is the neck-profile radius.

Sintering mechanisms depend on the composition of the metal particles, as well as on processing parameters. The mechanisms are *diffusion*, *vapor-phase transport*, and *liquid-phase sintering*. As the temperature increases, two adjacent powder particles begin to form a bond by a **diffusion mechanism** (*solid-state bonding*, Fig. 17.22a). As a result, the strength, density, ductility, and thermal and electrical conductivities of the compact increase. At the same time, the compact shrinks, thus allowances must be made for shrinkage, as are done in casting.

A second sintering mechanism is **vapor-phase transport** (Fig. 17.22b). Because the material is heated to close to its melting temperature, metal atoms are released to the vapor phase from the particles. At convergent geometries (the interface of two particles), the melting temperature is locally higher, and the vapor phase resolidifies; thus, the interface grows and strengthens while each particle shrinks as a whole.

If two adjacent particles are of different metals, *alloying* can take place at the interface of the two particles. If one of the particles has a lower melting point than the other, the particle will melt and will surround the particle that has not melted because of surface tension (Fig. 17.23). An example of this mechanism, known as *liquid-phase sintering*, is cobalt in tungsten-carbide tools and dies (see Section 22.4), and the parts made are denser and stronger.

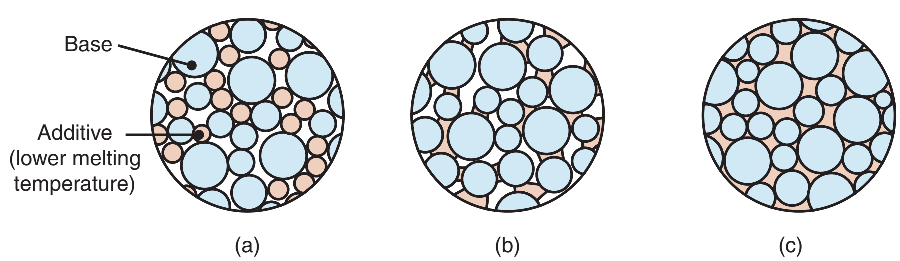


Figure 17.23: Schematic illustration of liquid phase sintering using a mixture of two powders. (a) Green compact of a higher melting point base metal and lower temperature additive; (b) liquid melting, wetting and reprecipitation on surfaces; (c) fully sintered solid material.

In *spark sintering*, loose metal powders are placed in a graphite mold, then heated by electric current, subjected to a high-energy discharge and compacted, all in one step. Another technique is microwave sintering, which reduces sintering time and thereby prevents grain growth, which can adversely affect strength.

Mechanical Properties. Depending on temperature, time, and processing history, different structures and porosities can be obtained in a sintered compact, thus affecting its properties. Porosity cannot be completely eliminated because (a) some voids remain after compaction and (b) gases evolve during sintering. Porosity may consist either of a *network* of interconnected pores or of *closed holes*. Generally, if the density of the part is less than 80% of its bulk density, the pores are interconnected. Although porosity reduces the strength of the PM product, it is an important characteristic for making metal filters and bearings, and to also allow for infiltration with liquid lubricants by surface tension.

Typical mechanical properties for several sintered PM alloys are given in Table 17.3. The differences in mechanical properties of wrought versus PM metals are given in Table 17.4. To further evaluate the differences between the properties of PM, wrought, and cast metals and alloys, compare these tables with the ones given in Parts I and II.

The effects of various manufacturing processes on the mechanical properties of a titanium alloy are shown in Table 17.5. Note that hot isostatic pressed (HIP) titanium has properties that are similar to those for cast and forged titanium. It should be noted, however, that unless they are precision forged, forgings generally require some additional machining or finishing operations whereas a PM component may not.

17.5 Secondary and Finishing Operations

In order to further improve the properties of sintered PM products or to impart special characteristics, several additional operations may be carried out following sintering:

1. **Coining** and **sizing** are compacting operations, performed in presses under high pressure. The purposes of these operations are to further density and impart better dimensional accuracy to the sintered part and to improve its strength and surface finish.

2. *Preformed and sintered* alloy-powder compacts subsequently may be cold or hot **forged** or *impact forged* to the desired final shapes. The parts made have good surface finish and dimensional tolerances, and uniform and fine grain size. The superior properties obtained make this technology particularly suitable for such applications as highly stressed automotive and jet-engine components.

3. Powder-metal parts also may be subjected to other finishing operations, such as:

    - **Machining**, for producing various geometric features by milling, drilling, and tapping.
    - **Grinding**, for improving dimensional accuracy and surface finish.
    - **Heat treating**, for increasing hardness and strength.

4. The inherent porosity of PM components can be utilized as a design advantage by **impregnating** them with a fluid. Bearings and bushings that are lubricated internally with up to 30% oil by volume are made by immersing the sintered bearing in heated oil. In service, an interruption in lubricant supply or a temporary high load or speed will cause the bearing to increase in temperature. Since the liquid oil has a higher thermal expansion than the solid bearing, it will exude or percolate from the surface, leading to a self-lubricated condition.

5. **Infiltration** is a process whereby, for example, a slug of a lower-melting-point metal is placed in contact with the sintered part; the assembly is then heated to a temperature sufficiently high to melt the slug (see Fig. 20.15). The molten metal infiltrates the pores by capillary action, producing a relatively pore-free part having good density and strength. The most common application is the infiltration

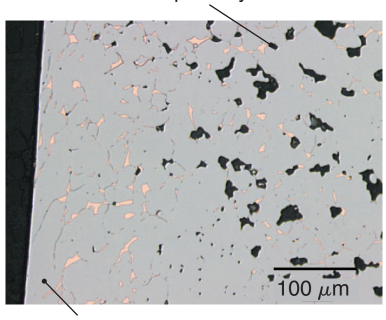


Figure 17.24: Micrograph of a PM material surface after roll densification. Note the low porosity near the surface, increasing the material's ability to support contact stresses and resist fatigue. *Source:* Courtesy of Capstan Atlantic Corp.

of iron-based compacts by copper or bronze. The hardness and tensile strength of the part are thus improved and the pores are filled, preventing moisture penetration, which could cause corrosion. Moreover, because some porosity is essential for infiltration, the part may be sintered only partially, resulting in lower thermal warpage.

6. **Electroplating** (Section 34.9) can be applied to PM parts, but special care is required to ensure that the electrolytic fluid is thoroughly removed since it presents health hazards. Under certain conditions, electroplating also can seal a part and eliminate its permeability.

7. **Densification**, or **roll densification**, is similar to roller burnishing (Section 34.2), where a small-diameter hard roll is pressed against a PM part, resulting in sufficiently high contact pressures to cause plastic deformation of its surface layers. Thus, instead of cold working the part, the effect is to cause an increase in density, or densification, of the part's surface layers (Fig. 17.24). PM gears and bearing races are generally subjected to roll densification; the surface layer becomes more fatigue resistant and is better able to support higher contact stresses than untreated components.

Case Study 17.2 Production of Tungsten Carbide for Tools and Dies

Tungsten carbide is an important tool and die material, mainly because of its hardness, strength, and wear resistance over a wide range of temperatures (see Section 22.4.1); it is made by PM techniques. First, powders of tungsten and carbon are blended together in a ball mill or a rotating mixer. The mixture (typically 94% tungsten and 6% carbon, by weight) is heated to approximately 1500°C in a vacuum-induction furnace; as a result, the tungsten is carburized, forming tungsten carbide in a fine powder form. A binding agent (usually cobalt) is then added to the tungsten carbide (together with an organic fluid, such as hexane), and the mixture is ball milled to produce a uniform and homogeneous mix. The process can take several hours, or even days.

The mixture is then dried and consolidated, usually by cold compaction, at pressures in the range of 200 MPa. Finally, the compact is sintered in a hydrogen atmosphere or a vacuum furnace, at a temperature of 1350–1600°C, depending on its composition. Powders may also be hot pressed at the sintering temperature, using graphite dies. At this temperature, the cobalt is in a liquid phase and acts as a binder for the carbide particles. During sintering, the tungsten carbide undergoes a linear shrinkage of about 16%, corresponding to a volume shrinkage of about 40%; thus, control of size and shape is important for producing tools with accurate dimensions. A combination of other carbides, such as titanium carbide and tantalum carbide, can likewise be produced, using mixtures made by the methods described in this example.

Table 17.3: Mechanical Properties of Selected PM Materials.

Material	Yield strength (MPa)	Ultimate tensile strength (MPa)	Elastic modulus (GPa)	Hardness	Elongation in 25 mm (%)	Density (g/cm^3)	Notes
Ferrous							
F-0008-20	170	200	85	35 HRB	< 1	5.8	F-008 is often most
F-0008-35	260	390	140	70 HRB	1	7.0	cost effective.
F-0008-55HT	—	450	115	22 HRC	< 1	6.3	
F-0008-85HT	—	660	150	35 HRC	< 1	7.1	
FC-0008-30	240	240	85	50 HRB	< 1	5.8	Copper added for strength,
FC-0008-60	450	520	155	84 HRB	< 1	7.2	hardness, and wear
FC-0008-95		720	150	43 HRC	< 1	7.1	resistance.
FN-0205-20	170	280	115	44 HRB	1	6.6	Good heat treated strength,
FN-0205-35	280	480	170	78 HRB	5	7.4	impact energy
FN-0205-180HT		1280	170	78 HRB	< 1	7.4	
FX-1005-40	340	530	160	82 HRB	4	7.3	Copper infiltrated steel
FX-1005-110HT		830	160	38 HRC	< 1	7.3	Copper infiltrated steel
Stainless Steels							
SS-303N1-38	310	470	115	70 HRB	5	6.9	Good machinability
SS-304N1-30	260	300	105	61 HRB	< 1	6.4	High corrosion resistance
SS-316N1-25	230	280	105	59 HRB	< 1	6.4	Good general-purpose alloy
SS-316N2-38	310	480	140	65 HRB	131	6.9	
Copper and Copper Alloys							
CZ-1000-9	70	120	80	65 HRH	9	7.6	General purpose structural parts
CZ-1000-11	80	160	100	80 HRH	12	8.1	General purpose structural parts
CZP-3002-14	110	220	90	88 HRH	16	8.0	High strength structural parts
CT-1000-13	110	150	60	82 HRH	4	7.2	Common self-lubricated bearing material
Aluminum Alloys							
Ax 123-T1	200	270	—	47 HRB	3	2.7	General purpose
Ax 123-T6	390	400	—	72 HRB	< 1	2.7	structural parts
Ax 231-T6	200	220	—	55 HRB	1	2.7	High wear resistance
Ax 231-T6	310	320	—	77 HRB	< 1	2.7	
Ax 431-T6	270	300	—	55 HRB	5	2.8	High strength structural parts
Ax 431-T6	440	470	—	80 HRB	2	2.8	
Titanium Alloys							
Ti-6Al-4V (HIP)	917	827	—	—	—	13	Most common titanium alloy
Superalloys							
Stellite 19	—	1035	—	—	49 HRC	< 1	

17.6 Design Considerations

Because of the unique properties of metal powders, their flow characteristics in the die, and the brittleness of green compacts, there are certain design principles that should be followed (Figs. 17.25 through 17.27):

1. The shape of the compact must be kept as simple and uniform as possible. Sharp changes in contour, thin sections, variations in thickness, and high length-to-diameter ratios should be avoided.

2. Provision must be made for ejection of the green compact from the die without damaging it. Holes or recesses should be parallel to the axis of punch travel. Chamfers should be provided to avoid damage to the edges during ejection.

Table 17.4: Comparison of Mechanical Properties of Some Wrought and Equivalent PM Metals (as Sintered).

Metal	Condition	Relative density[a] (%)	Ultimate tensile strength[a] (MPa)	Elongation in 50 mm (%)
Aluminum				
2014-T6	Wrought (W)	100	480	20
	PM	94	330	2
6061-T6	W	100	310	15
	PM	94	250	2
Copper, OFHC[b]	W, annealed	100	235	50
	PM	89	160	8
Brass, 260	W, annealed	100	300	65
	PM	89	255	26
Steel, 1025	W, hot rolled	100	590	25
	PM	84	235	2
Stainless steel, 303	W, annealed	100	620	50
	PM	82	360	2

Notes: [a] The density and strength of PM materials greatly increase with further processing, such as forging, isostatic pressing, and heat treatments.
[b] OFHC = oxygen-free, high conductivity.

3. In order to increase tool and die life and reduce production costs, PM parts should be made with the widest acceptable dimensional tolerances, consistent with their intended applications.

4. Part walls generally should not be less than 1.5 mm thick; however, with special care, walls as thin as 0.34 mm can be pressed successfully on components as little as 1 mm in length. Walls with length-to-thickness ratios greater than 8:1 are difficult to press, and density variations are virtually unavoidable.

5. Steps in parts can be produced if they are simple and their size doesn't exceed 15% of the overall part length. Larger steps can be pressed, but they require more complex, multiple-motion tooling.

6. Letters and numbers can be pressed if they are oriented perpendicular to the direction of pressing, and these can be raised or recessed. Raised letters are more susceptible to damage in the green stage, and also may prevent stacking during sintering.

Table 17.5: Mechanical Property Comparisons for Ti-6AL-4V Titanium Alloy.

Process	Relative density (%)	Yield strength (MPa)	Ultimate tensile strength (MPa)	Elongation (%)	Reduction of area (%)
Cast	100	840	930	7	15
Cast and forged	100	875	965	14	40
Blended elemental (P+S)	98	786	875	8	14
Blended elemental (HIP)	> 99	805	875	9	17
Prealloyed (HIP)	100	880	975	14	26
Electron-beam melting	100	910	970	16	—

* P+S = pressed and sintered, HIP = hot isostatically pressed. *Source:* Courtesy of R.M. German and Stratasys, Inc.

7. Flanges or overhangs can be produced by providing a step in the die; however, long flanges can be broken during ejection, thus requiring more elaborate tooling. A long flange should incorporate a draft around the flange, a radius at the bottom edge, and a radius at the juncture of the flange and/or component body, in order to reduce stress concentrations and thus the likelihood of fracture.

8. A true radius cannot be pressed into the edge of a PM part because it would require the punch to be feathered (gently tapered) to a zero thickness, as shown in Fig. 17.26c. Chamfers or flats are preferred for pressing, and a 45° angle in a 0.25-mm flat is a common design practice.

9. Keys, keyways, and holes used for transmitting torques on gears and pulleys can be formed during powder compaction. Bosses (see Fig. 10.16) also can be produced, provided that proper drafts are provided, and their length is small compared to the overall component dimensions.

10. Notches and grooves can be made if they are oriented perpendicular to the powder pressing direction; circular grooves should not exceed a depth of 20% of the overall component, and rectangular grooves should not exceed 15%.

11. Parts produced by PIM have design constraints similar to those produced by injection molding of polymers (Section 19.3). With PIM, wall thicknesses should be uniform to minimize distortion during sintering. Also, molds should be designed with smooth shape transitions, to prevent powder accumulation and to allow uniform distribution of metal powder.

12. Dimensional tolerances of sintered PM parts are usually on the order of ± 0.05 to 0.1 mm. Tolerances an be improved significantly with such subsequent operations as sizing, machining, and grinding.

17.7 Economics of Powder Metallurgy

Because powder metallurgy can produce parts at net or near-net shapes, thus eliminating some secondary finishing and assembly operations, it has become increasingly competitive with casting, forging, and machining. On the other hand, the high initial cost of punches, dies, and equipment for PM processing means that production volume must be sufficiently high to warrant this expenditure. Although there are exceptions, the process generally is economical for quantities over 10,000 pieces.

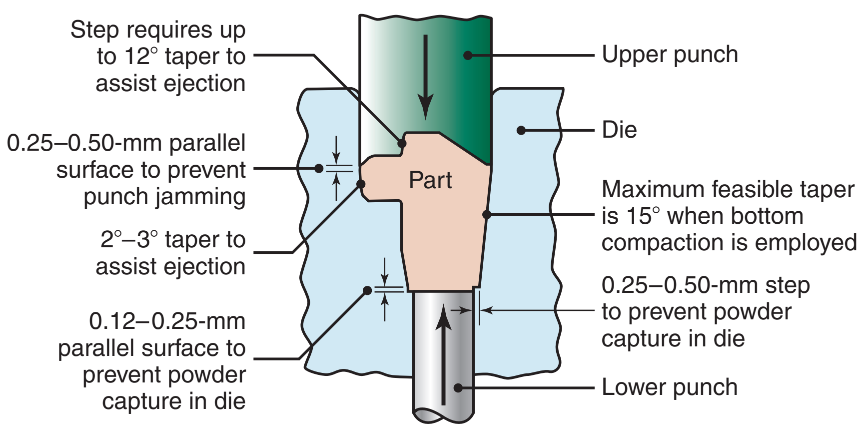


Figure 17.25: Die geometry and design features for powder-metal compaction. *Source:* Courtesy of the Metal Powder Industries Federation.

As in other processes, the cost of dies and tooling in powder metallurgy depends on part complexity and the method of processing the powders. Tooling costs for processes such as hot isostatic pressing and powder-injection molding are thus higher than more conventional powder processing. Because it is a near-net-shape manufacturing method, the cost of finishing operations in PM is low as compared to other processes. However, if there are certain features to the part, such as threaded holes, undercuts, and transverse cavities and holes, the finishing costs will increase. Consequently, following design guidelines in PM to minimize or avoid such additional operations can be more important in this process than in others.

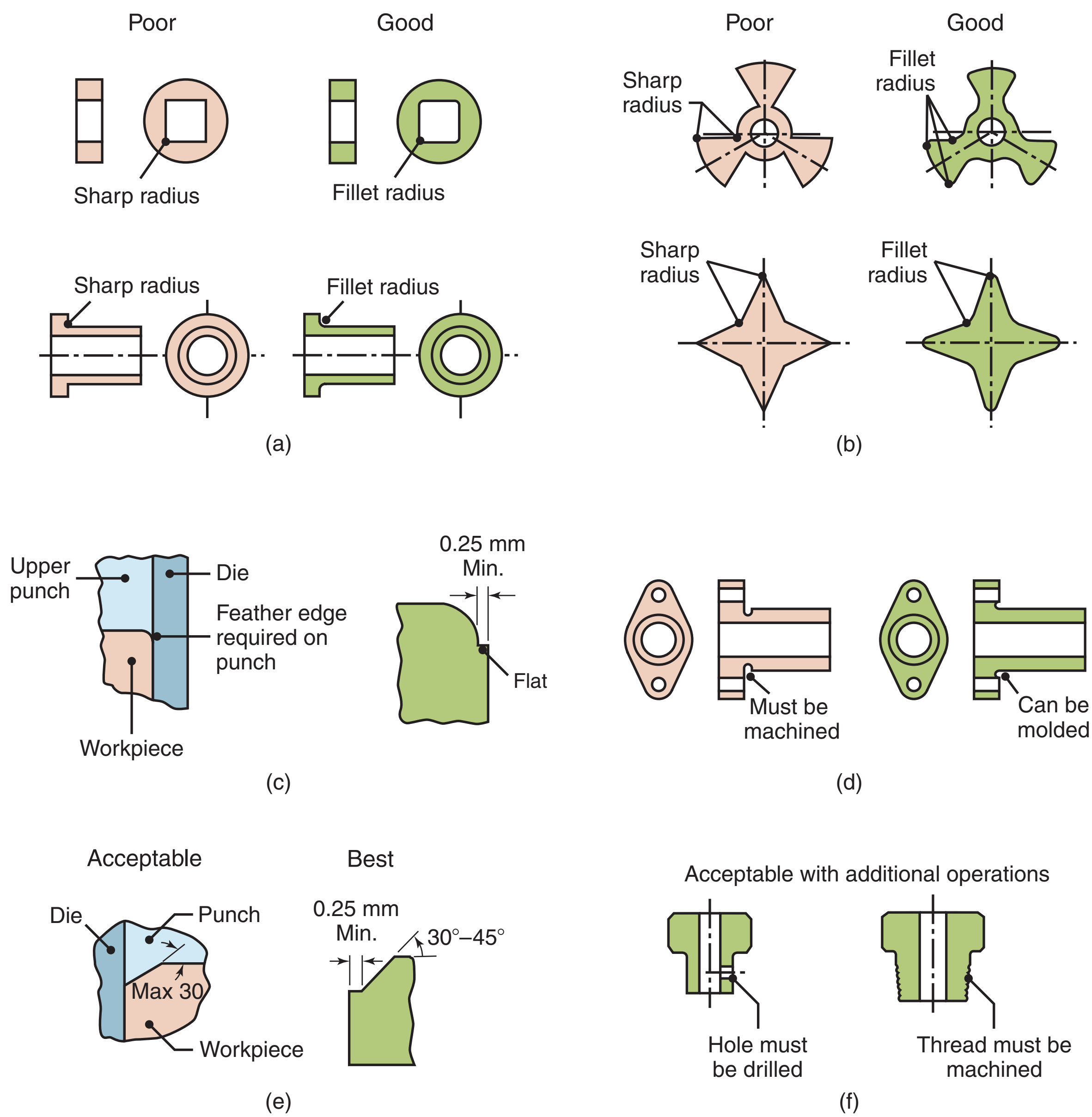


Figure 17.26: Examples of PM parts showing poor and good designs. Note that sharp radii and reentry corners should be avoided, and that threads and transverse holes have to be produced separately by additional machining operations. *Source:* Courtesy of the Metal Powder Industries Federation.

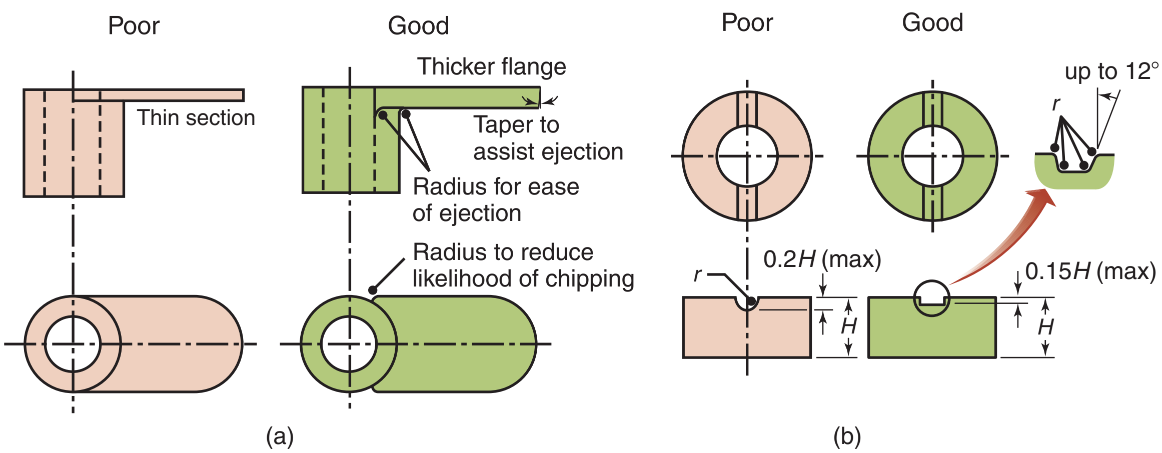


Figure 17.27: (a) Design features to use with unsupported flanges. (b) Design features for use with grooves. *Source:* Courtesy of Metal Powder Industries Federation.

Table 17.6: Forged and PM Titanium Parts and Cost Savings.

	Weight (kg)			
Part	**Forged billet**	**PM**	**Final part**	**Cost savings (%)**
F-14 Fuselage brace	2.8	1.1	0.8	50
F-18 Engine mount support	7.7	2.5	0.5	20
F-18 Arrestor hook support fitting	79.4	25	12.9	25
F-14 Nacelle frame	143	82	24.2	50

Equipment costs for conventional PM processing are somewhat similar to those for bulk deformation processing of metals, such as forging; however, the cost increases significantly when using methods such as HIP and PIM. Although the cost of materials has increased significantly (see Table 6.1), it has actually improved the economic viability of PM, since tooling and equipment costs are a smaller fraction of the total cost of production.

Labor costs for PM are not as high as those in some other processes, primarily because the individual operations, such as powder blending, compaction, and sintering, are performed on highly automated equipment.

The near-net-shape capability of PM significantly reduces or eliminates scrap. Weight comparisons for aircraft components, produced by forging and by PM processes, are shown in Table 17.6. Note that the PM parts are subjected to further machining processes; thus, the final parts weigh less than those made by either of the two processes alone.

Case Study 17.3 Powder Metallurgy Parts in a Snowblower

Some of the parts in the freewheeling steering system of a commercial snowblower are shown in Fig. 17.28. Among the 16 PM components, the sprocket is the largest, at around 140 mm in diameter.

The final assembly incorporates a stamped steel frame, bronze and plastic bearings, and a wrought-steel axle, to produce a highly functional and low-cost machine. Unique features compatible with PM manufacturing were incorporated into the design of these parts to enhance their functionality.

The PM components in the assembly range from single-level parts, with fixed features on punch faces and core rods, to intricate multilevel parts with complex die geometry, core rods, and transfer punches. These are unique features and they manage the powder for local density control. The clutch pawl, for example, is produced to a net-shape peripheral geometry that is not practical or economical with other manufacturing technologies. The material used is FLC4608-70 steel (a prealloyed powder of iron, with 1.9% Ni, 0.56% Mo, and 0.8% C mixed in with 2% Cu), with a tensile strength of 500 MPa and a density of 6.8 g/cm^3.

Part numbers are pressed into the face of the components, as a simple means of identifying them. Two of the components are made with especially close tolerances: The pawl latch gear has a 0.15-mm tolerance on the pitch diameter (PD), with 0.11 mm PD to ID run-out and 0.025 mm tolerance on the bore. The 32-tooth sprocket has a thin-walled 57.75 mm ID with a 0.05-mm tolerance. Both the pawl latch gear and the sprocket acquire a density of 6.7 g/cm^3 and a tensile strength of 690 MPa.

All components shown passed normal life-cycle testing and product-life testing, including shock loading by engaging the drive in reverse, while traveling at maximum forward speed down an incline. Clutch components, which were also subjected to salt-spray corrosion resistance, and proper operation in subzero temperatures, experienced no failures. No machining is required on these parts, as these are sufficiently net-shape components. The only additional operations, prior to final assembly, are vibratory deburring and honing of the 32-tooth sprocket, in order to produce a close-tolerance bore and surface finish. The clutch pawls, produced with sinter-hardened steel, are quenched in an atmosphere so that the porosity present can be filled with a lubricant, to provide lubricity at the interface of mating parts (see also Section 33.6).

Source: Courtesy of the Metal Powder Industries Federation and Burgess-Norton Manufacturing Co.

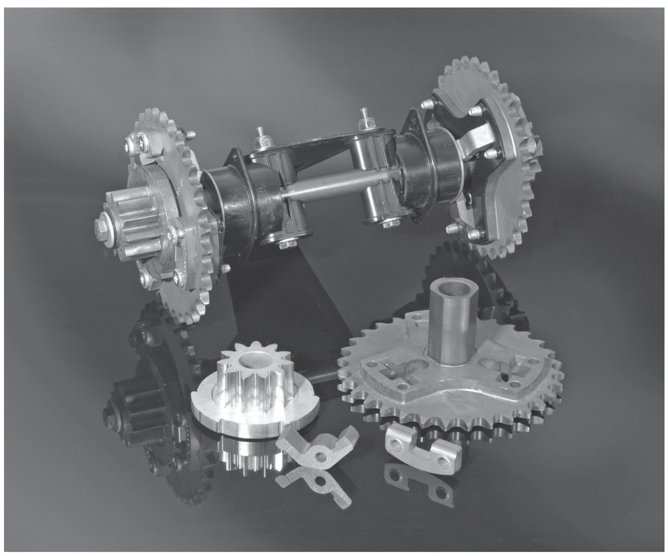


Figure 17.28: Powder metallurgy parts in a commercial snowblower. Courtesy of Metal Powder Industries Federation.

Summary

- Powder metallurgy is a net-shape or near-net shape forming sequence consisting of metal powder production, blending, compaction in dies, and sintering in order to impart strength, hardness, and toughness. Although the size and the weight of PM products are limited, the process is capable of producing relatively complex parts economically, in net-shape form, to close dimensional tolerances, and from a wide variety of metal and alloy powders.
- Secondary and finishing operations may be performed on PM parts to improve their dimensional accuracy, surface finish, mechanical and physical properties, and appearance. These operations include forging, heat treating, machining, grinding, plating, impregnation (as with oil), and infiltration (with lower melting-point metals).
- Control of powder shape and quality, process variables, and sintering atmospheres are important considerations to ensure product quality. Density and mechanical and physical properties can be controlled by tooling design and by controlling the compacting pressure.
- An important PM process is powder-injection molding, which involves mixing very fine metal powders with a polymer; the viscous mixture is then injected into molds to produce parts.
- Sintering is a process whereby a porous metal powder compact or shape is heated to a high temperature in a controlled atmosphere. Sintering fuses the particles, but also leads to shrinkage and potentially warpage.
- Design considerations for powder metallurgy include the shape of the part, the ability to eject the green compact from the die, and the dimensional tolerances that are acceptable for the particular application.
- The PM process is suitable for medium- to high-volume production runs and for relatively small parts.

Key Terms

Atomization
Blending
Carbonyls
Cold isostatic pressing
Comminution
Compaction
Diffusion
Electrolytic deposition
Green compact
Green strength
Hot isostatic pressing
Impregnation
Infiltration
Injection molding
Mechanical alloying
Metal injection molding
Powder injection molding
Powder metallurgy
Pressing
Pressureless compaction
Reduction
Screening
Shape factor
Sintering
Spark sintering

Bibliography

Allen, T., **Powder Sampling and Particle Size Determination**, Elsevier, 2003.
ASM Handbook, Vol. 7: **Powder Metallurgy**, ASM International, 2015.
Brewin, P.R., and Coube, O., **Modelling of Powder Die Compaction**, Elsevier, 2007.
Datta, B.K., **Powder Metallurgy: An Advanced Technique of Processing Metals**, 2nd ed., Prentice-Hall, 2014.
El-Eskandarany, M.S., **Mechanical Alloying**, 2nd ed., William Andrew, 2015.
Douvard, D. (ed.), **Powder Metallurgy**, ISTE Publishing, 2009.
Fayed, M., and Otten, L. (eds.), **Handbook of Powder Science and Technology**, 2nd ed., Chapman & Hall, 1997.
German, R.M., **Sintering: From Empirical Observations to Scientific Principles**, Butterworth-Heinemann, 2014.
German, R.M., **Powder Metallurgy and Particulate Materials Processing**, Metal Powder Industry, 2006.
Kang, S.-J., **Sintering: Densification, Grain Growth and Microstructure**, Butterworth-Heinemann, 2005.
Klar, E., and Samal, P.K., **Powder Metallurgy Stainless Steels: Processing, Microstructures and Properties**, ASM International, 2008.
Leander, F., and West, W.G., **Fundamentals of Powder Metallurgy**, Metal Powder Industries Federation, 2002.
Neikov, O.D., et al. (eds.), **Handbook of Non-Ferrous Metal Powders: Technologies and Applications**, Elsevier, 2009.
Powder Metallurgy Design Manual, 2nd ed., Metal Powder Industries Federation, 1995.
Selesca, S.M., Selesca, M., and Danninger, S., **Machinability of Powder Metallurgy Steels**, Cambridge International Science, 2005.
Upadhyaya, A., and Upadhyaya, G.S., **Powder Metallurgy: Science, Materials, and Technology**, Universities Press, 2012.
Upadhyaya, G.S., **Sintered Metallic and Ceramic Materials: Preparation, Properties and Applications**, Wiley, 2000,

Review Questions

17.1. Describe briefly the production steps involved in making powder-metallurgy parts.

17.2. Name the various methods of powder production and explain the types of powders produced.

17.3. Explain why metal powders may be blended.

17.4. Describe the methods used in metal powder compaction.

17.5. What is isostatic pressing? How is it different from pseudo-isostatic pressing?

17.6. What hazards are involved in PM processing? Explain their causes.

17.7. Describe what occurs to metal powders during sintering.

17.8. Describe the wet-bag and dry-bag techniques.

17.9. Why might secondary and finishing operations be performed on PM parts?

17.10. Explain the difference between impregnation and infiltration. Give some applications of each.

17.11. What is roll densification? Why is it done?

17.12. What is mechanical alloying? What are its advantages over the conventional alloying of metals?

17.13. What is the osprey process?

17.14. What is screening of metal powders? Why is it done?

17.15. Why are protective atmospheres necessary in sintering? What would be the effects on the properties of PM parts if such atmospheres were not used?

Qualitative Problems

17.16. Why is there density variation in the compacting of powders? How is it reduced?

17.17. What is the magnitude of the stresses and forces involved in powder compaction?

17.18. List the similarities and differences of forging and compacting metal powders.

17.19. Give some reasons that powder-injection molding is an important process.

17.20. How does the equipment used for powder compaction vary from those used in other metalworking operations in the preceding chapters?

17.21. Explain why the mechanical and physical properties depend on their density.

17.22. What are the effects of the different shapes and sizes of metal particles in PM processing?

17.23. Describe the relative advantages and limitations of cold and hot isostatic pressing.

17.24. How different, if any, are the requirements for punch and die materials in powder metallurgy from those for forging and extrusion operations? Explain.

17.25. The powder metallurgy process can be competitive with processes such as casting and forging. Explain why this is so.

17.26. What are the reasons for the shapes of the curves shown in Fig. 17.11 and for their relative positions on the charts?

17.27. Should green compacts be brought up to the sintering temperature slowly or rapidly? Explain your reasoning.

17.28. Because they undergo special processing, metal powders are more expensive than the same metals in bulk form, especially powders used in powder-injection molding. How is the additional cost justified in processing powder-metallurgy parts?

17.29. In Fig. 17.12c, it can be seen that the pressure is not uniform across the diameter of the compact at a particular distance from the punch. What is the reason for this variation?

17.30. Why do the compacting pressure and the sintering temperature depend on the type of powder metal?

17.31. What will be stronger: a blend of stainless steel and copper powder that is compacted and sintered, or a stainless steel powder that is compacted, sintered, and infiltrated by copper? Explain.

17.32. Name the various methods of powder production and sketch the morphology of powders produced.

Quantitative Problems

17.33. Estimate the maximum tonnage required to compact a brass slug 150 mm in diameter. Would the height of the slug make any difference in your answer? Explain your reasoning.

17.34. Refer to Fig. 17.11a. What should be the volume of loose, fine iron powder in order to make a solid cylindrical compact 40 mm in diameter and 20 mm high?

17.35. Determine the shape factors for (a) a cylinder with a dimensional ratio of 1:1:1 and (b) a flake with a ratio of 1:10:10.

17.36. Estimate the number of particles in a 400-g sample of iron powder if the particle size is 40 μm.

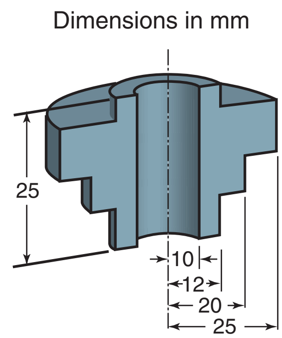
17.37. Assume that the surface of a copper particle is covered by an oxide layer 0.15 mm in thickness. What is the volume (and the percentage of volume) occupied by this layer if the copper particle itself is 60 μm in diameter?

17.38. A coarse copper powder is compacted in a mechanical press at a pressure of 275 MPa. During sintering, the green part shrinks an additional 5%. What is the final density?

17.39. A gear is to be manufactured from iron powders. It is desired that it have a final density 90% that of cast iron, and it is known that the shrinkage in sintering will be approximately 5%. For a gear that is 60 mm in diameter and has a 25 mm hub, what is the required press force?

17.40. What volume of powder is needed to make the gear in Problem 17.39?

17.41. The axisymmetric part shown in the accompanying figure is to be produced from fine copper powder and is to have a tensile strength of 150 MPa. Determine the compacting pressure and the initial volume of powder needed.



17.42. The part considered in Problem 17.41 is to be compressed in a dual-action press. The density of loose powder is 40%, but it is desired to produce a green compact with 80% of full density. Specify the initial positions of the vertical features in the die.

17.43. Coarse iron powder is compacted into a cylinder with a 25 mm diameter and 40 mm height. The green part has a measured mass of 130 g. Calculate (a) the apparent density; (b) the percentage of the theoretical full density; (c) an estimate of the compacting pressure used.

17.44. The part in Problem 17.43 is to be hot isostatically pressed to full density. If shrinkage is the same in all directions, estimate the final part dimensions.

17.45. Fine iron powder is compressed into a cylinder (d=30 mm, h=20 mm), achieving 70% of theoretical density. It is to be hot forged to a height of 5 mm. What diameter should be planned in order to achieve a 95% final density?

17.46. Assume that you are an instructor covering the topics described in this chapter and you are giving a quiz on the numerical aspects to test the understanding of the students. Prepare two quantitative problems and supply the answers.

Synthesis, Design, and Projects

17.47. Prepare an illustration similar to Fig. 13.1, showing the variety of PM manufacturing options.

17.48. Make sketches of PM products in which density variations (see Fig. 17.12) would be desirable. Explain why in terms of the functions of these parts.

17.49. Compare the design considerations for PM products with those for (a) casting and (b) forging. Describe your observations.

17.50. Are there applications in which you, as a manufacturing engineer, would not recommend a PM product? Explain.

17.51. Describe in detail other methods of manufacturing the parts shown in Fig. 17.1.

17.52. Using the Internet, locate suppliers of metal powders and compare the cost of the powder with the cost of ingots for five different materials.

17.53. Explain why powder-metal parts are commonly used for machine elements requiring good frictional and wear characteristics and for mass-produced parts.

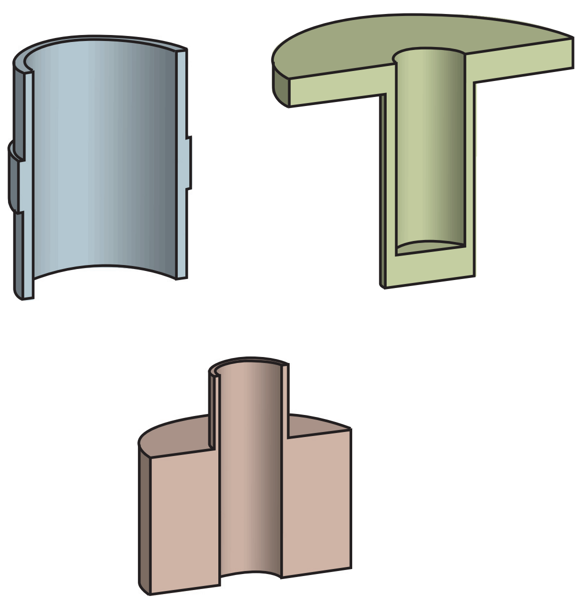
17.54. It was stated that powder-injection molding competes well with investment casting and small forgings for various materials, but not with zinc and aluminum die castings. Explain why.

17.55. Describe how the information given in Fig. 17.15 would be helpful to you in designing PM parts.

17.56. It was stated that, in the process shown in Fig. 17.20, shapes produced are limited to axisymmetric parts. Do you think it would be possible to produce other shapes as well? Describe how you would modify the design of the setup to produce other shapes, and explain the difficulties that may be encountered.

17.57. It has been noted that PM gears are very common for low-cost office equipment such as the carriage mechanism of inkjet printers. Review the design requirements of these gears and list the advantages of PM manufacturing approaches for these gears.

17.58. The axisymmetric parts shown in the accompanying figure are to be produced through PM. Describe the design changes that you would recommend.



17.59. Assume you are working in technical sales. What applications currently using non-PM parts would you attempt to develop? What would you say to your potential customers during your sales visits? What kind of questions do you think they would ask?

Chapter 18

Ceramics, Glasses, and Superconductors: Processing and Equipment

- This chapter presents the manufacturing processes for ceramics, glass, and superconductors.
- It first describes the preparation of ceramic powders, followed by operations that produce discrete parts through the basic processes of casting, pressing, extrusion, and molding.
- Drying and firing processes, followed by finishing operations for ceramics, are also discussed.
- Glass manufacture involves production of continuous shapes, such as plate, tube, and bars, through drawing, rolling, or floating methods; the operations for discrete products typically consist of molding, blowing, or pressing.
- The chapter ends with the processing of superconductors, which are produced mainly through the oxide-powder-in-tube process.

Typical products made: Ceramics: electrical insulators, rotors for gas turbines, lightweight components for high-speed machines, ball and roller bearings, seals, furnace components, ovenware, and tiles. Glass: glazing, laminated glass, bulletproof glass, bulbs, lenses, bottles, glass fibers, rods, and tubing. Superconductors: MRI magnets.

Alternative processes: Casting, forging, powder injection molding, blow molding, injection molding, additive manufacturing.

18.1 Introduction

The properties and applications of ceramics and glasses are described in Chapter 8. These materials have important characteristics, such as high-temperature strength and hardness, low electrical and thermal conductivity, chemical inertness, and resistance to wear and corrosion. The wide range of applications for these materials include electrical insulators, ball bearings, cutting tools, floor tiles, and dishes.

The processing methods employed for ceramics (Fig. 18.1) consist of (a) crushing the raw materials, (b) mixing/blending and shaping them by various means, (c) drying and firing, and (d) finishing operations, as needed, to impart the required dimensional tolerances and surface finish. For glasses, the processes involve (a) mixing and melting the raw materials in a furnace and (b) shaping them in molds using various techniques, depending on the shape and size of the part. Discrete products, such as bottles, and continuous products, such as flat glass, rods, tubing, and fibers, can be produced. Glasses can be strengthened by thermal or chemical means, as well as by laminating them with polymer sheets, as is done for automobile windshields and bulletproof glass.

18.2 Shaping Ceramics

Several techniques are available for processing ceramics into useful products (Table 18.1), depending on the type of ceramic involved and their shapes. Production of some ceramic parts, such as pottery, ovenware, and floor tiles, generally does not involve the same level of control of materials and processes as do high-tech parts made of such structural ceramics as silicon nitride, aluminum oxide, and silicon carbide. Generally, the procedure involves the following steps (Fig. 18.2):

1. Crushing or otherwise processing the raw materials into very fine particles

2. Mixing them with additives to impart certain specific characteristics, and water for formability

(a) (b)

Figure 18.1: (a) Examples of typical glass parts. (b) Examples of ceramic parts. *Source:* (a) Courtesy of Commercial Optical Manufacturing, Inc. (b) Courtesy of Shutterstock/Xiao Zhou.

Table 18.1: General Characteristics of Ceramics Processing.

Process	Advantages	Limitations
Slip casting	Large parts, complex shapes, low equipment cost	Low production rate, limited dimensional accuracy
Extrusion	Hollow shapes and small diameters, high production rate	Parts have constant cross section, limited thickness
Dry pressing	Close tolerances, high production rates (with automation)	Density variation in parts with high length-to-diameter ratios, dies require abrasive-wear resistance, equipment can be costly
Wet pressing	Complex shapes, high production rate	Limited part size and dimensional accuracy, tooling costs can be high
Hot pressing	Strong, high-density parts	Protective atmospheres required, die life can be short
Isostatic pressing	Uniform density distribution	Equipment can be costly
Jiggering	High production rate with automation, low tooling cost	Limited to axisymmetric parts, limited dimensional accuracy
Injection molding	Complex shapes, high production rate	Tooling can be costly

3. Shaping, drying, and firing the material
4. Finishing, such as by machining, grinding, and glazing, which requires an additional firing step.

The first step in processing ceramics is *crushing*, also called *comminution* or *milling*, of the raw materials. Crushing is generally done in a *ball mill* (see Fig. 17.6b), either dry or wet. Wet crushing is more effective, because it keeps the particles together and it also prevents the fine particles from contaminating the environment. The particles may then be *sized* by passing them through a sieve, followed by filtering and washing.

The ground particles are then mixed with *additives*, the functions of which are one or more of the following:

- **Binder**, for holding ceramic particles together.
- **Lubricant**, to reduce internal friction between particles during molding and also to help remove the part from the mold.
- **Wetting agent**, to improve mixing.

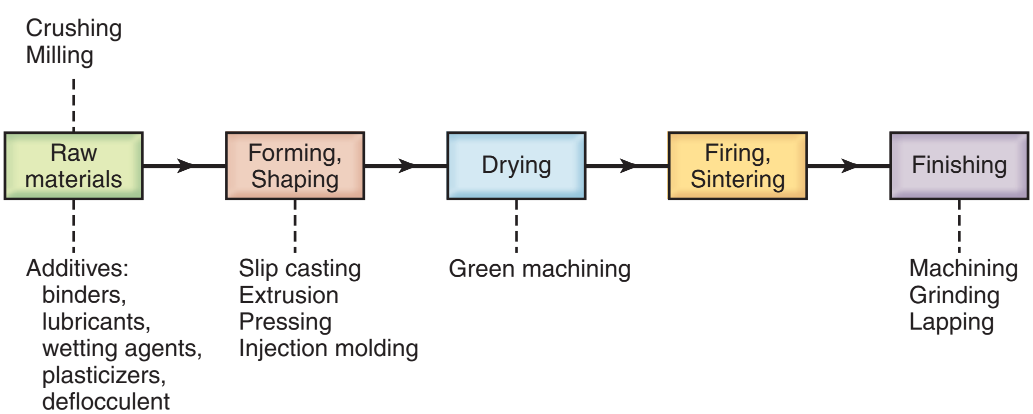


Figure 18.2: Processing steps involved in making ceramic parts.

- **Plasticizer**, to make the mix easier to shape.

- **Deflocculent**, to make the ceramic–water suspension more uniform, by changing the electrical charges on the particles of clay, so that the particles repel rather than attract each other. Typical deflocculents are Na_2CO_3 and Na_2SiO_3, in amounts of less than 1%.

The blended powder is then mixed with water to produce a pourable or formable mixture. The three basic shaping processes for ceramics are casting, plastic forming, and pressing. The parts made also may be subjected to additional processing, such as machining and grinding, for better control of their dimensions and surface finish.

18.2.1 Casting

The most common casting process is **slip casting**, also called **drain casting**, as illustrated in Fig. 18.3. A **slip** consists of ceramic particles suspended in a liquid, generally water. The slip is poured into a porous mold and may consist of several components, as also done in other shaping processes.

The slip must have sufficient fluidity and low viscosity for it to flow easily into the mold, much like the importance of fluidity of molten metals in casting operations, as described in Section 10.3. Pouring the slip must be done in a manner to avoid air entrapment, which can be significant during casting.

After the mold has absorbed some of the water from the outer layers of the suspension, it is inverted and the remaining suspension is poured out. The product is now a hollow object, as in the slush casting of metals, described in Section 11.4.3. The top of the part is then trimmed (note the trimming knife in Fig. 18.3d), the mold is opened, and the part is removed.

Large and complex parts, such as plumbing ware or art objects, can be made by slip casting. Although mold and equipment costs are low, dimensional control is poor and the production rate is low. In some applications, components of the product, such as handles for cups and pitchers, are made separately and then joined, using the slip as an adhesive. For solid-ceramic parts, the slip is supplied continuously into the mold to replenish the absorbed water, as otherwise the part will shrink. At this stage, the part is described as either a soft solid or semirigid. The higher the concentration of solids in the slip, the less water has to be removed. The part removed from the mold is referred to as a *green part*, as in powder metallurgy.

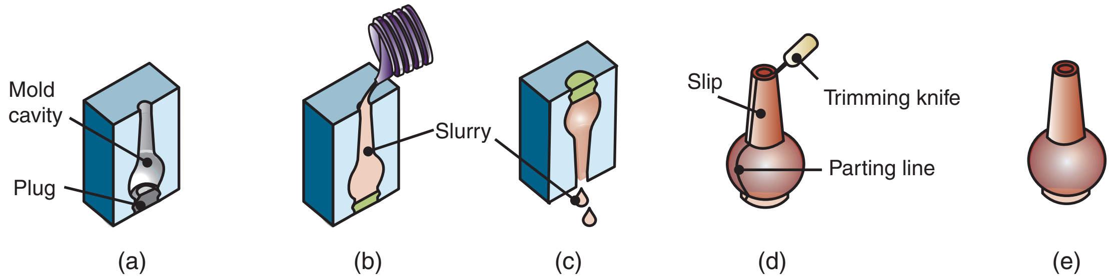


Figure 18.3: Sequence of operations in slip casting a ceramic part. (a) Mold is assembled and plug attached; some plugs incorporate draining features; (b) slurry, mixed from ceramic particles, binder and water, is poured into the mold; (c) the mold is inverted and the slurry is poured from the mold, leaving a thin coating over the mold cavity; (d) after an initial drying period, the slip is removed from the mold, and features such as parting lines and sprue lips are removed; (e) the slip is ready to be dried and fired in an oven, to develop strength and hardness.

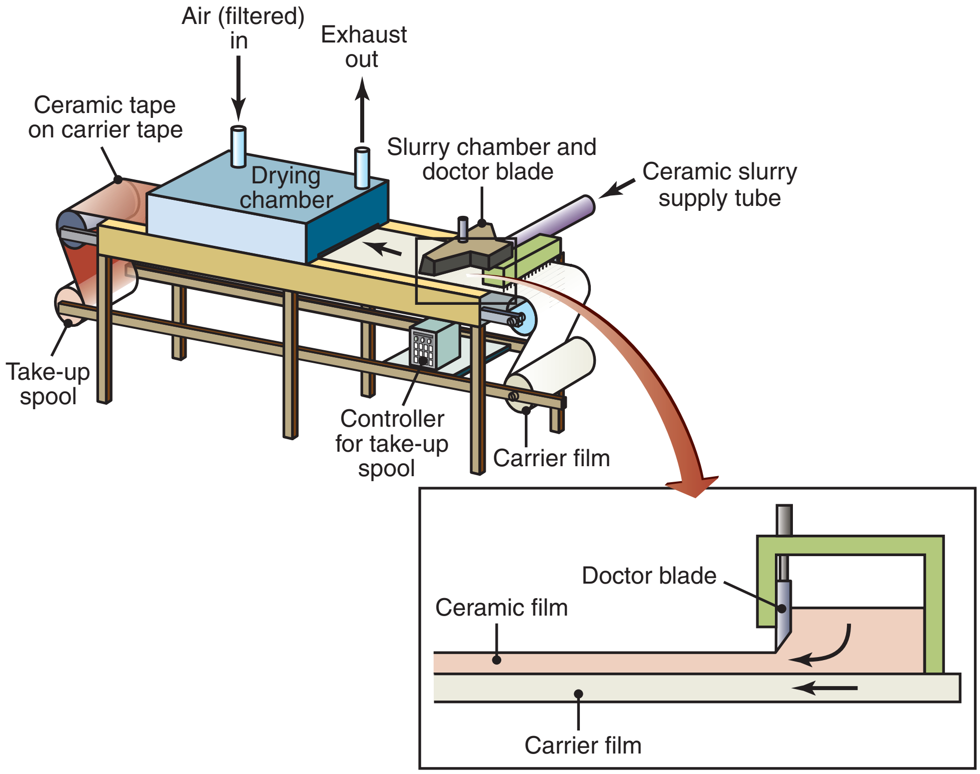


Figure 18.4: Production of ceramic sheets through the doctor-blade process.

While the parts are still green, they may be machined to produce certain features or for better dimensional accuracy. Because of the delicate nature of the green compacts, however, machining is usually done manually using simple tools. For example, the flashing in a slip casting may be removed gently with a fine wire brush or any holes can be drilled in the mold.

Doctor-blade and Other Processes. Thin sheets of ceramics, less than 1.5 mm thick, can be made by the *doctor-blade process* (Fig. 18.4). The slip is cast over a moving plastic belt, while its thickness is being controlled by a blade. Ceramic sheets also may be produced by such method as (a) *rolling* the slip between pairs of rolls and (b) *casting* the slip over a paper tape, which subsequently burns off during firing.

18.2.2 Plastic Forming

Plastic forming, also called *soft*, *wet*, or *hydroplastic forming*, can be carried out by several methods, such as extrusion, injection molding, or *jiggering* (Fig. 18.5). Plastic forming tends to orient the layered structure of the clay along the direction of material flow, and thus tends to cause anisotropic behavior of the material, both in subsequent processing and in the final properties of the ceramic product.

In *extrusion*, the clay mixture, containing 20% to 30% water, is forced through a die opening (see, for example, Fig. 19.3); the cross section of the extruded product is thus constant. There are limitations to wall thickness for hollow extrusions, because of the risk of fracture during firing (Section 18.2.4). Production rates are high and tooling costs are low. The green extrusion may be subjected to additional shaping operations.

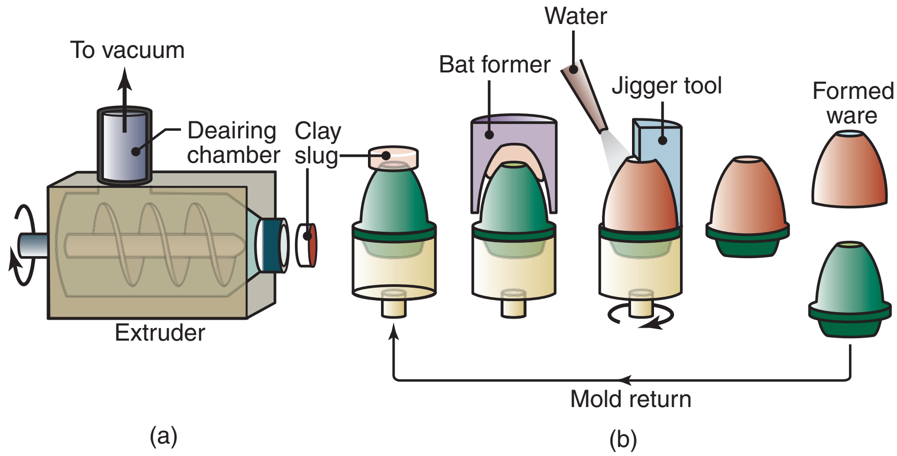


Figure 18.5: (a) Extruding and (b) jiggering operations.

18.2.3 Pressing

Dry Pressing. This is a technique similar to powder-metal compaction, as described in Section 17.3. *Dry pressing* is used for relatively simple shapes, such as whiteware, refractories for furnaces, and abrasive products. The moisture content of the mixture is generally below 4%, although it may be as high as 12%. Organic and inorganic binders (such as stearic acid, wax, starch, and polyvinyl alcohol) are usually added to the mixture; these additives give strength and also act as lubricants to aide in compaction. Dry pressing has the same high production rates and close control of dimensional accuracy as does compaction in powder metallurgy.

The pressing pressure ranges from 35 to 200 MPa. Density can vary significantly in dry-pressed ceramics, as in PM compaction (see Fig. 17.12), because of friction among the particles and at the mold walls. Density variations cause warping during firing, which is particularly severe for parts having high length-to-diameter ratios, the recommended maximum ratio being 2:1. Several methods may be used to minimize density variations, including (a) proper design of tooling, (b) vibratory pressing and impact forming, particularly for nuclear-reactor fuel elements, and (c) isostatic pressing. Modern presses for dry pressing are highly automated. The dies, usually made of carbides or hardened steel, must have high wear resistance to withstand the abrasive ceramic particles.

Wet Pressing. In *wet pressing*, the part is formed in a mold while under high pressure, in a hydraulic or mechanical press. Moisture content usually ranges from 10 to 15%. Production rates are high; however, (a) part size is limited, (b) dimensional control is difficult because of shrinkage during drying, and (c) tooling costs can be high. Wet pressing is generally used for making parts with intricate shapes, such as filters and electronic packaging.

Isostatic Pressing. This process is used for ceramics in order to obtain a uniform density distribution throughout the part during compaction (see Section 17.3.2). The white insulators for automotive spark-plugs, for example, are made by this method and at room temperature. Silicon-nitride vanes for high-temperature applications (see Fig. 8.1) are made by *hot isostatic pressing*.

Hot isostatic pressing (Section 17.3.2) also may be used, particularly to improve shape accuracy and the quality of high-technology ceramics, such as silicon carbide and silicon nitride. Glass-encapsulated HIP processing has been shown to be effective for this purpose. Hot isostatic pressing is usually necessary for fatigue or wear applications.

Jiggering. As an example of jiggering, consider ceramic dinner plates that are made by a series of steps (Fig. 18.5). First, clay slugs are extruded and formed into a *bat* over a plaster mold; they are then jiggered on a rotating mold. *Jiggering* is a motion in which a clay bat is formed by means of templates or rollers; the part is then dried and fired. This process is confined to axisymmetric parts, and has limited dimensional accuracy.

Injection Molding. *Injection molding* is used extensively for precision forming of ceramics in demanding applications such as for rocket-engine components. The raw material is first mixed with a binder, such as a thermoplastic polymer (polypropylene, low-density polyethylene, or ethylene vinyl acetate) or wax, and injection molded. The binder is usually removed by pyrolysis (inducing chemical changes by heat), and the part is then fired.

The injection-molding process can produce thin sections, typically less than 10 to 15 mm thick, from most engineering ceramics such as alumina, zirconia, silicon nitride, silicon carbide, and sialon (see Chapter 8). Thicker sections require careful control of the materials used and of the processing parameters, in order to avoid such defects as internal voids and cracks.

Hot Pressing. In this process, also called *pressure sintering*, pressure and heat are applied simultaneously, thereby reducing porosity in the part and making it denser and stronger. Graphite is commonly used as a punch and die material, and protective atmospheres usually are employed during the pressing step.

18.2.4 Drying and Firing

The next step in ceramic processing is to dry and fire the part to give it the proper strength and hardness. *Drying* is a critical stage to reduce the tendency for the part to warp or to crack from variations in its moisture content and thickness. Control of atmospheric humidity and ambient temperature during drying is important in order to reduce warping and cracking.

Loss of moisture during drying causes shrinkage of the part by as much as 20% from the original, moist size (Fig. 18.6). In a humid environment, the evaporation rate is low, and thus the moisture gradient across the thickness of the part is lower than that in a dry environment. The low moisture gradient prevents a large, uneven gradient in shrinkage from the surface to the interior, reducing the tendency for excessive warping or cracking.

A ceramic part that has been shaped by any of the methods described thus far is in the *green* state. It can be machined in order to bring it closer to a near-net shape. Although the green part should be handled carefully, machining it is not particularly difficult, because of the relative softness of the materials.

Firing, also called **sintering**, involves heating the part to an elevated temperature in a controlled environment; although some shrinkage occurs during firing, the ceramic part becomes stronger and harder. The

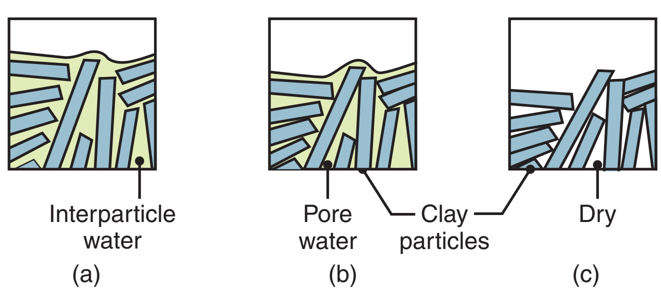


Figure 18.6: Shrinkage of wet clay caused by the removal of water during drying. Shrinkage may be as much as 20% by volume. *Source:* After F.H. Norton.

improvement in mechanical properties is due to (a) the development of strong bonds among the complex oxide particles in the ceramic body and (b) reduced porosity. **Microwave sintering** is also possible, and it can be significantly faster and less expensive than conventional sintering for larger production runs.

Nanophase ceramics, described in Section 8.2.5, can be sintered at lower temperatures than those for conventional ceramics. They are easier to fabricate because they can be (a) compacted at room temperature to high densities, (b) hot pressed to attain theoretical density, and (c) formed into net-shaped parts without using any binders or sintering aids.

18.2.5 Finishing Operations

Because firing causes dimensional changes, additional operations may be performed to (a) give the ceramic part its final shape, (b) remove any surface flaws, and (c) improve surface finish and dimensional accuracy. Although ceramics typically are hard and brittle, major advances have been made in producing **machinable ceramics** and **grindable ceramics**, thus enabling the production of ceramic components with high dimensional accuracy and good surface finish. An example is silicon carbide, which can be machined into final shapes from sintered blanks.

The finishing processes employed can be one or more of the following operations, described in detail in various sections in Part IV:

1. *Grinding*, using a diamond wheel
2. *Lapping* and *honing*
3. *Ultrasonic machining*
4. *Drilling*, using a diamond-coated drill
5. *Electrical-discharge machining*
6. *Laser-beam machining*
7. *Abrasive water-jet cutting*
8. *Tumbling*, to remove sharp edges and grinding marks.

Process selection is an important consideration because of the brittle nature of most ceramics and the additional costs involved in using some of these processes. The effect of the finishing operation on the final properties of the product also must be considered. For example, because of notch sensitivity (Section 2.9), the finer the finish of the part, the higher are its strength and load-carrying capacity, particularly its fatigue strength. Ceramic parts also can undergo *static fatigue*, as described for glass in Section 18.5.

Glazing. To improve their appearance and strength and to make them impermeable, ceramic products often are coated with a **glaze** or **enamel** (Section 34.12), which forms a glassy coating after firing.

Example 18.1 Dimensional Changes During the Shaping of Ceramic Components

Given: A solid, cylindrical ceramic part is to be made, with a final length of L=20 mm. For this material, it has been established that linear shrinkages during drying and firing are 7% and 6%, respectively, based on the dried dimension, L_d.

Find: Calculate (a) the initial length, L_o, of the part and (b) the dried porosity, P_d, if the porosity of the fired part, P_f, is 3%.

Solution:

a) On the basis of the information given and noting that firing is preceded by drying,

$$\frac{(L_d - L)}{L_d} = 0.06,$$

or

$$L = (1 - 0.06)\, L_d.$$

Hence,

$$L_d = \frac{20}{0.94} = 21.28 \text{ mm}$$

and

$$L_o(1 + 0.07)L_d = (1.07)(21.28) = 22.77 \text{ mm}.$$

b) Since the final porosity is 3%, the actual volume, V_a, of the solid material in the part is

$$V_a = (1 - 0.03)\, V_f = 0.97V_f,$$

where V_f is the volume of the part after firing. Because the linear shrinkage during firing is 6%, the dried volume, V_d of the part can be determined as

$$V_d = \frac{V_f}{(1 - 0.06)^3} = 1.2V_f.$$

Hence,

$$\frac{V_a}{V_d} = \frac{0.97}{1.2}, \text{ or } 81\%$$

Therefore, the porosity, P_d, of the dried part is 19%.

18.3 Forming and Shaping of Glass

Glass is processed basically by melting and shaping it, either in molds or by blowing. The shapes produced include flat sheets and plates, rods, tubing, glass fibers, and discrete products, such as bottles, lenses, automobile headlights, and cookware. Glass products may be as thick as those for large telescope mirrors, and as thin as those for holiday tree ornaments. The strength of glass can be improved by thermal and chemical treatments, inducing compressive surface residual stresses, or by laminating it with a thin sheet of tough plastic.

Glass products generally can be categorized as follows:

1. **Flat sheets** or **plates**, ranging in thickness from about 0.8 to 10 mm, and used as window glass, glass doors, and tabletops.

2. **Rods** and **tubing**, used for neon lights, decorative artifacts, and for processing and handling chemicals.

3. **Discrete products**, such as bottles, vases, and eyeglasses.

4. **Glass fibers**, as reinforcements in composite materials (Section 9.2.1) and for use in fiber optics.

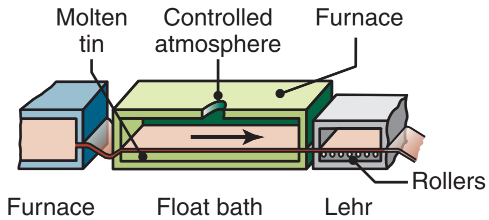


Figure 18.7: The float method of forming sheet glass.

All forming and shaping processes begin with molten glass, at a temperature typically in the range of 1000° to 1200°C, and has the appearance of a red-hot, viscous liquid.

18.3.1 Flat-sheet and Plate Glass

Flat-sheet glass can be made by any of the following three methods from the molten state, with glass supplied from a melting furnace or tank:

1. In the **float method** (Fig. 18.7), molten glass from the furnace is fed into a long bath in which the glass, under a controlled atmosphere and at a temperature of 1150°C, floats over a bath of molten tin. The glass, at a temperature of about 650°C, then moves over rollers into another chamber (*lehr*) where it solidifies. *Float glass* has smooth (*fire-polished*) surfaces, thus further finishing operations, such as grinding or polishing, are not necessary; the width can be as much as 4 m. Both thin and plate glass are made by this process.

2. The **drawing** process for making flat glass sheets or plates involves passing the molten glass through a pair of rolls (Fig. 18.8a). The solidifying glass is squeezed between these two rolls, forming it into a flat sheet; it then moves forward over a set of smaller rolls.

3. In the **rolling** process (Fig. 18.8b), the molten glass is squeezed between powered rollers, thereby forming a sheet, with a surface that is somewhat rough. The surfaces of the glass may also be embossed with a pattern, using *textured* roller surfaces, thus the glass surface becomes a replica of the roll surface.

18.3.2 Tubing and Rods

Glass *tubing* is basically produced by the process shown in Fig. 18.9. Molten glass is wrapped around a rotating (cylindrical or cone-shaped) hollow mandrel, and is then drawn out by a set of rolls. Air is blown

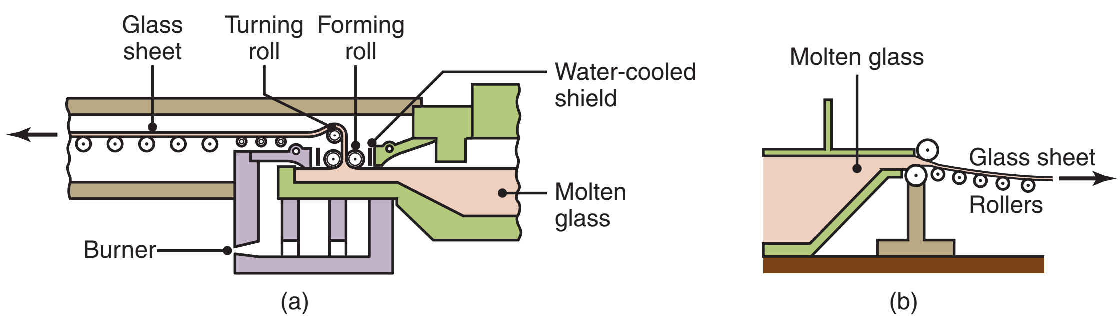


Figure 18.8: (a) Drawing process for drawing sheet glass from a molten bath. (b) Rolling process.

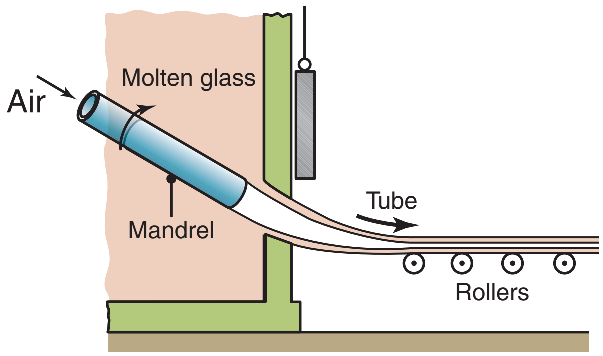


Figure 18.9: Manufacturing process for glass tubing. Air is blown through the mandrel to keep the tube from collapsing. Glass tubes for fluorescent bulbs are made by this method.

through the mandrel to prevent the glass tube from collapsing. The machines may be horizontal, vertical, or slanted downward. This method is also used in making glass tubes for fluorescent bulbs.

An alternative method for making glass tubes is by *extruding* a strip of glass (with a thin rectangular cross section), which is then wrapped obliquely (at an angle) around a rotating mandrel. The molten glass strips bond together along their edges, forming a continuous tube; it is then drawn off the mandrel in a continuous manner. Glass *rods* are extruded or drawn directly from a molten bath, without the need for internal pressurization.

18.3.3 Discrete Glass Products

Blowing. Hollow and thin-walled glass items, such as bottles, vases, and flasks, are made by *blowing*, a process that is similar to blow molding of thermoplastics (Section 19.4). The steps involved in the production of an ordinary glass bottle by the blowing process are shown in Fig. 18.10. Blown air expands a hollow *gob* of heated glass against the inner walls of the mold. The mold surfaces are usually coated with a parting agent, such as oil or emulsion, to prevent the glass from sticking to the mold surfaces. Blowing may be followed by a second blowing operation to finalize the product shape, called the *blow and blow* process.

The surface finish of glass parts made by blowing is acceptable for most applications, such as bottles and jars. It is difficult to precisely control the wall thickness of the product, because of the lack of an inner mold, but the process is economical for high-rate production.

Pressing. In the *pressing* method, a gob of molten glass is placed into a mold, and is pressed into a confined die cavity with the use of a plunger; the process is thus similar to closed-die forging (Section 14.3). The mold may be made in one piece, such as that shown in Fig. 18.11, or it may be a *split mold* (Fig. 18.12). After pressing, the solidifying glass acquires the shape of the mold-plunger cavity. Because of the confined environment, the product has a better dimensional accuracy than can be obtained with blowing.

Pressing in one-piece molds cannot be used for (a) shapes of parts from which the plunger cannot be retracted or (b) thin-walled items, because of high forces needed to produce the walls, and distortion upon part removal. For example, split molds are used for bottles, whereas pressing can be combined with blowing for thin-walled items, in a process known as *press and blow*. The pressed part is subjected to air pressure (hence the term blow), which further expands the molten glass into the mold.

Centrifugal Casting. Also known as *spinning* or *rotocasting* (Fig. 18.13), this process is similar to that used for casting metals (see Section 11.4.6), whereby the centrifugal force pushes the molten glass against the mold walls, where it begins to solidify. Typical products made are large lenses for research telescopes and architectural shapes.

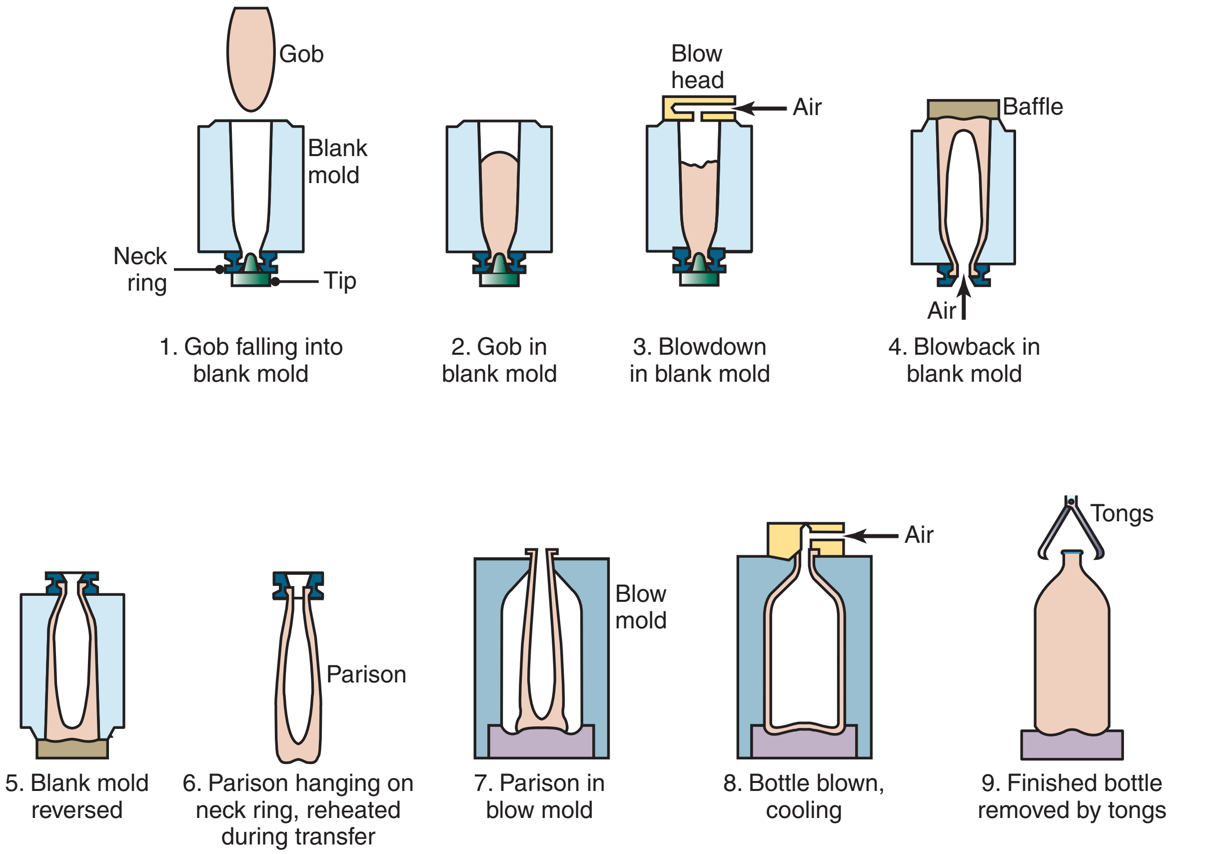


Figure 18.10: Steps in manufacturing an ordinary glass bottle. *Source:* After F.H. Norton.

Sagging. Shallow dish-shaped or lightly embossed glass parts can be made by the *sagging* process. A sheet of hot glass is placed over a mold and heated, whereby the glass sags by its own weight and takes the shape of the mold. The process is similar to the thermoforming of thermoplastics (Section 19.6), but no pressure or vacuum is involved. Typical parts made are dishes, sunglass lenses, mirrors for telescopes, and lighting panels.

Glass Ceramics Manufacture. Glass ceramics (trade names: *Pyroceram, Corningware*) contain large proportions of several oxides, as noted in Section 8.5. Their manufacture involves a combination of the methods used for ceramics and glasses. Glass ceramics are shaped into discrete products, such as dishes and baking pans, and then heat treated, whereby glass becomes **devitrified** (recrystallized).

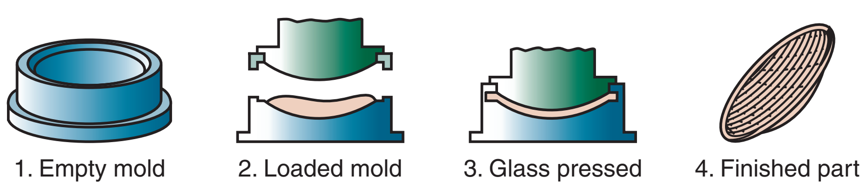


Figure 18.11: Manufacturing a glass item by pressing molten glass into a mold. *Source:* Courtesy of Corning Glass Works.

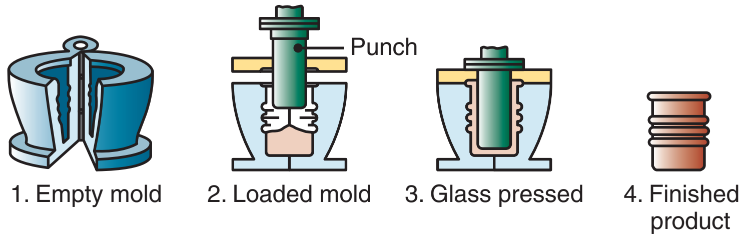


Figure 18.12: Pressing molten glass into a split mold. *Source:* After E.B. Shand.

18.3.4 Glass Fibers

Continuous glass fibers are drawn through multiple orifices (200 to 400 holes) in heated platinum plates, at speeds as high as 500 m/s. Fibers as small as 2 μm in diameter can be produced by this method. To protect their surfaces, fibers are subsequently coated with chemicals, known as *sizing*, which are mainly silane compounds in water, but other sizing blends are also used. Short or *chopped* fibers are produced by subjecting long fibers to compressed air or steam as they leave the orifice, being broken into very short pieces.

Glass wool, which consists of short glass fibers, are used as a thermal insulating material and for acoustic insulation. They are made by a *centrifugal spraying process*, in which molten glass is ejected (*spun*) from a rotating head. The diameter of the fibers typically ranges from 20 to 30 μm.

18.4 Techniques for Strengthening and Annealing Glass

Glass can be strengthened by several processes, and discrete glass products may be subjected to annealing and to other finishing operations to impart desired properties and surface characteristics.

Thermal Tempering of Glass. In this process, also called *physical tempering* or *chill tempering*, the surfaces of hot glass are rapidly chilled by a blast of air (Fig. 18.14). As a result, the surfaces shrink and, at first, tensile stresses develop on the surfaces. The bulk of the glass then begins to cool, and because it contracts, the already solidified surfaces of the glass also are forced to contract. Consequently, compressive residual stresses develop on the surfaces, while the interior develops tensile stresses (see also Section 2.11). Compressive surface stresses improve the strength of the glass in the same way that they do in metals and other materials (see Section 2.11).

The higher the coefficient of thermal expansion of the glass and the lower its thermal conductivity, the higher will be the level of residual stresses developed, and hence, the stronger the glass becomes. Thermal

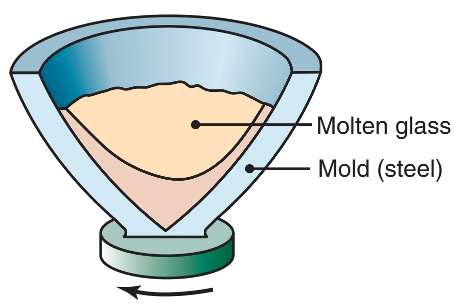


Figure 18.13: Centrifugal casting of glass. *Source:* Courtesy of Corning Glass Works.

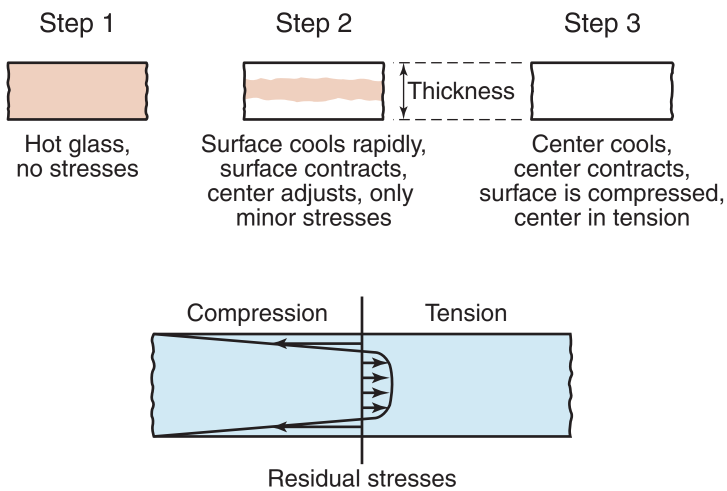


Figure 18.14: (a) Stages involved in inducing compressive surface residual stresses for improved strength. (b) Residual stresses in a tempered glass plate. *Source:* Courtesy of Corning Glass Works.

tempering takes a relatively short time (minutes), and can be applied to most glasses. Because of the high amount of energy stored in residual stresses, **tempered glass** shatters into numerous pieces when broken. The broken pieces are not as sharp and as hazardous as those from a broken ordinary window glass, which has a sharp jagged fracture path.

Chemical Tempering. In this process, the glass is heated in a bath of molten KNO_3, K_2SO_4, or $NaNO_3$, depending on its type. *Ion exchanges* then take place, with larger atoms replacing the smaller atoms on the surface of the glass; as a result, residual compressive stresses develop on the surface.

Chemical tempering may be done at various temperatures. At low temperatures, part distortion is minimal, thus complex part shapes can be tempered. At elevated temperatures, there may be some distortion of the part, but the product can then be used at higher temperatures without loss of strength. The time required for chemical tempering is much longer than that for thermal tempering.

Laminated Glass. Laminated glass consists of two or more pieces of flat glass with a thin sheet of tough plastic between each layer. As a result, when laminated glass cracks, the pieces are held together by the plastic sheet and it becomes far less hazardous. This is a phenomenon commonly observed in shattered automobile windshields.

Flat glass for glazing windows and doors can be strengthened with wire netting with a hexagonal mesh, such as *chicken wire*, which is embedded in the glass during its production. When a hard object strikes its surface, the glass breaks but the pieces will be held together because of the embedded wire. This type of glass will have both toughness and flexibility (see also Section 2.2.4).

Bulletproof Glass. Laminated glass has significant ballistic impact resistance, and it can prevent the full penetration of solid objects because of the presence of a tough thermoplastic polymer film in between the two layers of glass (see Section 7.3). *Bulletproof glass*, used in some automobiles, armored bank vehicles, and buildings, is a more challenging design, because of the very high speed and energy level of the bullet and the small size and the shape of the bullet tip.

Although there are several designs, bulletproof glass basically consists of glass plates laminated with a polymer sheet. The capacity of a bulletproof glass to stop a bullet depends on (a) the type and thickness of the glass; (b) the size, shape, weight, and speed of the bullet; and (c) the properties and thickness of the polymer sheet. Polycarbonate is commonly used because of its high toughness and flexibility. Laminated with thick glass, it can stop a bullet, although the glass itself develops a circular shattered region. In order to maintain the transparency of the bulletproof glass and minimize its distortion, the *index of refraction* of the glass and the polymer must be nearly identical.

18.4.1 Finishing Operations

As in metal products, residual stresses can develop in glass products if they are not cooled at a sufficiently low rate. In order to ensure that the product is free of these stresses, it is **annealed**, by a process similar to the stress-relief annealing of metals (Section 4.11). The glass is first heated to a certain temperature, and then cooled slowly. Depending on the size, thickness, and type of the glass, annealing times may range from a few minutes to as long as 10 months, as in the case of a 600-mm mirror for a telescope in an observatory.

In addition to annealing, glass products may be subjected to further operations, such as cutting, drilling, grinding, and polishing. Sharp edges and corners can be made smooth by (a) grinding, as can be seen in glass tops for desks and shelves, or (b) holding a torch against the periphery (**fire polishing**), which rounds the edges by localized softening of the glass and by surface tension.

In all finishing operations on glass, as well as other brittle materials, care should be exercised to ensure that there is no surface damage, especially the presence of stress raisers such as rough surface finish and scratches. Because of its notch sensitivity, even just a single scratch on glass can cause premature failure, especially if the scratch is in a direction where the tensile stresses are a maximum.

18.5 Design Considerations for Ceramics and Glasses

Ceramic and glass products require careful selection of composition, processing methods, finishing operations, and methods of assembly with other components. With such properties as poor tensile strength, sensitivity to internal and external defects, low impact toughness, and static fatigue, the consequences of part failure are always a significant factor in designing ceramic and glass products. On the other hand, these limitations must be balanced against such desirable and important material characteristics as hardness, scratch resistance, compressive strength at room and elevated temperatures, and a wide range of diverse physical properties.

As noted in Section 8.3.1, ceramics and glasses undergo a phenomenon called *static fatigue*, whereby after a period of time they can suddenly and without any warning fracture under a static load. Although this phenomenon does not occur in a vacuum or in dry air, provisions must be made to prevent such failure. A general rule is that, in order for a glass item to withstand a certain load for 1000 hours or longer, the maximum stress that can be applied to it is about one-third of the maximum stress that it can withstand during the first second of loading. The control of processing parameters and of the quality and level of impurities in the raw materials are also important.

Dimensional changes, warping, the possibility of cracking during processing, and service life are significant factors in selecting methods for shaping glass and ceramics. When a part made of such a material is a component of a larger assembly, its *compatibility* with other components is an important consideration. Particularly significant are the type of external forces and thermal expansion, such as in seals and windows with metal frames. Recall that Table 3.1 displayed a wide range of thermal expansion coefficients for various metallic and nonmetallic materials. Thus, when a plate glass fits too tightly within a metal window frame, factors such as temperature variations within the glass (sun shining on only a portion of a window) can cause thermal stresses so high that they may lead to cracking; this is a phenomenon often observed in some tall buildings. A common solution is placing rubber seals along the glass and the window frame, to allow for dimensional changes.

18.6 Processing of Superconductors

Although *superconductors* (Section 3.7) have major energy-saving potential in the generation, storage, and distribution of electrical power, their processing into useful shapes and sizes for practical applications has presented significant difficulties. The following are two basic types of superconductors:

1. Metals, called **low-temperature superconductors** (LTSC), include combinations of niobium, tin, and titanium. For example, niobium–tin alloys, cooled by liquid helium, constitute the superconducting magnet used in most *magnetic resonance imaging* (MRI) scanners for medical imaging.

2. Ceramics, called **high-temperature superconductors** (HTSC), include various copper oxides. Here, "high" means closer to ambient temperature, although the commercially important HTSCs maintain superconductivity above the boiling point of liquid nitrogen (-196°C).

Ceramic superconducting materials are available in powder form. The difficulty in manufacturing them is their (a) inherent brittleness and (b) anisotropy, making it difficult to align the grains in the proper direction to achieve high efficiency. The smaller the grain size, the more difficult it is to align the grains.

The basic manufacturing process for superconductors consists of the following steps:

1. Preparing the powders, blending, and grinding them in a ball mill (see Fig. 17.6b) to a grain size of 0.5 to 10 μm.

2. Forming the powder into the desired shape.

3. Heat treating the product to enhance properties.

The most common forming process is the **oxide-powder-in-tube** (OPIT) method. The powder is first packed into silver tubes (silver has the highest electrical conductivity of any metal; see Table 3.1), and sealed at both ends. The tubes are then shaped, by such processes as swaging, drawing, extrusion, isostatic pressing, and rolling. The final product may be wire, tape, coil, or in bulk form. Other methods of processing superconductors are (a) coating silver wire with superconducting material, (b) depositing superconductor films by laser ablation (removing material), (c) forming by the doctor-blade process (Section 18.2.1), (d) explosive cladding, and (e) chemical spraying. The shaped part subsequently may be heat treated to improve the grain alignment of the superconducting powder.

Case Study 18.1 Production of High-temperature Superconducting Tapes

Two bismuth-based oxides are preferred as superconducting ceramic materials for various commercial and military applications, such as electrical propulsion for ships and submarines, shallow-water and ground minesweeping systems, transmission cable generators, and *superconducting magnetic energy storage* (SMES). Different processing methods have been explored to produce wires and multifilament tapes. The powder-in-tube process (Fig. 18.15) has been used successfully to fabricate long lengths of bismuth-based wires and tapes, with specific desirable properties. The approach uses the following steps:

1. First, a composite billet is produced, using a silver casing and ceramic powder. The casing is made of an annealed high-purity silver, filled with the bismuth-ceramic powder in an inert atmosphere. A steel ram is used to compact the casing in several increments, up to a 30% relative density. In order to minimize density gradients, such as those shown in Fig. 17.12, about 1 g of powder is added to the billet for each stroke of the ram. Each billet is weighed and measured to verify the initial packing density. The billet ends are then sealed with a silver alloy, to avoid contamination during subsequent processing.

2. The billet is then extruded and drawn, to reduce its diameter and increase the powder density. Billets are drawn on a draw bench (see Fig. 15.24) down to a wire with a final diameter of 1.63 mm. It takes 12 passes, with a 20% reduction per pass, to perform this task. The dies have a semicone angle of 8°, and the drawing speed is approximately 1.4 m/min. A semisoluble oil or zinc-stearate spray is used as a lubricant.

3. Following the drawing process, the wire is transformed progressively into tape in a single-stand rolling mill in two-high and four-high configurations. For the four-high case, the diameter of the backup rolls (which are the work rolls for the two-high configuration) is 213 mm, and the diameter of the work rolls is 63.5 mm. The final tape dimensions are 100 to 200 μm in thickness and 2 to 3 mm in width, with a ceramic core ranging from 40 to 80 μm in thickness and 1.0 to 1.5 mm in width.

Source: Courtesy of S. Vaze and M. Pradheeradhi, Concurrent Technologies Corporation.

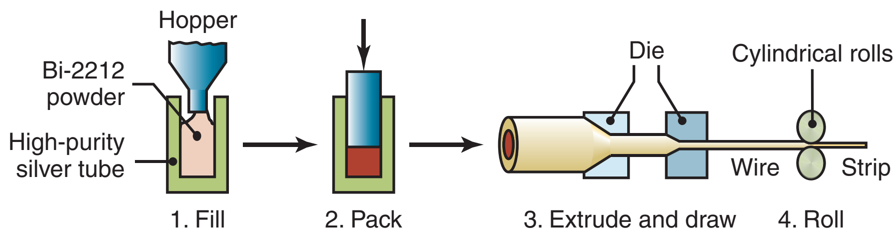


Figure 18.15: Schematic illustration of the powder-in-tube process.

Summary

- Ceramic products are shaped by various casting, plastic forming, or pressing techniques; the parts are then dried and fired to impart strength and hardness. Finishing operations, such as machining and grinding, may be performed to give the part its final shape and dimensional accuracy, or to subject it to surface treatments. Because of their brittleness, ceramics are processed with due consideration of distortion and cracking. The control of raw-material quality and processing parameters also are important.

- Glass products are made by several shaping processes, similar to those used for ceramics and plastics. They are available in a wide variety of forms, compositions, and mechanical, physical, and optical properties. Their strength can be improved by thermal and chemical treatments.

- Continuous methods of glass processing are drawing, rolling, and floating. Discrete glass products can be manufactured by blowing, pressing, centrifugal casting, or sagging. The parts subsequently may be annealed to relieve residual stresses.

- Design considerations for ceramics and glasses are guided by such factors as their low tensile strength and toughness, and their sensitivity to external and internal defects. Warping and cracking during production are also important considerations.

- Manufacturing superconductors into useful consumer and industrial products can be challenging because of the anisotropy and brittleness of the materials involved. Although new processes are being developed, the basic process consists of packing the powder into a silver tube and forming it into desired shapes.

Key Terms

Binder
Blow and blow
Blowing
Bulletproof glass
Centrifugal casting
Chemical tempering
Deflocculent
Doctor-blade process
Drawing
Fire polishing
Firing
Float glass
Gob
High-temperature superconductors
Hot pressing
Injection molding
Jiggering
Laminated glass
Low-temperature superconductors
Microwave sintering
Oxide-powder-in-tube process
Plastic forming
Plasticizer
Press and blow
Pressing
Sagging
Slip
Slip casting
Static fatigue
Tempered glass
Thermal tempering
Wetting agent

Bibliography

Advanced Ceramic Technologies & Products, Springer, 2012.
Bansal, N.P., and Boccaccini, A.R., (eds.), **Ceramics and Composites: Processing Methods**, Wiley-American Ceramic Society, 2012.
Bansal, N.P., and Lamon, J., **Ceramic Matrix Composites**, Wiley, 2014.
Basu, B., and Balani, K., **Advanced Structural Ceramics**, Wiley-American Ceramic Society, 2011.
Bengisu, M., **Engineering Ceramics**, Springer, 2010.
Bourhis, E.L., **Glass: Mechanics and Technology**, 2nd ed., Wiley, 2014.
Carter, C.B., and Norton, W.G., **Ceramic Materials: Science and Engineering**, 2nd ed., Springer, 2013.
Holand, W., and Beall, G.H., **Glass Ceramic Technology**, 2nd ed., Wiley-American Ceramic Society, 2012.
Marchi, J., **Biocompatible Glasses**, Springer, 2016.
Rahaman, M.N., **Ceramics Processing**, 2nd ed., CRC Press, 2017.
Richerson, D.W., **Modern Ceramic Engineering: Properties, Processing, and Use in Design**, 3rd ed., CRC Press, 2005.
Riley, F.L., **Structural Ceramics: Fundamentals and Case Studies**, Cambridge, 2009.
Shackelford, G.F., and Doremus, R.H., **Ceramic and Glass Materials: Structure, Properties and Processing**, Springer, 2008.
Shelby, J.E., **Introduction to Glass Science and Technology**, 2nd ed., Royal Society of Chemistry, 2005.

Review Questions

18.1. Outline the steps involved in processing (a) ceramics and (b) glasses.

18.2. List and describe the functions of additives in ceramics.

18.3. Describe the doctor-blade process.

18.4. Explain the advantages of isostatic pressing.

18.5. What is jiggering? What shapes does it produce?

18.6. Name the parameters that are important in drying ceramic products.

18.7. What types of finishing operations are used on ceramics? On glass? Why?

18.8. Describe the methods by which sheet glass is made.

18.9. What is float glass?

18.10. What is a gob?

18.11. How is glass tubing produced?

18.12. What is the difference between physical and chemical tempering of glass?

18.13. What is the structure of laminated glass? Bulletproof glass?

18.14. How are glass fibers made? What are their sizes?

18.15. Describe the processes of chemical and thermal tempering of glass.

18.16. What is the doctor-blade process?

18.17. Is diamond a ceramic? Explain.

Qualitative Problems

18.18. Inspect various products; noting their shape, color, and transparency, identify those that are made of (a) ceramic, (b) glass, and (c) glass ceramics.

18.19. Describe the differences and similarities in processing metal powders vs. ceramics.

18.20. Which property of glasses allows them to be expanded to large dimensions by blowing? Can metals undergo such behavior? Explain.

18.21. Explain why ceramic parts may distort or warp during drying. What precautions should be taken to avoid this situation?

18.22. What properties should plastic sheets have to be used in laminated glass? Why?

18.23. It is stated that the higher the coefficient of thermal expansion of a glass and the lower its thermal conductivity, the higher the level of the residual stresses developed. Explain why.

18.24. Are any of the processes used for making discrete glass products similar to ones described in preceding chapters? Describe them.

18.25. Injection molding is a process that is used for powder metals, polymers, and ceramics. Explain why is this so.

18.26. Explain the phenomenon of static fatigue and how it affects the service life of a ceramic or glass component.

18.27. Describe and explain the differences in the manner in which each of the following would fracture when struck with a heavy piece of rock: (a) ordinary window glass, (b) tempered glass, and (c) laminated glass.

18.28. Is there any flash that develops in slip casting? How would you propose to remove such flash?

18.29. Explain the difficulties involved in making large ceramic components. What recommendations would you make to improve the process?

Quantitative Problems

18.30. Using Example 18.1, calculate (a) the porosity of the dried part if the porosity of the fired part is to be 8% and (b) the initial length, L_o of the part if the linear shrinkages during drying and firing are 7% and 5%, respectively.

18.31. What would be the answers to Problem 18.30 if the quantities given were halved?

18.32. Assume that you are an instructor covering the topics described in this chapter and you are giving a quiz on the numerical aspects to test the understanding of the students. Prepare two quantitative problems and supply the answers.

Synthesis, Design, and Projects

18.33. List some similarities and differences between the processes described in this chapter and those in (a) Part II on metal casting and (b) Part III on forming and shaping.

18.34. Consider some ceramic products with which you are familiar, and outline a sequence of processes that you think were used to manufacture them.

18.35. Make a survey of the technical literature, and describe the differences, if any, between the quality of glass fibers made for use in reinforced plastics and those made for use in fiber-optic communications. Comment on your observations.

18.36. How different, if any, are the design considerations for ceramics from those for other materials? Explain.

18.37. Visit a ceramics/pottery shop, and investigate the different techniques used for coloring and decorating a ceramic part. What are the methods of applying a metallic finish to the part?

18.38. Give examples of designs and applications in which static fatigue should be taken into account.

18.39. Construct a table that describes the approach for manufacturing plate from (a) metals; (b) thermoplastics; (c) ceramics; (d) powder metal; (e) glass. Include descriptions of process capabilities and shortcomings in your descriptions.

18.40. Pyrex cookware displays a unique phenomenon: it functions well for a large number of cycles and then shatters into many pieces. Investigate this phenomenon, list the probable causes, and discuss the manufacturing considerations that may alleviate or contribute to such failures.

18.41. It has been noted that the strength of brittle materials such as ceramics and glasses are very sensitive to surface defects such as scratches (notch sensitivity). Obtain some pieces of these materials, make scratches on them, and test them by carefully clamping in a vise and bending them. Comment on your observations.

18.42. Describe your thoughts on the processes that can be used to make (a) small ceramic statues, (b) whiteware for bathrooms, (c) common brick, (d) floor tile.

18.43. Perform a literature search, and make a list of automotive parts or components that are made of ceramics. Explain why they are made of ceramics.

18.44. Describe your thoughts on the processes that can be used to make (a) a small ceramic ball, (b) a small statue, (b) whiteware for bathrooms, (c) common brick, and (d) floor tile.

18.45. Describe any special design considerations in products that use ceramics with a near-zero coefficient of thermal expansion.

Chapter 19

Plastics and Composite Materials: Forming and Shaping

- This chapter describes the manufacturing processes for producing polymers and composite materials.
- Extrusion is widely used to make rods, tubing, and also pellets as a base stock for producing plastic parts, sheet, and film.
- The chapter describes several molding operations for producing discrete parts, including injection- and reaction-injection molding, transfer molding, rotational molding, and compression molding.
- The processes associated with shaping composite materials are then described, including compression and vacuum molding, contact molding, pultrusion, and filament winding.
- The chapter ends with a description of the characteristics of the machinery used, mold design principles, and economic considerations in polymer processing.

Typical parts made: Extensive variety of consumer and industrial products with a range of colors and characteristics.

Alternative processes: Casting, forming, additive manufacturing, powder metallurgy, and machining.

19.1 Introduction

Processing of plastics and elastomers involves operations similar to those used for forming and shaping of metals described in preceding chapters. Processing of rubbers and elastomers began in the 1800s, with the discovery of vulcanization by C. Goodyear in 1839. Plastics began to be developed in the 1920s, and rapid progress in the 1940s and onward led to important advances in polymeric materials, design, and manufacturing to make numerous products in large quantities and low cost. In the 1970s, reinforced plastics began to be introduced, leading the way for rapid progress in the use of composite materials with unique properties and applications, as well as the associated challenges in producing them.

As noted in Chapter 7, thermoplastics *melt* and thermosets *cure* at relatively low temperatures. Hence, unlike metals, they are relatively easy to handle, and requiring much less force and energy to process them. Plastics, in general, can be molded, cast, shaped, and machined into complex shapes, with relative ease and at high production rates (Table 19.1). Plastics can be joined by a variety of techniques (Section 32.6) and also can be coated using various techniques (Chapter 34).

Plastics are shaped into discrete or continuous products, such as sheets, plates, rods, and tubing; they may then be shaped by secondary processes into products. The types and properties of polymers, and the complexity of the parts that can be produced, are greatly influenced by their manufacturing and processing characteristics.

Plastics are usually shipped to manufacturing plants as pellets, granules, or powders, and are (a) softened or melted (for thermoplastics) just before shaping them and (b) cured and set (for thermosetting plastics). Liquid plastics that cure into solid form also are used, especially in making thermosets and reinforced-plastic parts. With increasing awareness of the environment, raw materials may consist of reground, chopped, or melted plastics, delivered from recycling centers or produced from natural sources.

Following the outline shown in Fig. 19.1, this chapter describes the basic processes, operations, machinery, and economics of forming and shaping plastics. The processing techniques for reinforced plastics and metal-matrix and ceramic-matrix composites are also described. The chapter begins with melt-processing techniques, starting with extrusion, and continuing on to various molding processes.

Table 19.1: General Characteristics of Forming and Shaping Processes for Plastics and Composite Materials.

Process	Characteristics
Extrusion	Continuous, uniformly solid or hollow, and complex cross sections; high production rates; relatively low tooling costs; wide tolerances
Injection molding	Complex shapes of various sizes; thin walls; very high production rates; costly tooling; good dimensional accuracy
Structural foam molding	Large parts with high stiffness-to-weight ratio; less expensive tooling than in injection molding; low production rates
Blow molding	Hollow, thin-walled parts and bottles of various sizes; high production rates; relatively low tooling costs
Rotational molding	Large, hollow items of relatively simple shape; relatively low tooling costs; relatively low production rates
Thermoforming	Shallow or relatively deep cavities; low tooling costs; medium production rates
Compression molding	Parts similar to impression-die forging; expensive tooling; medium production rates
Transfer molding	More complex parts than compression molding; higher production rates; high tooling costs; some scrap loss
Casting	Simple or intricate shapes made with rigid or flexible low-cost molds; low production rates
Processing of composite materials	Long cycle times; expensive operation; tooling costs depend on process

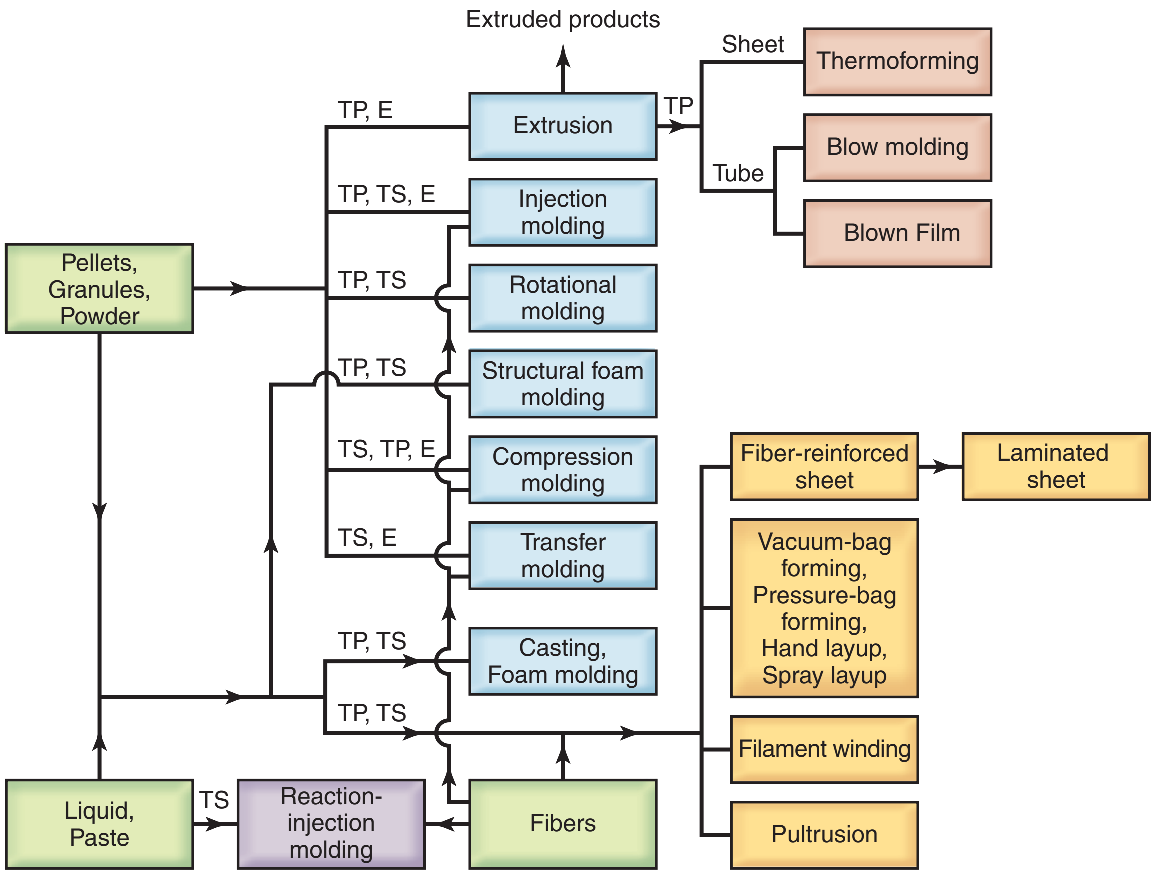


Figure 19.1: Outline of forming and shaping processes for plastics, elastomers, and composite materials. (TP = Thermoplastic; TS = Thermoset; E = Elastomer.)

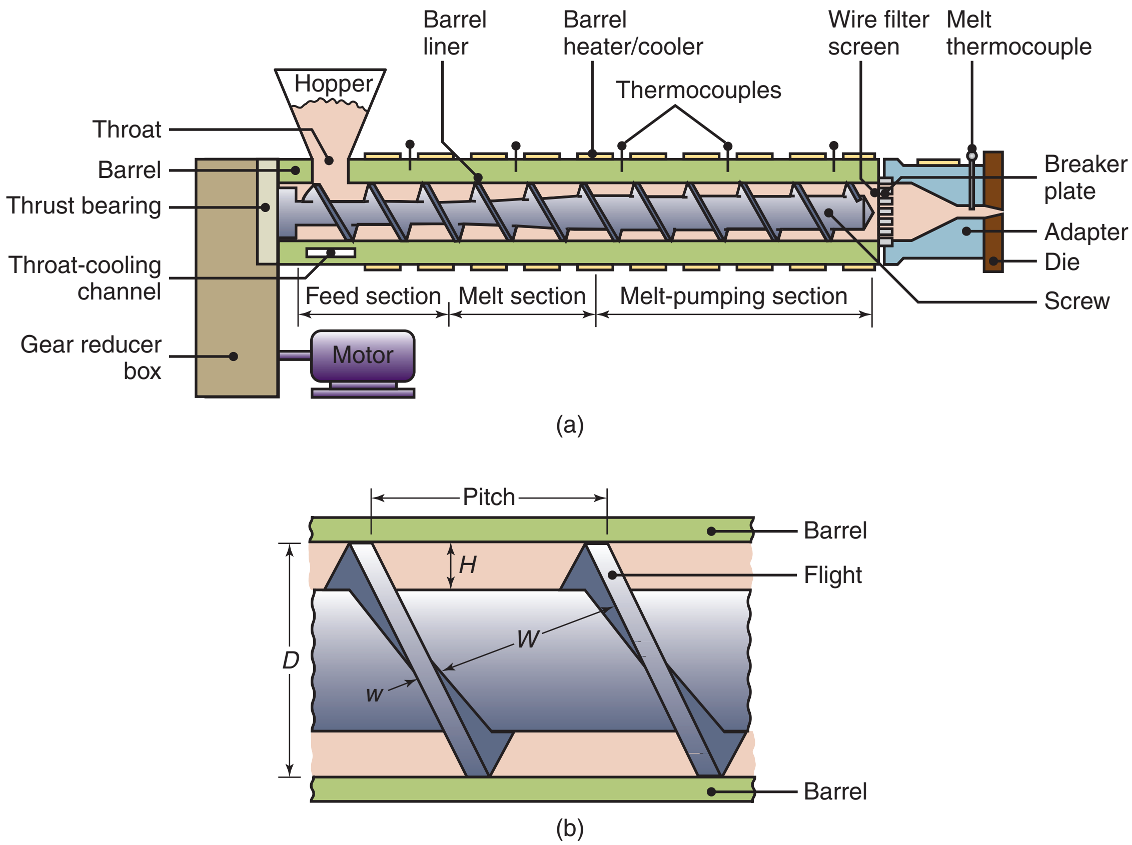


Figure 19.2: Common extrusion die geometries: (a) Schematic illustration of a typical screw extruder. (b) Geometry of an extruder screw metering or pumping section. Complex shapes can be extruded with relatively simple and inexpensive dies.

19.2 Extrusion

In *extrusion*, which produces the largest volume of plastics made, raw materials in the form of thermoplastic **pellets,** granules, or powder, are placed into a *hopper* and fed into the barrel of a *screw extruder*. The barrel is equipped with a helical screw that builds up pressure in the barrel, blends the pellets, and conveys them down the barrel towards the extrusion die. The barrel heaters and the internal friction from the mechanical action of the screw heat the pellets, and *liquefies* them.

Screws have three distinct sections:

1. *Feed section*: Conveys the material from the hopper into the central region of the barrel.

2. *Melt section*, also called *compression* or *transition section*: Where the heat, generated by the viscous shearing of the plastic pellets and by the external heaters around the barrel, cause melting to begin.

3. *Metering* or *pumping section*: Where additional shearing and melting take place, with pressure building up at the die entrance.

The lengths of these individual sections can be changed to accommodate the melting characteristics of different types of plastics. A metal-wire filter screen is usually placed just before the die to filter out unmelted or congealed resin. The screen, which is replaced periodically, also causes back pressure in the

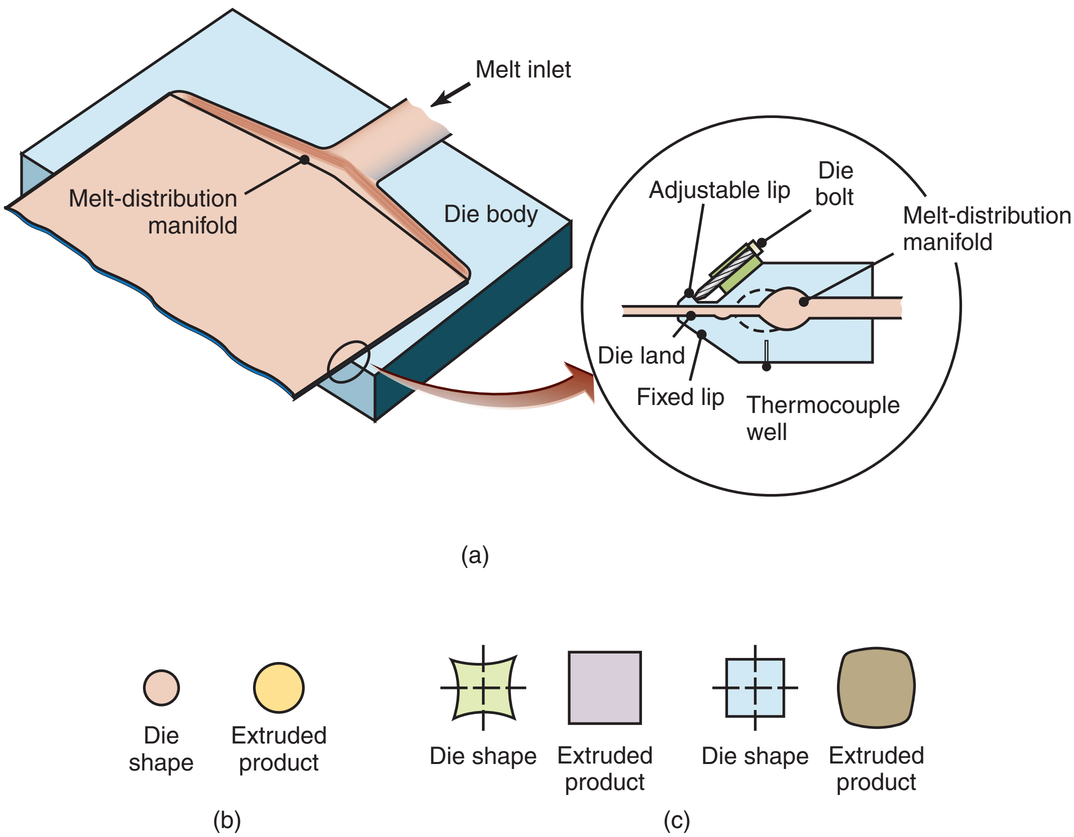


Figure 19.3: Common extrusion die geometries: (a) coat-hanger die for extruding sheet; (b) round die for producing rods; and (c) nonuniform recovery of the part after it exits the die.

barrel, which has to be overcome by the extruder screw. Between the screen and the die is a *breaker plate*, with several small holes in it, which helps improve mixing of the polymer prior to its entering the die. The extruded material is cooled, generally by exposing it to blowing air or by being passed through a water-filled channel (trough).

Controlling the rate and uniformity of cooling is important in order to minimize product shrinkage and distortion. In addition to single-screw extruders, other types include *twin* (two parallel screws side by side) and *multiple screws*, for polymers that are difficult to extrude or require additional blending (see also *reciprocating screw*, Section 19.3).

A typical helical screw metering section is shown in indicating the important parameters that affect the mechanics of polymer extrusion. At any point in time, the molten plastic is in the shape of a helical ribbon, with thickness H and width W; it is conveyed towards the extruder outlet by the rotating screw *flights*. The shape, pitch, and *flight angle* of the helical screw are important parameters, as they affect the flow of the polymer through the extruder. The ratio of the barrel length, L, to its diameter, D, is also important. In typical commercial extruders, the L/D ratio ranges from 5 to 30, and barrel diameters generally are in the range from 25 to 200 mm.

Process Characteristics. Because there is a continuous supply of raw material from the hopper, long products, such as solid rods, sections, channels, sheet, tubing, pipe, and architectural components can be extruded continuously. Complex shapes with constant cross sections can be extruded with relatively inexpensive tooling. Some common die profiles are shown in Fig. 19.3b. Polymers usually undergo much greater and uneven shape recovery than is encountered in metal extrusion (Chapter 15). Because the polymer will *swell* after exiting the die, the openings shown in Fig. 19.3b are smaller than the extruded cross sections. After it has cooled, the extruded product may subsequently be drawn (*sized*) by a puller, then coiled or cut into desired lengths.

The control of processing parameters, such as extruder-screw rotational speed, barrel-wall temperatures, die design, and rate of cooling of the extrudate are all important in order to ensure product quality and uniform dimensional accuracy. *Defects* observed in extruding plastics are similar to those in metal extrusion (Section 15.5). Die shape is important, as it can induce high stresses in the product, causing it to develop surface fractures, as also occur with metals. Other surface defects are *bambooing* and *sharkskin effects*, which are due to a combination of friction at the die–polymer interfaces, elastic recovery, and nonuniform deformation of the outer layers of the product with respect to its bulk during extrusion.

Extruders generally are rated by the diameter, D, of the barrel and the length-to-diameter (L/D) ratio of the barrel. The cost of machinery is on the order of $300,000, including the cost for the equipment for downstream cooling and winding of the extruded product.

19.2.1 Miscellaneous Extrusion Processes

There are several variations of the basic extrusion process for producing a number of different polymer products.

Plastic Tubes and Piping. These products are produced in an extruder, using a *spider die* (Fig. 19.4a; see also Fig. 15.8 for details). For the production of reinforced tubing to withstand higher pressures, woven fiber or wire reinforcements also may be fed through specially designed dies; a typical product is a reinforced plastic water hose. Extrusion of tubes is also a first step for related processes, such as blown film production or extrusion blow molding (Section 19.4).

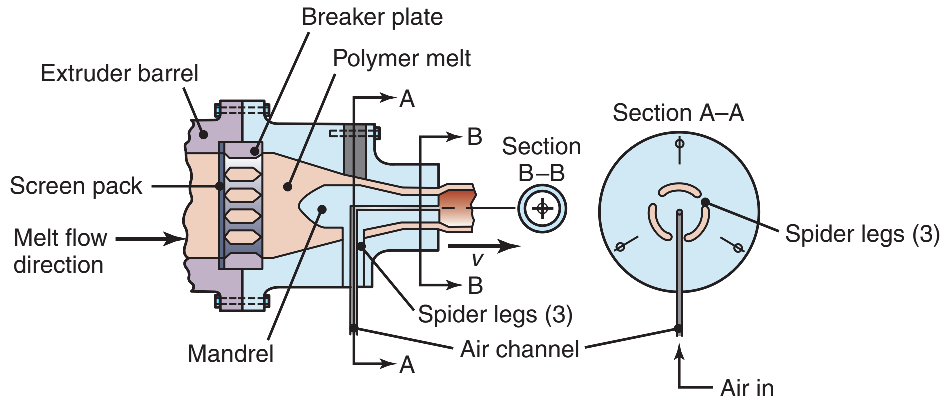
Rigid Plastic Tubing. Extruded by a process in which the die is *rotated*, rigid plastic tubing causes the polymer to be sheared and biaxially oriented during extrusion. As a result, the tube has a higher crushing strength and a higher strength-to-weight ratio than conventionally extruded tubing.

Coextrusion. Shown in Fig. 19.4b, *coextrusion* involves simultaneous extrusion of two or more polymers through a single die. The product cross section thus contains different polymers, each with its own characteristics and functions. Coextrusion produces such shapes as flat sheets, films, and tubes, and is used especially for food packaging where different layers of polymers have different functions. These include: (a) providing inertness for contact with food products, (b) serving as barriers to fluids such as water or oil, and (c) labeling of the product.

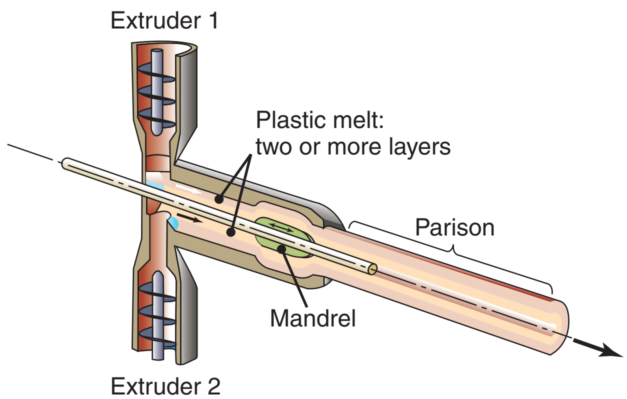
Plastic-coated Electrical Wire. Electrical wire, cable, and strips are simultaneously extruded and coated with plastic. The metal wire is fed into the die opening at a controlled rate with the extruded plastic, depositing a uniform coating on the wire. To ensure proper insulation, extruded wires are checked continuously for their electrical resistance as they exit the die. At the same time, the wire is also marked with ink from a roller for identification purposes. Common *plastic-coated wire paper clips*, with different colors, also are made by this process.

Polymer Sheets and Films. Generally, polymer *sheet* is considered to be thicker than 0.5 mm, and *film* is thinner than 0.5 mm. They are produced using a specially designed flat extrusion die (Fig. 19.3a). Also known as a *coat-hanger die*, it is designed to distribute the polymer melt evenly throughout the width of the die opening. The polymer is extruded by forcing it through the die, after which the extruded sheet is taken up by water-cooled rolls to cool the sheets, then by a pair of rubber-covered pull-off rolls.

Thin Polymer Films. Common *plastic bags* and other thin polymer films are made from **blown film**, which itself is made from a thin-walled tube produced by an extruder. In this process, a tube is extruded vertically (Fig. 19.5), while it is continuously pulled upward and expanded into a balloon shape. Air is continuously



(a)



(b)

Figure 19.4: Extrusion of tubes. (a) Extrusion using a spider die (see also Fig. 15.8) and pressurized air. (b) Coextrusion for producing a parison for a bottle.

blown through the center of the extrusion die until the specified film thickness is reached. Because of the molecular orientation of thermoplastics (Section 7.3), a *frost line* develops on the balloon, which reduces its transparency.

The ratio of the blown diameter to the extruded tube diameter is called the *blow ratio*, which is usually between 1.5 and 5, but is about 3:1 in Fig. 19.5. Note that, as described in Section 2.2.7, the polymer must possess a high strain-rate sensitivity exponent, m, to be successfully blown in this process without tearing (see also Example 9.1).

The balloon is usually cooled by air from a cooling ring, which also acts as a physical barrier to further diametral expansion of the balloon, thus controlling its major dimensions. The cooled tube is then continuously slit lengthwise, becoming *film*. It can also be pinched/welded and cut off, to produce a *plastic bag*. The width of the film produced after slitting can be on the order of 6 m or more.

Plastic Films. Plastic films, especially polytetrafluoroethylene (PTFE; with the trade name of *Teflon*), can be produced by *shaving* the circumference of a solid round plastic billet using specially designed knives. The process is similar to producing *veneer* from a large block of round wood, in a process called **skiving** (see also Section 24.4).

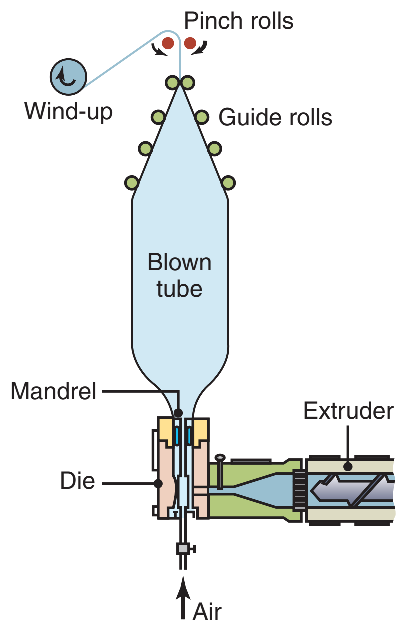


Figure 19.5: Schematic illustration of the production of thin film and plastic bags from tube–first produced by an extruder and then blown by air.

Pellets. Used as raw material for other plastic-processing methods described in this chapter, pellets are made by extrusion. A small-diameter, solid rod is extruded continuously, then chopped into short lengths, called pellets. With some modifications, extruders also can be used as simple melters for other shaping processes, such as injection molding and blow molding.

Example 19.1 Blown Film

Given: A typical plastic shopping bag made from blown film has a lateral dimension (width) of 400 mm.

Find: (a) Determine the extrusion-die diameter. (b) These bags are relatively strong in use. How is this strength achieved?

Solution:

(a) The perimeter of the flat bag is (2)(400) = 800 mm. Since the original cross section of the film is round, the blown diameter should be $\pi D = 800$ thus $D = 255$ mm. Recall that in this process, a tube is expanded from 1.5 to 5 times the extrusion-die diameter. Taking the maximum value of 5, the die diameter is $255/5 = 51$ mm.

(b) After extrusion, the balloon is pulled upward by the pinch rolls. Thus, in addition to diametral stretching from the internal pressure and the attendant molecular orientation, the film is *stretched* and *oriented* in the longitudinal direction. The resulting biaxial orientation of the polymer molecules significantly improves the strength and toughness of the plastic bag.

19.2.2 Production of Polymer Reinforcing Fibers

Polymer fibers have numerous important applications. In addition to their use as reinforcement in composite materials, they are used in a wide variety of consumer and industrial products, including clothing, carpeting, fabrics, rope, and packaging tape.

Most synthetic fibers in reinforced plastics are polymers, extruded through one to several hundred holes in a *spinneret*, resembling a shower head, forming continuous filaments of semisolid polymer. If the polymer is a thermoplastic, it is first melted in an extruder (Section 19.2). Thermosetting polymers also can be formed into fibers, by first dissolving or chemically treating them so that they can be extruded. These operations are performed at high rates and with very high reliability.

As the filaments emerge from the holes, the liquid polymer is first converted to a rubbery state, then it solidifies. This process of extrusion and solidification of continuous filaments is called **spinning**, a term also used for natural textile fibers, such as cotton or wool. There are four methods of spinning fibers: melt, wet, dry, and gel spinning.

1. In **melt spinning** (Fig. 19.6), the polymer melt is extruded through the spinneret, and then it solidifies directly by cooling. A typical spinneret for this operation is about 5 mm thick and has about 50 holes, about 0.25 mm in diameter. The fibers that emerge from the spinneret are cooled by forced-air convection; they are simultaneously pulled so that their final diameter becomes much smaller than the holes of the spinneret. Nylon, olefin, polyester, and PVC are produced in this manner. Melt-spun fibers also can be extruded from the spinneret in various other cross sections, such as trilobal (a triangle with curved sides), pentagonal, octagonal, and hollow shapes. Hollow fibers trap air, providing additional thermal insulation.

2. **Wet spinning**, the oldest process for fiber production, is used for polymers that have been dissolved in a solvent, by submerging the spinneret in a chemical bath. As the filaments emerge, they precipitate in the bath, producing a fiber which is then wound onto a **bobbin** (*spool*). The term "wet" refers to the use of a precipitating liquid bath, resulting in wet fibers; they require drying before they can be used. Acrylic, rayon, and aramid fibers are produced in this manner.

3. **Dry spinning** is used for thermosets dissolved by a fluid. Instead of precipitating the polymer by dilution, as in wet spinning, solidification is achieved by evaporating the solvent fluid in a stream

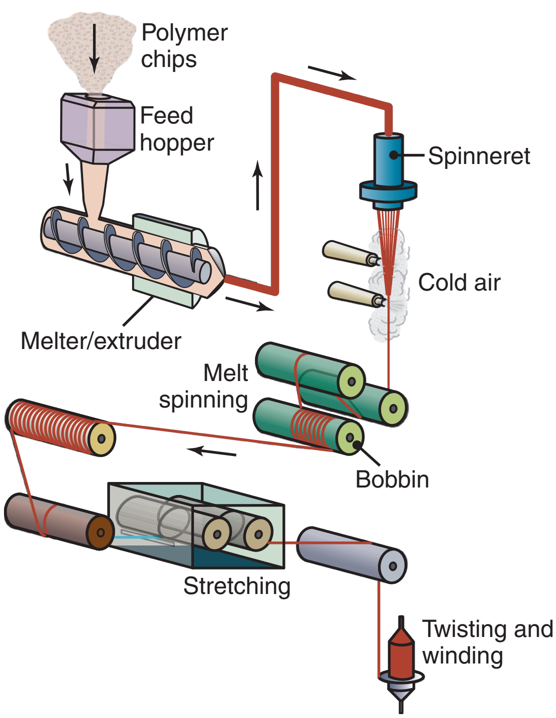


Figure 19.6: The melt-spinning process for producing polymer fibers. The fibers are used in a variety of applications, including reinforcements for composite materials and fabrics for clothing. In the stretching box, the right roll rotates faster than the left roll.

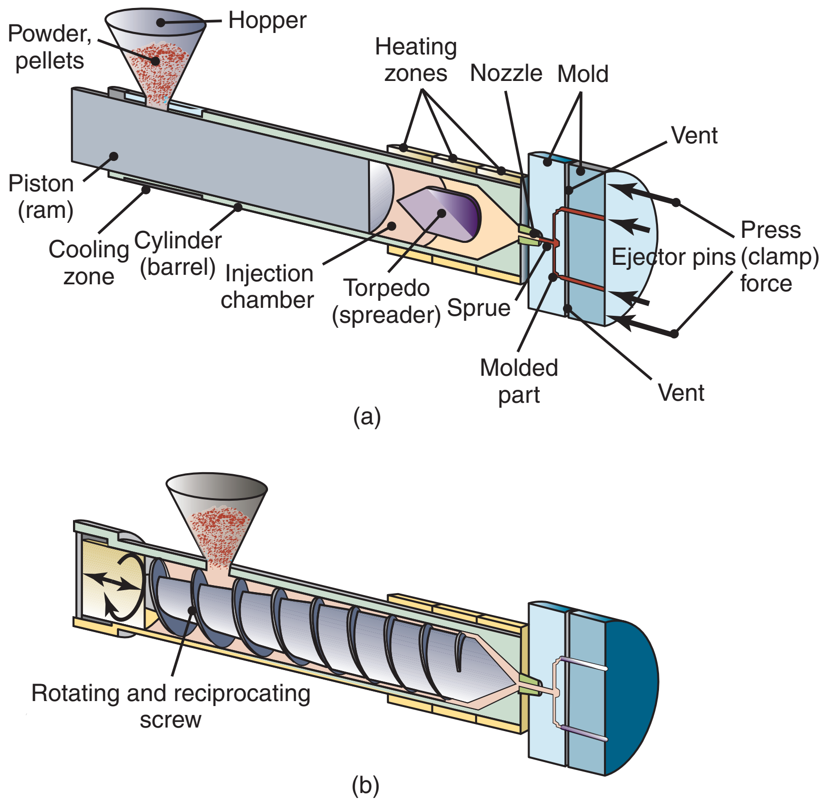


Figure 19.7: Schematic illustration of injection molding with (a) a plunger and (b) a reciprocating rotating screw.

of air or inert gas. The filaments do not come in contact with a precipitating liquid, thus eliminating the need for drying. Acetate, triacetate, polyether-based elastane, and acrylic fibers are made by this process.

4. **Gel spinning** is a special process used to impart high strength or other properties to the fibers. Some polyethylene and aramid fibers are produced by gel spinning. The polymer is not melted completely, or dissolved in a liquid, but the molecules bond together at various points in liquid-crystal form. This operation produces strong interchain forces in the filaments produced, significantly increasing their tensile strength. Moreover, the liquid crystals are aligned along the fiber axis by the strain encountered during extrusion. Thus, the filaments emerge from the spinneret with a high degree of orientation relative to each other, further enhancing their strength. This process is also called *dry–wet spinning*, because the filaments first pass through air and are then cooled further in a liquid bath.

An essential step in the production of most fibers is significant *stretching* to induce orientation of the polymer molecules in the fiber direction; the strain induced can be as high as 800%. Orientation is the main reason for the high strength of the fibers as compared with the polymer in bulk form. Stretching can be done while the polymer is still pliable (just after emerging from the spinneret) or it can be a cold-drawing operation.

Graphite fibers are produced from polymer fibers by *pyrolysis*. In this operation, controlled heat is applied to the polymer fiber (typically polyacrylonitrile, PAN) to drive off all elements except the carbon; temperatures in the range of 1500° to 3000°C are used. The fiber is under tension in order to impart a high degree of orientation in the resulting fiber structure (see also Section 9.2.1 on the properties of graphite fibers and other details).

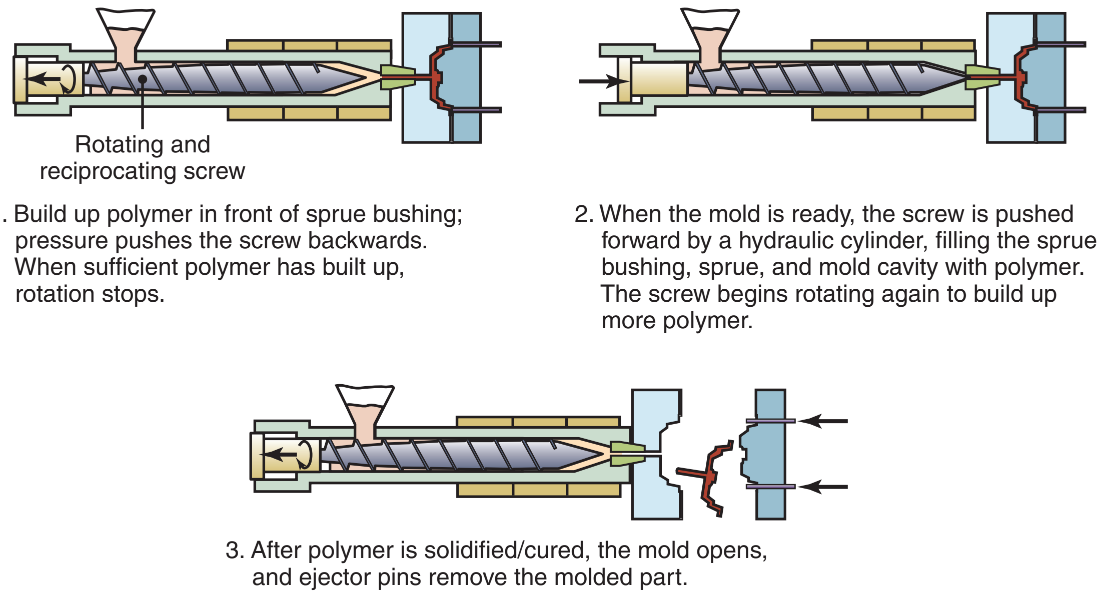


Figure 19.8: Sequence of operations in the injection molding of a part with a reciprocating screw. This process is used widely for numerous consumer and commercial products, such as toys, containers, knobs, and electrical equipment (see Fig. 19.9).

19.3 Injection Molding

Injection molding is similar to hot-chamber die casting (Fig. 19.7 and Section 11.4.5). The pellets or granules are fed into the heated cylinder, and the melt is forced into the mold, either by a hydraulic *plunger* or by a *rotating screw* system similar to that of an extruder. As in plastic extrusion, the barrel (cylinder) is heated externally to promote melting of the polymer. In injection-molding machines, however, a far greater portion of the heat transferred to the polymer is due to frictional heating.

Reciprocating or *plasticating screw* type injection molding machines (Fig. 19.7b) use the sequence of operations shown in Fig. 19.8. As the pressure builds up at the mold entrance, the rotating screw begins to move backwards under pressure to a predetermined distance. This movement controls the volume of material to be injected. The screw then stops rotating, and is pushed forward hydraulically, forcing the molten plastic into the mold cavity. The pressures developed typically range from 70 to 200 MPa.

Several injection-molded products are shown in Fig. 19.9; others include cups, containers, housings, tool handles, knobs, toys, plumbing fixtures, and components for electrical and communications-equipment. For thermoplastics, the molds are kept relatively cool, at about 90°C. Thermoset parts are molded in heated molds, at about 200°C, where *polymerization* and *cross-linking* take place (Section 7.2.1). Elastomers are also injection molded into discrete products using these processes.

After the part has cooled sufficiently (for thermoplastics) or cured (for thermosets), the molds are opened, and the part is ejected. The molds are then closed and the process is repeated automatically. Because the polymer is molten when injected into the mold, complex shapes with good dimensional accuracy can be made. However, because of uneven cooling of the part in the mold, residual stresses will develop.

Molds with moving and unscrewing mandrels also are used in injection molding, as they allow molding of parts with multiple cavities or internal and external threaded features. To accommodate part design, molds may have several components (Fig. 19.10), including runners (such as those used in metal-casting dies, Fig. 11.20), cores, cavities, cooling channels, inserts, knockout pins, and ejectors.

Figure 19.9: Typical products made by injection molding, including examples of insert molding. *Source:* After Rayco Mold and Mfg. LLC.

There are three basic types of molds for injection molding:

1. **Cold-runner, two-plate mold:** This design is the simplest and most common, as shown in Fig. 19.11a.

2. **Cold-runner, three-plate mold** (Fig. 19.11b): The runner system is separated from the part when the mold is opened.

3. **Hot-runner mold** (Fig. 19.11c), also called **runnerless mold**: The molten plastic is kept hot in a heated runner plate.

In cold-runner molds, the solidified plastic remaining in the channels connecting the mold cavity to the end of the barrel must be removed, usually by trimming. Later, this scrap can be chopped and recycled. In hot-runner molds, which are more expensive, there are no gates, runners, or sprues attached to the molded part. Cycle times are shorter, because only the molded part must be cooled and ejected.

Multicomponent injection molding, also called *co-injection* or *sandwich molding*, allows forming of parts that have a combination of various colors and shapes. An example is molding of automobile rear-light covers, made of different materials and colors, such as red, amber, and white. For some parts, printed film also can be placed in the mold cavity, thus providing decoration or labeling during molding.

Insert molding involves *metallic* components, such as screws, pins, and strips, that are placed in the mold cavity prior to injection of polymer, and become an integral part of the molded product (Fig. 19.9). The most common examples of such combinations are hand tools, where the handle is insert molded onto a metal component; other examples include electrical and automotive components.

Overmolding. This is a process for making such products as hinge joints and ball-and-socket joints in one operation, without requiring postmolding assembly. Two different types of plastics are molded to ensure that no bonds will form between the molded halves of the joint, as otherwise their motion would be impeded.

In **ice-cold molding**, the same type of plastic is used for both components of the joint. The operation is carried out in a conventional injection-molding machine and in one cycle. A two-cavity mold is used, with cooling inserts positioned in the area of contact between the first and second molded components of the joint. In this way, no bonds develop between the two pieces, and thus the two components have free movements, as in a hinge or a sliding mechanism.

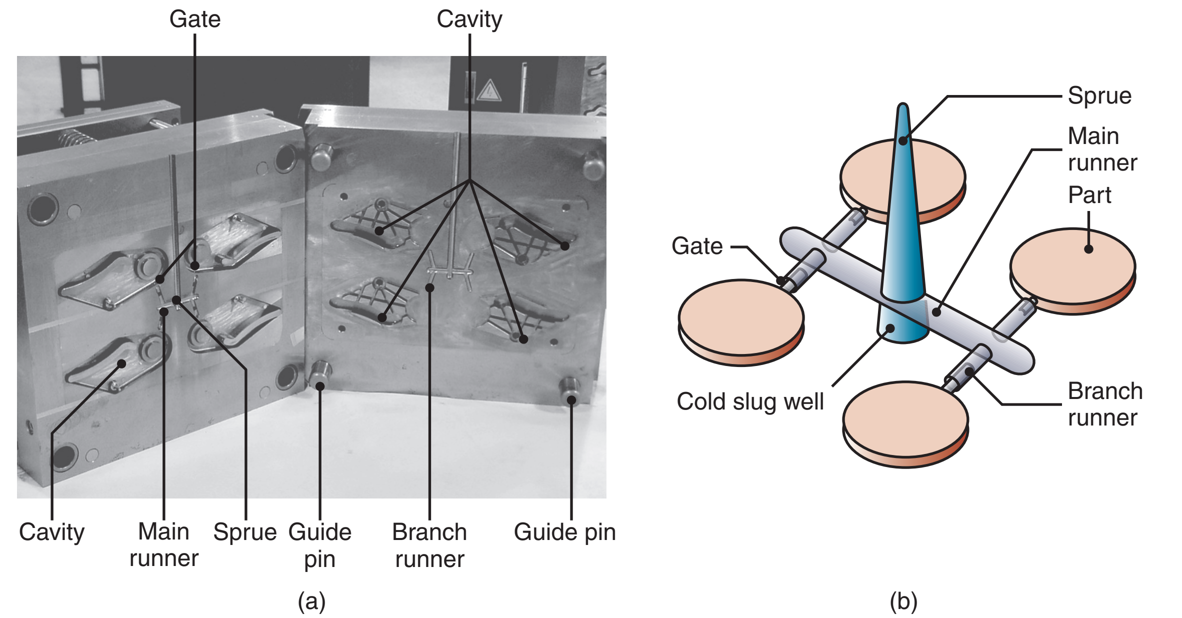


Figure 19.10: Illustration of mold features for injection molding. (a) Two-plate mold with important features identified. (b) Schematic illustration of the features in a mold. *Source:* Courtesy of Tooling Molds West, Inc.

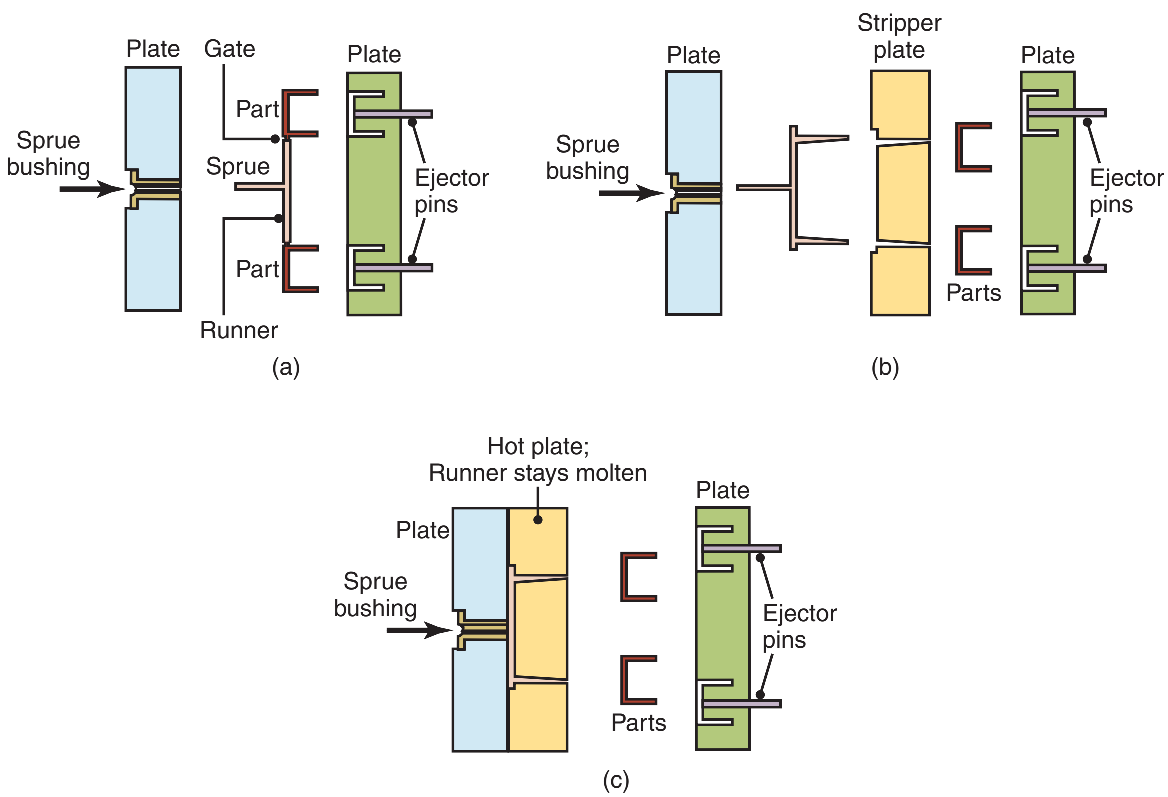


Figure 19.11: Types of molds used in injection molding: (a) two-plate mold; (b) three-plate mold; and (c) hot-runner mold.

Process Capabilities. Injection molding is a high-rate production process and permits good dimensional control. Although most parts generally weigh from about 100 to 600 g, they also can be much heavier. Typical cycle times range from 5 to 60 s, although they can be several minutes for thermosetting materials. Cycle times can be reduced by as much as 50% through the use of *conformal cooling* channels constructed in the mold (see Section 20.6.1 and Fig. 20.18).

Injection molding is a versatile process and capable of producing complex shapes, with good dimensional accuracy. Mold design and control of material flow in the die cavities are important factors in the quality of the product, and in avoiding defects. Because of the basic similarities to metal casting regarding material flow and heat transfer, *defects* observed in injection molding are somewhat similar to those in casting. For example:

- In Fig. 10.15g, the molten metal flows in from two opposite runners, and then meets in the middle of the mold cavity. Thus, a cold shut in casting is equivalent to *weld lines* in injection molding.

- If the runner cross sections are too small, the polymer may solidify prematurely, thus preventing complete filling of the mold cavity. Solidification of the outer layers in thick sections can cause *porosity* or *voids* due to shrinkage, as in the metal parts shown in Fig. 12.2.

- If, for some reason, the dies do not close completely or because of die wear, *flash* will form, in a manner similar to flash formation in impression-die forging (see Figs. 14.6 and 19.17c).

- A defect known as *sink marks* or *pull-in*, similar to that shown in Fig. 19.30c, is observed in injection-molded and cast parts.

- Methods of avoiding defects include proper control of temperatures, pressures, and mold design modifications, using simulation software.

Modeling techniques and *simulation software* continue to be developed for studying optimum gating systems, mold filling, mold cooling, and part distortion. Commercial software tools are available to expedite the design process for molding parts with good dimensional tolerances and other characteristics. The programs take into account such factors as injection pressure, temperature, heat transfer, and the condition of the polymers, and often work within a CAD software environment such as Autodesk and ProEngineer.

Machines. Injection-molding machines are usually horizontal (Fig. 19.12); vertical machines are used for making small, close-tolerance parts and for insert molding. The clamping force on the dies is typically supplied by hydraulic means, although lighter and quieter electrically powered clamps also are used. Modern machines are equipped with microprocessors in a control panel and monitor all aspects of the molding operation.

Injection-molding machines are rated according to the capacity of the mold and the clamping force. In most machines, this force ranges from 0.9 to 2.2 MN (90 to 225 metric tons). The largest machine in operation has a capacity of 75 MN (7500 metric tons), and it can produce parts weighing up to 40 kg. The cost of a 90-metric ton machine ranges from about $60,000 to about $90,000 and of a 270-metric ton machine from about $85,000 to about $250,000. Mold costs typically range from $20,000 to $200,000. Consequently, high-volume production is essential to justify such high expenditures.

The molds generally are made of tool steels, beryllium–copper, or aluminum. They may have multiple cavities, so that more than one part can be molded in one cycle (see also Fig. 11.20). Mold life may be on the order of two million cycles for steel molds, but it can be only about 10,000 cycles for aluminum molds. Some *rapid tooling* approaches (Section 20.10) can produce metal, polymer, or hybrid molds, but they often have a life restricted to a few thousand parts. For low melting point metals, vulcanized rubber molds can be produced, although these are useful only for small batch production runs.

Figure 19.12: A 2.2-MN (225-metric ton) injection-molding machine. The tonnage is the force applied to keep the dies closed during the injection of molten plastic into the mold cavities and hold it there until the parts are cool and stiff enough to be removed from the die. *Source:* Shutterstock/Value ho

Example 19.2 Force Required in Injection Molding

Given: A 2.25-MN injection-molding machine is to be used to make spur gears, 100 mm in diameter and 15 mm thick. The gears have a fine-tooth profile.

Find: How many gears can be injection molded in one set of molds? Does the thickness of the gears affect the force?

Solution: Because of the detail involved (fine gear teeth), assume that the pressure required in the mold cavity will be on the order of 100 MPa. The cross-sectional (projected) area of the gear is $\pi(0.1)^2/4 = 0.00785\ \text{m}^2$. Assuming that the parting plane of the two halves of the mold is in the mid-plane of the gear, the force required is $(100 \times 10^6)(0.00785) = 785$ kN.

Since the capacity of the machine is 2.25 MN, the mold can accommodate $(2250/785)$=2.8 gears. Thus, there is some excess capacity, but this will make sure that the runners and spurs do not cause the molds to open. Therefore, the injection molding machine can be used to produce two gears per cycle. Because it does not influence the cross-sectional area of the gear, gear thickness does not directly affect the pressures involved and thus does not change the answer.

19.3.1 Reaction-injection Molding

In the *reaction-injection molding* (RIM) process, a monomer (Section 7.2) and two or more reactive fluids are forced into a mixing chamber at high speed and a pressure of 10 to 20 MPa, and then into the mold cavity (Fig. 19.13). Chemical reactions rapidly take place in the mold cavity, where the polymer solidifies. Typical polymers used are polyurethane, nylon, and epoxy. Cycle times may range up to about 10 minutes, depending on the materials, part size, and shape complexity.

Major applications of this process include automotive parts, such as bumpers, fenders, steering wheels, and instrument panels, thermal insulation for refrigerators and freezers, water skis, and stiffeners for structural components. Weight of parts made may range up to about 50 kg. Reinforcing fibers, such as

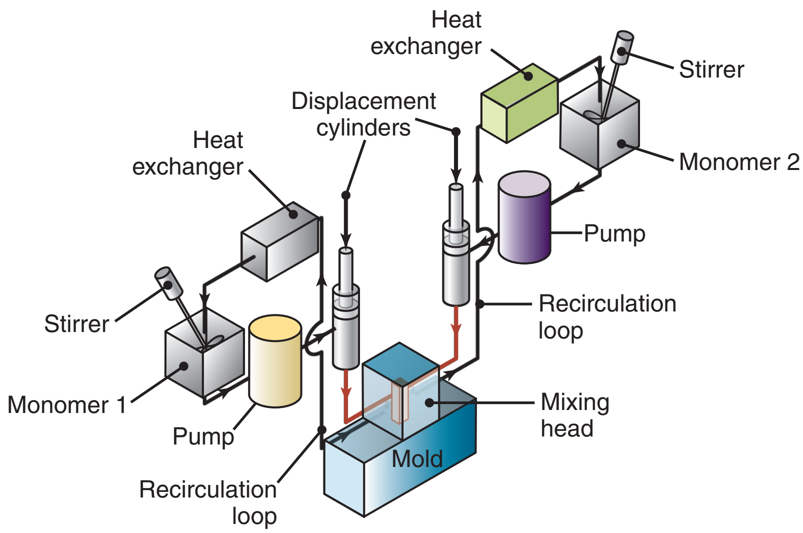


Figure 19.13: Schematic illustration of the reaction-injection molding process. Typical parts made are automotive-body panels, water skis, and thermal insulation for refrigerators and freezers.

glass or graphite, also may be used to improve the strength and stiffness of the part. Depending on the number of parts to be made and part quality required, molds can be made of such materials as steel or aluminum.

19.4 Blow Molding

Blow molding is a modified extrusion and injection-molding process. In **extrusion blow molding**, a tube or preform, usually oriented so that it is vertical, is first extruded. It is then clamped into a mold with a cavity that is much larger than the tube diameter, and blown outward to fill the mold cavity (Fig. 19.14a). Depending on the material, the blow ratio may be as high as 7:1. Blowing usually is done with a hot-air blast, at a pressure ranging from 350 to 700 kPa. Plastic drums, with a volume as large as 2000 liters, can be made by this process. Typical die materials are steel, aluminum, and beryllium copper.

In **injection blow molding**, a short tubular piece (**parison**) is injection molded (Fig. 19.14b) into cool dies. The dies are then opened, and the parison is transferred to a blow-molding die, using an indexing mechanism (Fig. 19.14c). Hot air is injected into the parison, expanding it to contact the walls of the mold cavity. Typical products made are beverage bottles made of polyethylene or polyetheretherketone (PEEK). A related process is **stretch blow molding**, in which the parison is simultaneously expanded and elongated, subjecting the polymer to biaxial stretching and enhancing its properties.

Multilayer blow molding involves the use of coextruded tubes or parisons, and thus permitting the production of a multilayer structure (Fig. 19.4b). A typical example of a product is plastic packaging for food and beverages, having such characteristics as odor and permeation barrier, taste and aroma protection, scuff resistance, printable, and the ability to be filled with hot fluids. Other applications of this process include containers in the cosmetics and the pharmaceutical industries.

19.5 Rotational Molding

Most thermoplastics and some thermosets can be shaped into large, hollow parts by *rotational molding*. A thin-walled metal mold is first made in two pieces (*split-female mold*), designed to be rotated about two

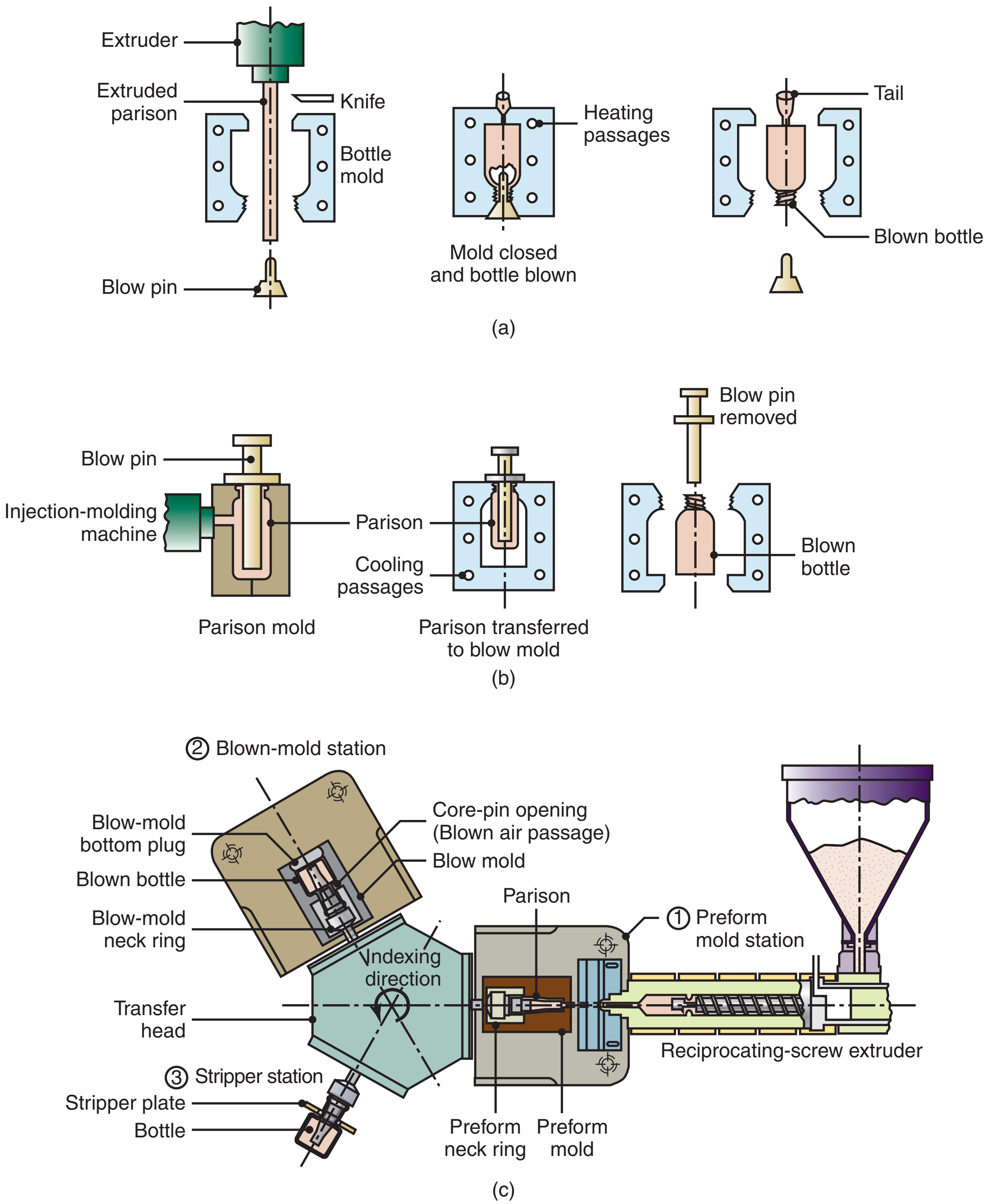


Figure 19.14: Schematic illustrations of (a) the extrusion blow-molding process for making plastic beverage bottles; (b) the injection blow-molding process; and (c) a three-station injection blow-molding machine for making plastic bottles.

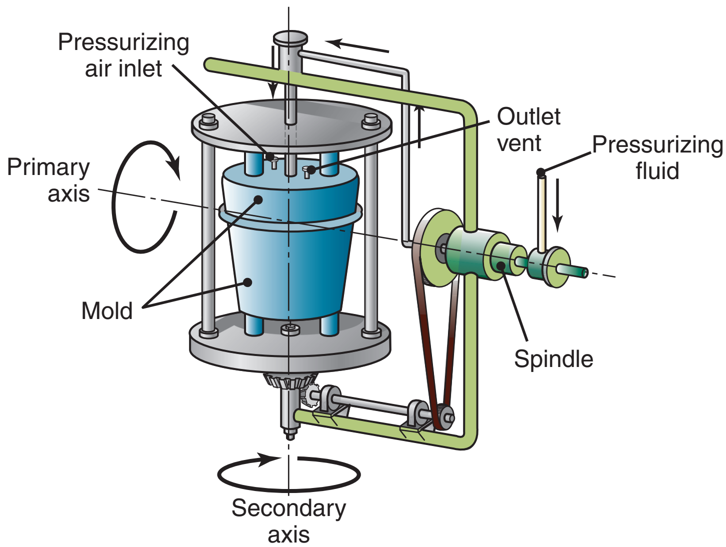


Figure 19.15: The rotational molding (rotomolding or rotocasting) process. Trash cans, buckets, and plastic footballs can be made by this process.

perpendicular axes (Fig. 19.15). For each production cycle, a premeasured quantity of powdered plastic material is placed inside the warm mold. The powder is from a polymerization process that precipitates a powder from a liquid. The mold is then heated, usually in a large oven, and rotated continuously about its two principal axes.

The powder is tumbled against the mold, where the heat fuses the powder without melting it. For thermosetting parts, a chemical agent is added to the powder; cross-linking occurs after the part is formed in the mold. The machines are highly automated, with parts moved by an indexing mechanism, similar to that shown in Fig. 19.14c.

A wide variety of parts are made by rotational molding, such as storage tanks, trash cans, boat hulls, buckets, housings, large hollow toys, carrying cases, and footballs. The outer surface finish of the part is a replica of the surface finish of the inside mold walls. Various metallic or plastic inserts also may be molded integrally into the parts.

In addition to powders, liquid polymers (**plastisols**) can be used in rotational molding–PVC plastisols being the most common material. In this operation, called **slush molding** or *slush casting*, the mold is heated and rotated simultaneously. Due to the tumbling action, the polymer is forced against the inside walls of the mold, where it melts and coats the mold walls. The part is cooled while it is still rotating, and removed by opening the mold. Parts made are typically thin-walled products, such as boots, buckets for aerial cranes, and toys.

Process Capabilities. Rotational molding can produce parts with complex, hollow shapes with wall thicknesses as small as 0.4 mm. Cycle times are longer than in other molding processes. Quality-control considerations usually involve accurate weight of the powder, proper rotational speed of the mold, and temperature–time relationships during the oven cycle.

19.6 Thermoforming

Thermoforming is a process for forming thermoplastic sheets and films over a mold through the application of heat and pressure (Fig. 19.16). A sheet is first clamped and heated to the *sag point* (above the *glass-transition temperature*, T_g, of the polymer; Table 7.2), usually by radiant heating, and then forced against the

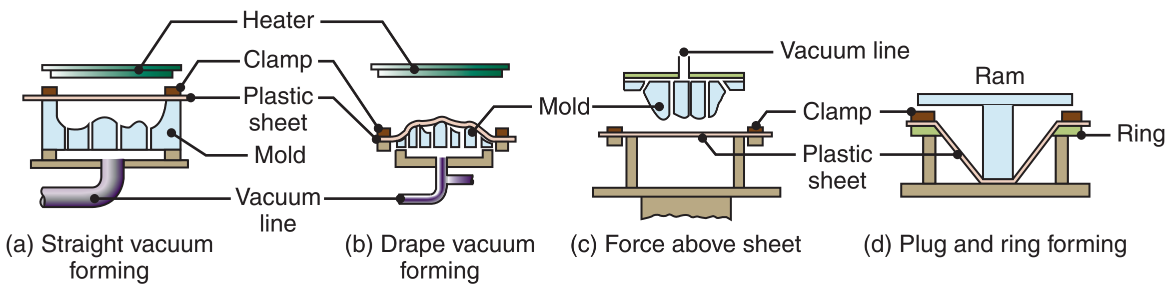


Figure 19.16: Various thermoforming processes for a thermoplastic sheet. These processes are commonly used in making advertising signs, cookie and candy trays, panels for shower stalls, and packaging.

mold surfaces by applying a vacuum or air pressure. The sheets are available as a coiled strip or as blanks, with lengths and widths of various sizes.

The mold is generally at room temperature, thus the shape produced becomes set upon contact with the mold. Because of the low strength of the materials shaped, the pressure difference caused by a vacuum is usually sufficient for forming. Thicker and more complex parts require air pressure, ranging from 100 to 2000 kPa, depending on the type of polymer and its thickness. Variations of the basic thermoforming process are shown in Fig. 19.16.

Process Capabilities. Typical parts made by thermoforming are packaging, trays for cookies and candy, advertising signs, refrigerator liners, appliance housings, and panels for shower stalls. Parts with openings or holes cannot be formed by this process because the pressure difference cannot be maintained during forming. Because thermoforming is basically a combination of *drawing* and *stretching* operations, much like some sheet-metal forming processes, the material must exhibit high, uniform elongation, as otherwise it will neck and tear. Thermoplastics have high capacities for uniform elongation, by virtue of their high strain-rate sensitivity exponent, m (Section 2.2.7).

Molds for thermoforming usually are made of aluminum, because high strength is not required and the machinability and thermal conductivity of aluminum is advantageous; thus, tooling is relatively inexpensive. The molds have small through-holes in order to aid vacuum forming. These holes typically are less than 0.5 mm in diameter, as otherwise they may leave circular marks on the parts being formed. *Defects* encountered in thermoforming include (a) tearing of the sheet during forming, (b) excessive nonuniform wall thickness, (c) improperly filled molds, (d) poor part definition, and (e) lack of surface details.

19.7 Compression Molding

In *compression molding*, a preshaped charge of polymer, a premeasured volume of powder, or a viscous mixture of liquid-resin and filler material is placed directly into a heated mold cavity, which typically is around 200°C but can be much higher. Forming is done under pressure from a plug or from the upper half of the die (Fig. 19.17), thus the process is somewhat similar to closed-die forging of metals (Section 14.3). Polymers also can be molded by cold or hot *isostatic pressing* (Section 17.3.2).

Pressures range from about 10 to 150 MPa. As can be seen in Fig. 19.17, there is a flash that forms. Typical parts made are fittings, electrical and electronic components, washing-machine agitators, and housings. Fiber-reinforced parts, with chopped fibers, also are formed by this process.

Compression molding is mainly used with thermosetting plastics, with the original material being in a *partially polymerized* state; thermoplastics and elastomers are also processed by compression molding. Curing times are in the range of 0.5 to 5 minutes, depending on the material and part thickness and its shape. The thicker the material, the longer the time required to cure.

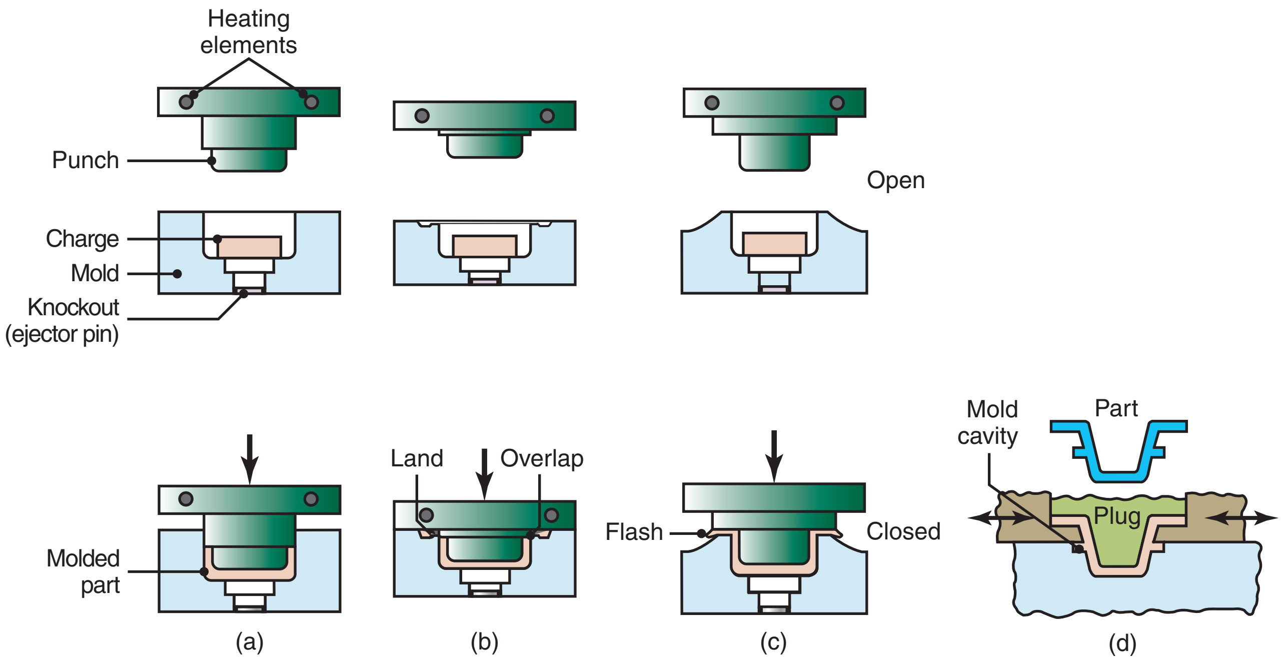


Figure 19.17: Types of compression molding–a process similar to forging: (a) *positive*; (b) *semipositive*; and (c) *flash*, in which the flash is later trimmed off. (d) Die design for making a compression-molded part with external undercuts.

Process Capabilities. Three types of compression molds are made:

- *Flash type*, for shallow or flat parts
- *Positive type*, for high-density parts
- *Semipositive type*, for quality production.

Undercuts in parts are not recommended; however, dies can be designed to open sideways (Fig. 19.17d) to allow removal of the molded part. In general, the complexity of parts produced is less than that from injection molding, but the dimensional control is better. Surface areas of compression-molded parts may range up to about 2.5 m^2. Because of their relative simplicity, dies for compression molding generally are less costly than those for injection molding. Die materials typically are tool steels; they may be chrome plated or polished for improved surface finish of the molded part.

19.8 Transfer Molding

Transfer molding is a further development of the compression molding process. The uncured thermosetting resin is placed in a heated transfer pot or chamber (Fig. 19.18), where it is heated, then injected into heated closed molds. Depending on the type of machine, a ram, plunger, or rotating-screw feeder forces the polymer to flow through the narrow channels into the mold cavity at pressures up to 300 MPa. The viscous flow generates considerable heat, raising the temperature and homogenizing the polymer; curing takes place by cross-linking. Because the resin is in a molten state as it enters the molds, the complexity of the parts made and their dimensional control approach those of injection molding.

Process Capabilities. Typical parts made by transfer molding are electrical connectors, electronic components, rubber and silicone parts, and encapsulation of microelectronic devices. The process is especially

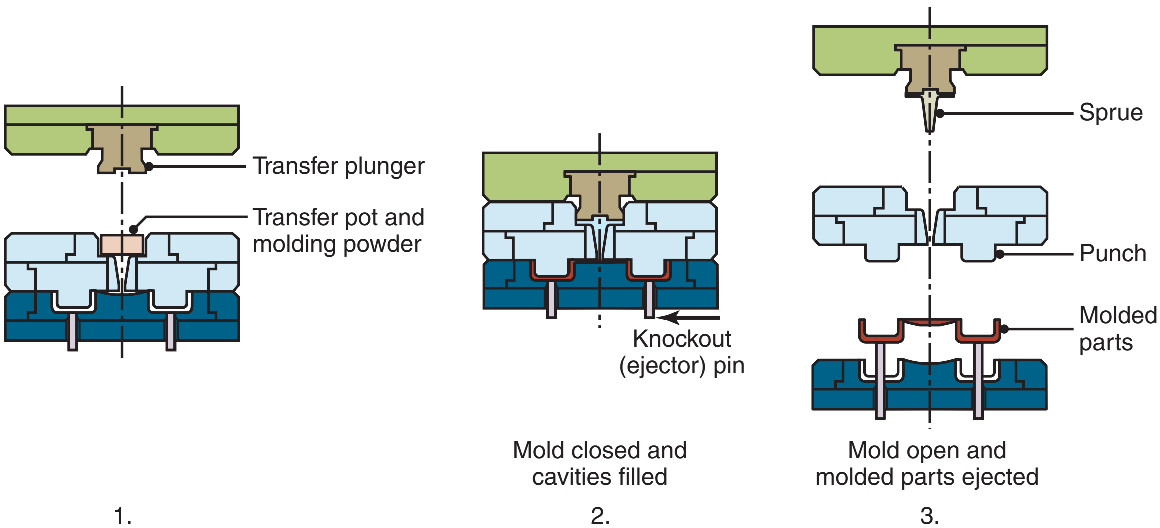


Figure 19.18: Sequence of operations in transfer molding for thermosetting plastics. This process is suitable particularly for intricate parts with varying wall thickness.

suitable for intricate shapes with varying wall thicknesses. The molds tend to be more expensive than those for compression molding, and some excess material is left in the channels of the mold during filling, which is later removed.

19.9 Casting

Some thermoplastics, such as nylons and acrylics, and thermosetting plastics, such as epoxies, phenolics, polyurethanes, and polyester, can be *cast* into a variety of shapes, using either rigid or flexible molds (Fig. 19.19). Compared to other methods of processing plastics, casting is a slow but simple and inexpensive process. Also, the polymer must have sufficiently low viscosity in order to flow easily into the mold. Typical parts cast are gears (especially nylon), bearings, wheels, thick sheets, lenses, and components requiring resistance to abrasive wear (Section 33.5).

In conventional casting of thermoplastics, a mixture of monomer, catalyst, and various additives (*activators*) is heated to above its melting point, T_m, and poured into the mold. The part is shaped after polymerization takes place at ambient pressure. Intricate shapes can be produced using *flexible molds*, which are

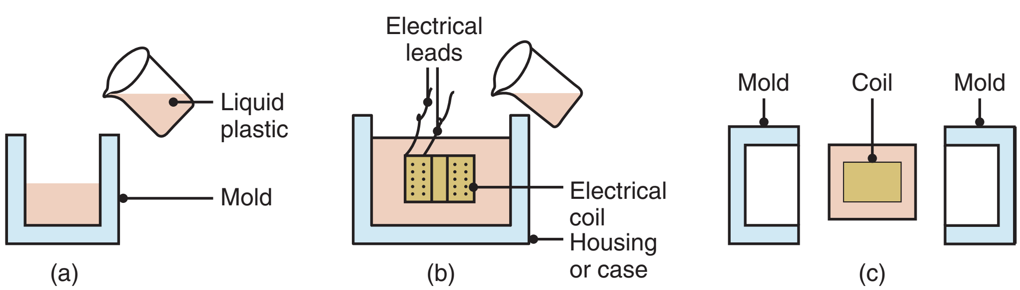


Figure 19.19: Schematic illustration of (a) casting, (b) potting, and (c) encapsulation processes for plastics and electrical assemblies, where the surrounding plastic serves as a dielectric.

then peeled off and reused. As with metals, thermoplastics may be continuously cast (Section 5.4), with the polymer being carried over continuous stainless-steel belts and polymerized by external heat.

Centrifugal Casting. This process is similar to centrifugal metal casting, described in Section 11.4.6, and is used with thermoplastics, thermosets, and reinforced plastics with short fibers.

Potting and Encapsulation. In a variation of casting, *potting* and *encapsulation* involve casting the plastic material (typically a liquid resin, such as epoxy) around an electrical component, such as a transformer, to embed it in the plastic. *Potting* (Fig. 19.19b) is carried out in a housing or case, which becomes an integral part of the component. In *encapsulation* (Fig. 19.19c), the component is coated with a layer of the plastic, surrounding it completely and then solidifying.

In both of these processes, the plastic material can serve as a *dielectric* (nonconductor); consequently, it must be free of moisture and porosity, which would require processing in a vacuum. Mold materials may be metal, glass, or various polymers. Small structural members, such as hooks and studs, may be encapsulated partially, by dipping them in a hot thermoplastic. A wide variety of polymer colors and hardnesses are available. These process are particularly important in the electrical and electronics industries.

19.10 Foam Molding

Styrofoam cups, food containers, thermally insulating blocks, and shaped packaging materials, such as for shipping appliances, computers, and electronics, are made by *foam molding*, using expandable **polystyrene beads** as the raw material. The parts made have a **cellular structure**, wherein they may have *open and interconnected* porosity (for polymers with low viscosity) or have *closed cells* (for polymers with high viscosity).

There are several techniques that are used in foam molding. In the basic operation, polystyrene beads, obtained by polymerization of styrene monomer, are placed in a mold with a blowing agent, typically pentane (a volatile hydrocarbon) or inert gas (nitrogen), and are exposed to heat, usually by steam. The beads expand, to as much as 50 times their original size, and take the shape of the mold cavity. The amount of expansion can be controlled by varying the temperature and time. Various other particles, including hollow glass beads or plastic spheres, may be added to impart specific structural characteristics to the foam produced.

Polystyrene beads are available in three sizes: (a) small, for cups with a finished part density of about 50 kg/m^3, (b) medium, for molded shapes, and (c) large, for molding insulating blocks, with a finished part density of about 15 to 30 kg/m^3; they can all then be cut to size. The bead size selected also depends on the minimum wall thickness of the product; the thinner the part, the smaller is the size. Beads can be colored prior to expanding them, making a part integrally colored. Both thermoplastics and thermosets can be used for foam molding, but thermosets are in a liquid-processing form, and are thus in a condition similar to that of polymers in reaction-injection molding (Section 19.3.1).

A common method of foam molding is using *pre-expanded polystyrene beads*, in which the beads are expanded *partially* by steam, hot air, hot water, or an oven, in an open-top chamber. The beads are then placed in a storage bin and allowed to stabilize for a period of 3 to 12 hours; they then can be molded into desired shapes.

Structural Foam Molding. This process is used to make plastic parts that have a *solid outer skin* and a *cellular core structure* such as inexpensive furniture components, computer and business-machine housings, and moldings, thus replacing more expensive wood moldings. In this process, thermoplastics are first mixed with a blowing agent (usually an inert gas such as nitrogen), then injection molded into cold molds of desired shapes. The rapid cooling next to the cold mold surfaces produces a skin that is rigid, which can be as much as 2 mm thick; the core of the part is cellular in structure. The overall density of the part made can be as low as 40% of the density of the solid plastic. With a rigid skin and a less dense bulk, molded parts thus have a high stiffness-to-weight ratio (see also Fig. 3.2).

Polyurethane Foam Processing. Such products as furniture cushions and insulating blocks are made by *polyurethane foam processing*. The operation starts with mixing two or more components; chemical reactions then take place after the mixture is (a) poured into molds of various shapes or (b) sprayed over surfaces with a spray gun, thus providing sound and thermal insulation. Several types of low-pressure and high-pressure machines are available, having computer controls to ensure proper mixing. The mixture solidifies into a cellular structure, the characteristics of which depend on the type and proportion of the components used.

19.11 Cold Forming and Solid-phase Forming

Processes for cold working of metals, such as rolling, closed-die forging, coining, deep drawing, and rubber forming, described in Part III, also can be used to shape thermoplastics at room temperature (*cold forming*). Typical materials formed are polypropylene, polycarbonate, ABS, and rigid PVC. Important considerations are that (a) the polymer must be sufficiently ductile at room temperature, thus polystyrenes, acrylics, and thermosets cannot be formed, and (b) its deformation must be nonrecoverable, in order to minimize springback and creep of the shaped part.

The advantages of cold forming over other methods of shaping plastics are:

- Strength, toughness, and uniform elongation are increased
- Polymers with high molecular weight (Section 7.2) can be used to make parts with superior properties
- Cycle times generally are shorter than those in molding processes
- Forming speeds are not affected by part thickness, because, unlike other processing methods, there is no heating or cooling involved.

Solid-phase Forming. Also called *solid-state forming*, this process is carried out at a temperature 10° to 20°C below the melting temperature of the plastic for a crystalline polymer. The shaping operation takes place while the polymer is still in a solid state. The main advantages of this process over cold forming are that forming forces and springback are lower. These processes are not used as widely as hot-processing methods and generally are restricted to special applications.

19.12 Processing Elastomers

Recall from Section 7.9 that, in terms of its processing characteristics, a thermoplastic *elastomer* is a *polymer*; in terms of its function and performance, it is a *rubber*. The raw material is basically a compound of rubber with various additives and fillers. Additives include carbon black, an element that enhances elastomer properties such as tensile and fatigue strength, abrasion and tear resistance, ultraviolet protection, and resistance to chemicals.

These materials are then mixed to break them down and to lower their viscosity; the mixture is subsequently **vulcanized**, using sulfur as the vulcanizing agent. This compound is then ready for such processes as calendering, extrusion, and molding, which may include placing fiber reinforcements. During final processing, the part becomes cross-linked, imparting the desirable elastic properties of rubber products, ranging from rubber boots to pneumatic tires.

Thermoplastic elastomers are commonly shaped by extrusion or injection molding, extrusion being the more economical and faster process; they also can be shaped by blow molding or thermoforming. Thermoplastic polyurethane, for example, can be shaped by any of the conventional methods. It also can be blended with thermoplastic rubbers, polyvinyl chloride compounds, ABS, and nylon to impart specific properties.

The temperatures for elastomer extrusion are typically in the range from 170° to 230°C, and for molding are up to 60°C. Dryness of the materials is important for product integrity. Reinforcements can be used in conjunction with extrusion, to impart greater strength. Examples of extruded elastomer products are tubing,

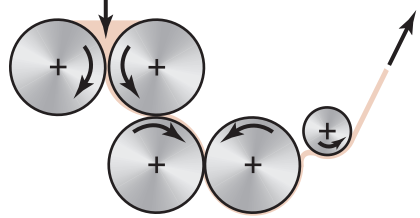


Figure 19.20: Schematic illustration of calendering. Sheets produced by this process subsequently are used in thermoforming. The process also is used in the production of various elastomer and rubber products.

hoses, moldings, and inner tubes. Injection-molded elastomer products cover a broad range of applications, including components for automobiles and appliances.

Rubber and some thermoplastic sheets are made by the **calendering** process (Fig. 19.20), wherein a warm mass of the compound is fed into a series of rolls, and **masticated** (compressed and kneaded into a pulp). Their thickness is typically 0.3 to 1 mm, but can be made less by stretching the material. The calendered rubber then may be molded into various products, such as tires and belts for machinery. Rubbers and thermoplastics also may be formed over both surfaces of a tape, paper, fabric, or plastics, thus making them permanently *laminated*. Roll surfaces may also be textured to produce a rubber sheet with various patterns and designs.

Discrete rubber products, such as gloves, balloons, and swim caps, are made by repeatedly *dipping* or **dip molding** a solid metal form, such as in the shape of a hand for making gloves, into a liquid compound; the liquid that coats the mold takes the shape of the form. A typical compound is *latex*, a milk-like sap obtained from the inner bark of a tropical tree. The compound is then vulcanized (cross-linked), usually in steam, and then stripped from the form, becoming a discrete product.

19.13 Processing Polymer-matrix Composites

As described in Chapter 9, *polymer-matrix composites* (PMC), also called **reinforced plastics**, are *engineered materials*, with unique mechanical properties, especially high strength-to-weight ratio, stiffness-to-weight ratio, fatigue strength, creep resistance, and directional properties. Because of their complex structure, however, reinforced plastics require special methods to shape them into consumer and industrial products (Fig. 19.21).

Fabrication to ensure reliable properties in composite parts and structures can be challenging, particularly over the long range of their service life because of the presence of two or more different types of materials. The matrix and the reinforcing fibers in the composite have, *by design*, very different properties and characteristics; consequently, they have different responses to the methods of processing (Section 9.2).

The several steps required for manufacturing reinforced plastics and the time and care involved make processing costs very high. This situation has necessitated the proper assessment and integration of design and manufacturing processes (*concurrent engineering*), in order to take advantage of the unique properties of these composites. The approach is to minimize manufacturing costs while maintaining long-range product integrity, reliability, and production rate.

19.13.1 Fiber Impregnation

For good bonding between the reinforcing fibers and the polymer matrix, and to protect them during handling, fibers are surface treated by impregnation (*sizing*). When impregnation is carried out as a separate step, the resulting partially cured sheets are called by various terms, as described below.

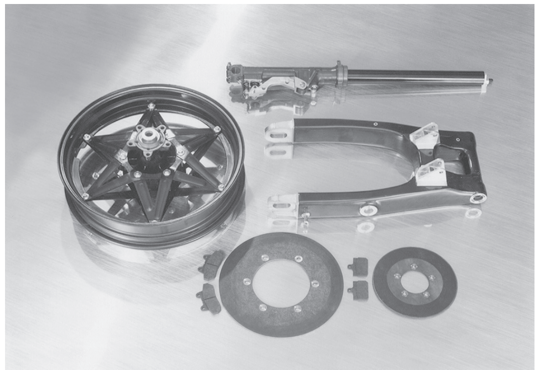


Figure 19.21: Reinforced-plastic components for a Honda motorcycle. The parts shown are front and rear forks, a rear swing arm, a wheel, and brake disks.

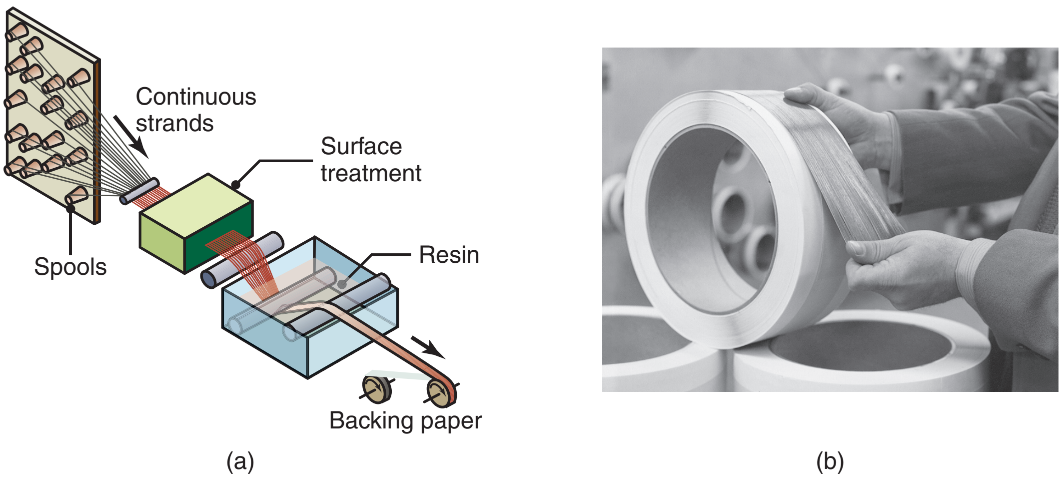


Figure 19.22: (a) Manufacturing process for polymer-matrix composite tape. (b) Boron-epoxy prepreg tape. These tapes are then used in making reinforced plastic parts and components with high strength-to-weight and stiffness-to-weight ratios, particularly important for aircraft and aerospace applications and sports equipment. *Source:* (a) After T.W. Chou, R.L. McCullough, and R.B. Pipes. (b) Courtesy of Avco Specialty Materials/Textron.

Prepregs. In a typical procedure for making fiber-reinforced plastic *prepregs* (meaning *pre-impregnated with resin*), the continuous fibers are first aligned and subjected to a surface treatment to enhance their adhesion to the polymer matrix (Fig. 19.22a). They then are coated by dipping them in a resin bath and are made into a *tape* (Fig. 19.22b), typically in widths of 75 to 150 mm. Individual segments of prepreg tape are then cut and assembled into *laminated structures*, such as for the horizontal stabilizer for fighter aircraft.

Typical composites made from prepregs are flat or corrugated architectural paneling, panels for construction and electrical insulation, and structural components of aircraft requiring good property retention over a period of time and under adverse conditions. These requirements include fatigue strength under hot or wet conditions, typically encountered by military aircraft.

Because the process of laying prepreg tapes is a time-consuming and labor-intensive operation, highly automated *computer-controlled tape-laying machines* have been built for this purpose. Prepreg tapes are automatically cut from a reel and placed on a mold in the desired patterns, with much better dimensional control than can be achieved by hand. The layout patterns can be modified easily, quickly and with high reliability by computer control.

Sheet-molding Compound. In making *sheet-molding compound* (SMC), continuous strands of reinforcing fibers are first chopped into short fibers (Fig. 19.23), and deposited in random orientations over a layer of resin paste. Generally, the paste is a polyester mixture, which may contain fillers, such as various mineral powders, and is carried on a polymer film, such as polyethylene. A second layer of resin paste is then deposited on top, and the sheet is pressed between rolls.

The product is then shaped into rolls or is placed into containers, in several layers, and stored until it has undergone a maturation period and has reached the desired viscosity. The maturing process takes place under controlled conditions of temperature and humidity, and usually taking about one day.

The molding compounds should be stored at a temperature sufficiently low to delay curing. They have a limited shelf life, usually about 30 days, and hence must be processed within this period. Alternatively, the resin and the fibers can be mixed together only at the time they are to be placed into the mold.

Bulk-molding Compound. *Bulk-molding compounds* (BMC) are in the shape of billets (hence the term "bulk"), and generally are up to 50 mm in diameter. They are made in the same manner as SMCs and are extruded to produce a bulk form. When processed into products, BMCs have flow characteristics that are similar to those of dough, hence they also are called *dough-molding compounds* (DMC).

Thick-molding Compound. *Thick-molding compounds* (TMC) combine the lower cost of BMCs with the higher strength of SMCs. They are generally injection molded, using chopped fibers of various lengths. One application is in electrical components because of the high dielectric strength of TMCs.

19.13.2 Molding of Reinforced Plastics

There are several molding processes used for reinforced plastics.

Compression Molding. The material is placed between two molds, and pressure is applied. The molds may be either at room temperature or heated to accelerate hardening of the part. The material may be a bulk-molding compound, powder, or it may be an uncured thermoset with a dough-like consistency. Generally, it is molded into the shape of a log, which subsequently is cut or sliced into the desired shape. Fiber lengths typically range from 3 to 50 mm, although longer fibers of 75 mm also may be used.

Sheet-molding compounds also can be processed by compression molding; they are similar to bulk-molding compounds, except that the resin–fiber mixture is laid between plastic sheets to make a sandwich that can be handled easily. The sheets have to be removed before placing the SMC in the mold.

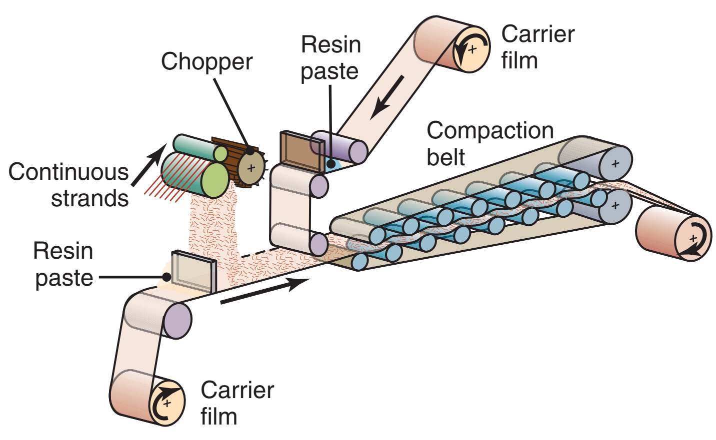


Figure 19.23: Schematic illustration of the manufacturing process for producing fiber-reinforced plastic sheets. The sheet still is viscous at this stage and later can be shaped into various products. *Source:* After T.-W. Chou, R.L. McCullough, and R.B. Pipes.

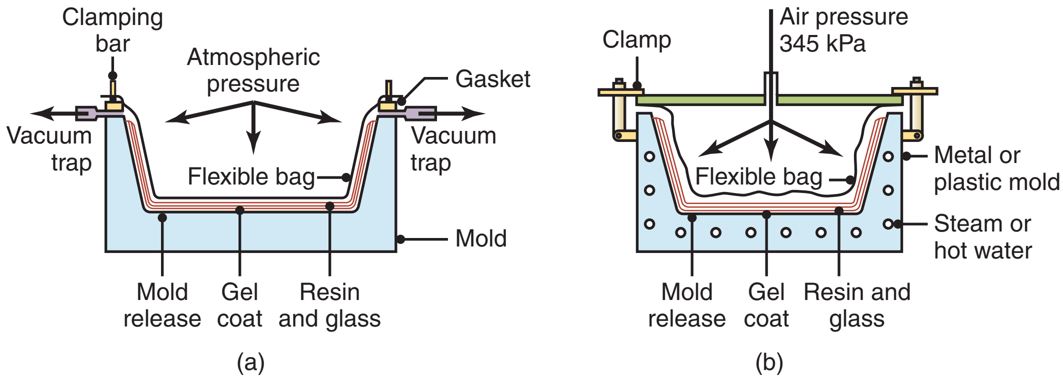


Figure 19.24: Schematic illustration of (a) vacuum-bag forming and (b) pressure-bag forming. These processes are used in making discrete reinforced plastic parts. *Source:* After T.H. Meister.

Vacuum-bag Molding. In this process (Fig. 19.24a), prepregs are laid in a mold to form the desired shape. The pressure required to shape the product and to ensure good bonding is applied by covering the layup with a plastic bag and creating a vacuum. Curing takes place at room temperature or in an oven.

A variation of this process is **pressure-bag molding** (Fig. 19.24b). A flexible bag is placed over the resin and reinforcing fiber mixture, and pressure is then applied over the mold, at a range typically from 200 to 400 kPa. If higher heat and pressure are needed to make parts with higher density and fewer voids, the entire assembly is placed into an *autoclave* (a chamber under heat and pressure).

Care should be exercised to maintain fiber orientation if specific directional properties are desired. With chopped-fiber materials, no specific orientation is intended. In order to prevent the resin from sticking to the vacuum bag, and also to facilitate removal of excess resin, several sheets of various materials, called *release cloth* or *bleeder cloth*, are placed on top of the prepreg sheets. Molds can be made of metal, usually aluminum, but more often they are made from the same resin (with reinforcement) as the material to be cured. This practice eliminates any difficulties caused by the difference in thermal expansion between the mold and the part.

Contact Molding. Also referred to as *open-mold processing*, this is a series of processes that uses a single male or female mold, made of such materials as reinforced plastics, wood, metal, or plaster (Fig. 19.25). The operation is a wet method, in which the materials are applied in layers, and the reinforcement is impregnated with the resin at the time of molding. Contact molding is used in making *laminated products*, with high surface area–to-thickness ratios, hence the process is also called *contact lamination*. Typical examples of products made are backyard swimming pools, boat hulls, automotive-body panels, tub and shower units, and housings.

The simplest method of contact molding is **hand layup**. The materials are first placed in proper order (resins and reinforcements), brushed with a liquid monomer, and shaped in the mold by hand using a roller (Fig. 19.25a). The squeezing action of the roller expels any trapped air bubbles while compacting the part. The reinforcements placed in the mold may consist of various shapes, including prepregs; their orientation in the final product can thus be controlled.

In **spray layup**, molding is done by spraying the materials into the mold. As seen in Fig. 19.25b, both the resin and the chopped fibers are sprayed over the mold surfaces. Rolling the deposited materials to remove any porosity may be necessary, as is done in layup. Because the chopped fibers have random orientations, directional properties cannot be imparted in products made by spray layup. Note also that only the mold-side surface of the formed part is smooth, because they have been in contact with the mold surfaces.

Resin-transfer Molding. This process is based on transfer molding of polymers (Section 19.8). A resin is fist mixed with a catalyst, and is then forced by a piston-type, positive-displacement pump into the mold

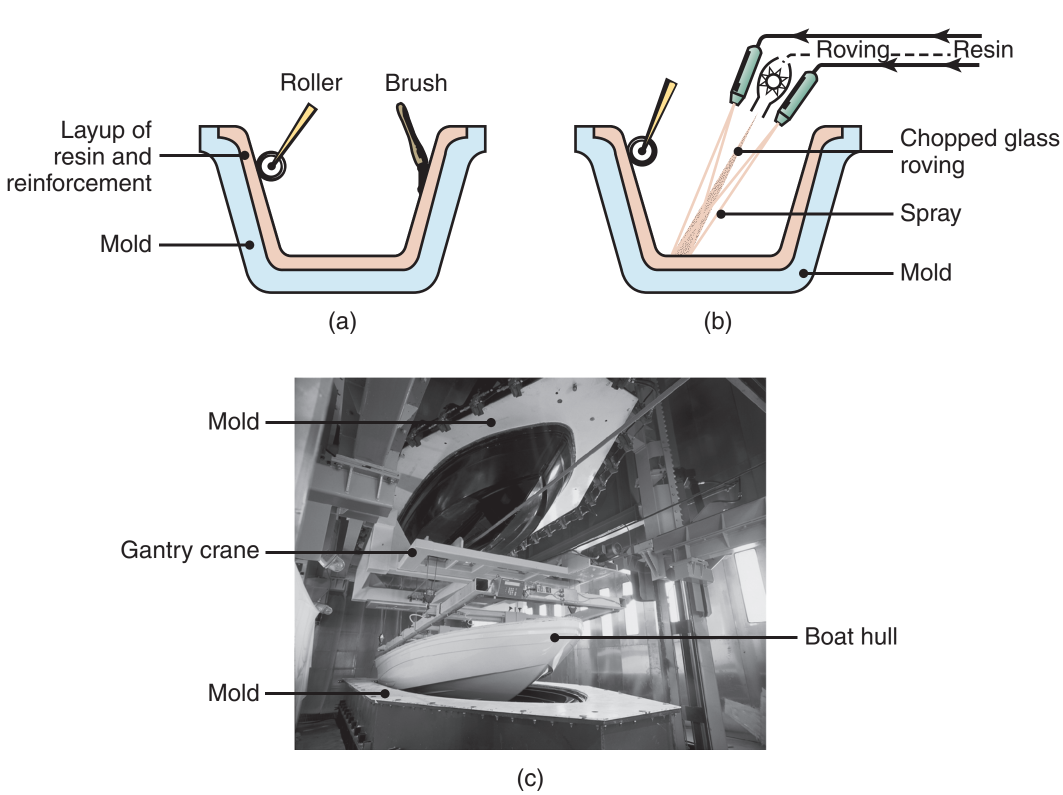


Figure 19.25: Manual methods of processing reinforced plastics: (a) hand layup and (b) spray layup. Note that, even though the process is slow, only one mold is required. The figures show a female mold, but male molds are used as well. These methods also are called *open-mold processing*. (c) A boat hull being made by these processes. *Source:* Courtesy of VEC Technology, LLC.

cavity, which has been filled with a fiber reinforcement. The process is a viable alternative to hand layup, spray layup, or compression molding for low- or intermediate-volume production.

Transfer/Injection Molding. This is an automated operation that combines the processes of compression-molding, injection-molding, and transfer-molding. This combination has the good surface finish, dimensional stability, and mechanical properties obtained in compression molding; it also has the high-automation capability and low cost of injection molding and transfer molding.

19.13.3 Filament Winding, Pultrusion, and Pulforming

Filament Winding. This is a process in which the resin and the fibers are combined at the time of curing in order to impart a composite structure (Fig. 19.26a). Axisymmetric parts, such as pipes and storage tanks, and even some nonsymmetric parts, are produced on a rotating mandrel. The reinforcing filament, tape, or roving is wrapped continuously around the form. The reinforcements are impregnated by passing them through a polymer bath.

The products made by filament winding are very strong, because of their highly reinforced structure, and have high strength-to-weight ratios. Parts as large as 4.5 m in diameter and 20 m long have been made by this process. The process also has been used for strengthening cylindrical or spherical pressure vessels (Fig. 19.26b), made of such materials as aluminum and titanium, where the presence of a metal inner lining makes the part impermeable. Seven-axis computer-controlled machines can automatically dispense several unidirectional prepregs to make such nonsymmetric parts as aircraft engine ducts, fuselages, propellers, blades, and struts.

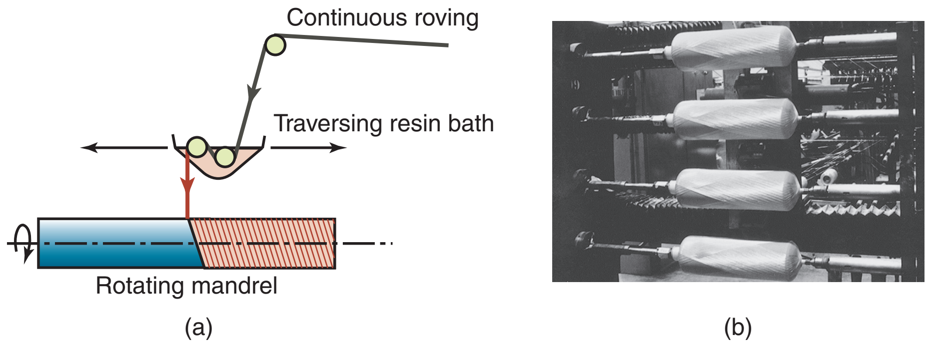


Figure 19.26: (a) Schematic illustration of the filament-winding process; (b) fiberglass being wound over aluminum liners for slide-raft inflation vessels for the Boeing 767 aircraft. *Source:* Courtesy of Brunswick Corporation.

Pultrusion. Long parts with various uniform cross sections, such as rods, profiles, flat strips, and tubing, are made continuously by the pultrusion process; the sequence of operations is shown in Fig. 19.27. The reinforcements are typically glass roving or fabric, made of E type calcium aluminosilicate glass fiber (Section 9.2.1). It is continuously supplied through several bobbins. The bundle is first pulled through a thermosetting polymer bath (usually polyester), then through a preforming die, and finally through a heated steel die.

The product is cured during its travel through the heated die, with a length of up to 1.5 m, and a speed that allows sufficient time for the polymer to set. Note that this is an operation similar to continuously baking bread or cookies, or making resin-bonded grinding wheels. The exiting material is then cut into desired lengths. Cross sections as large as 1.5 m $\times$ 0.3 m have been made by this process. Typical products made by pultrusion, which may contain up to about 75% reinforcing fiber, are golf clubs, ski poles, fishing poles, drive shafts, ladders, and handrails.

Pulforming. Continuously reinforced products, other than those with constant cross-sectional profiles, are made by *pulforming*. After being pulled through the polymer bath, the composite is clamped between the two halves of a die and cured into a finished shape. Commonly made products are hammer handles reinforced by glass fibers and curved automotive leaf springs.

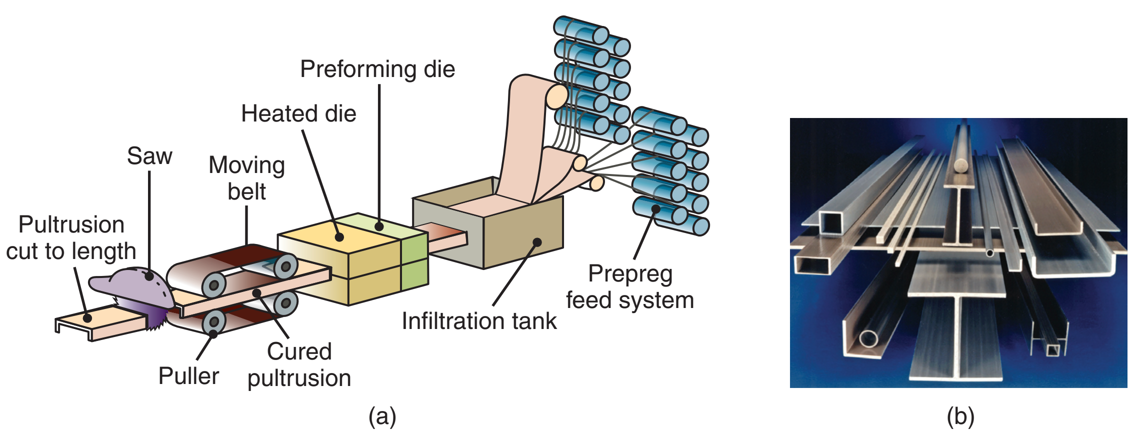


Figure 19.27: (a) Schematic illustration of the pultrusion process. (b) Examples of parts made by pultrusion. The major components of fiberglass ladders (used especially by electricians) are made by this process. They are available in different colors, but are heavier because of the presence of glass fibers. *Source:* Courtesy of Strongwell.

Case Study 19.1 Polymer Automotive-body Panels Shaped by Various Processes

Polymeric materials are commonly used for automobile bodies; this example outlines typical applications of polymers. Three commonly used and competing processing methods are: (a) injection-molding of thermoplastics and elastomers, (b) reaction-injection molding (RIM) of polyurea/polyurethanes, and (c) compression-molding of sheet-molding compound (SMC) with resin-transfer-molded polyester and vinylester.

Typical examples of parts made are:

- Body panels and other large exterior components made by injection molding
- Front fenders and rear quarter panels made of polyphenylene-ether/nylon or thermoplastic polyester
- Outer door panels made of polycarbonate/ABS
- Fascias made of thermoplastic polyolefin.

These materials are selected for design flexibility, impact strength and toughness, corrosion resistance, high durability, and low mass. Vertical panels and fascias are made in multicavity molds on large injection-molding machines; they are then assembled mechanically to a steel frame.

Large exterior-body parts are made of reaction-injection molded polyurethane, although polyureas are important for body panels and bumpers. Thermoset fascias are made of reinforced RIM polyurethane and polyureas, because of their higher thermal stability, low-temperature toughness, and lower cycle times. Large horizontal exterior-body panels, such as hoods, roofs, and rear decks, are made of reinforced polyester or vinylester in the form of compression-molded sheet-molding compounds. Lower volume parts are made by resin-transfer molding.

Environmental and recycling considerations in material and process selection for automobiles, as well as other products, continue to be important. For example, polyphenylene oxide is being replaced with polycarbonate, which is made out of 100% recycled or reclaimed materials.

19.13.4 Quality Considerations in Processing Reinforced Plastics

The major quality considerations in the processes described thus far concern internal voids and gaps between successive layers of material. Volatile gases that develop during processing must be allowed to escape from the layup through the vacuum bag, in order to avoid porosity due to trapped gases. Microcracks may develop during improper curing or during the transportation and handling of parts. These defects can be detected using ultrasonic scanning and other techniques described in Section 36.10.

Case Study 19.2 Manufacturing of Head Protector® Tennis Racquets

Competitive tennis is a demanding sport; there is a continuing demand to produce exceptionally *lightweight and stiff* racquets to improve performance. A tennis racquet consists of a number of regions (Fig. 19.28), of which the sweet spot is of particular interest. When the tennis ball is struck at the sweet spot, the player has optimum control and power, and vibration is minimized. Several innovative racquet-head designs have been developed over the years to maximize the size of the sweet spot. A stiff composite material, with high-modulus graphite fibers in an epoxy matrix (Chapter 9), is used to make the racquet head. Orientation of the fibers varies in different locations of the racquet; the main tube for the racquet, for example, consists of carbon-epoxy prepreg, oriented at $\pm 30^\circ$ from layer to layer.

The advantages in using such materials are obvious, in that stiff racquets allow higher forces to be applied to the ball. However, the use of these advanced materials also has led to an increased frequency of *tennis elbow*, a painful condition associated with tendons that anchor muscles to the bones at the elbow. The condition is due not only to the higher forces involved, but also to the associated greater vibration of the racquet encountered with every stroke, especially when balls are struck away from the sweet spot.

An innovative design for a racquet, the Protector® (made by Head Sport AG) uses lead zirconate titanate (PZT) fibers, as an integral layer of the composite racquet frames. PZT is well known as a piezoelectric material (Section 3.7); that is, it produces an electric response when deformed. Modules of the fibers, called Intellifibers®, are integrated into the throat on all sides of the racket. The module consists of about 50 PZT fibers, each approximately 0.3 mm in diameter, sandwiched between two polyamide layers, with printed electrodes for generating the potential difference when the fibers are bent.

During impact, the vibrations constantly excite the Intellifibers®, generating a very high voltage potential but at low current. The energy is stored, in real time, in coils on the printed circuit board (Chipsystem®) incorporated in the racquet handle, and released back to the Intellifibers®, in the optimal phase and waveform for the most efficient damping. The stored energy is sent back to the Intellifibers® in a phase that causes a mechanical force opposite to the vibration, thereby reducing it. The Chipsystem® is tuned to the first natural frequency of the racket, and it damps vibrations only within a range of its design frequency.

Making a Protector® tennis racquet involves a number of steps:

1. A carbon-epoxy prepreg is first produced, as described in Section 19.13.1.
2. The prepreg is cut to the proper size and placed on a flat, heated bench to make the matrix material tackier, resulting in better adhesion to adjacent layers.
3. A polyamide sleeve (bladder) is then placed over a rod, and the prepreg is rolled over the sleeve.
4. When the bar is removed, the result is a tube of carbon-epoxy prepreg with a polyamide sleeve, that can be placed in a mold and internally pressurized to develop the desired cross section.

The throat piece is molded separately by wrapping the prepreg around sand-filled polyamide preforms or expandable foam. Since there is no easy way to provide air pressure to the throat, the preform develops its own internal pressurization because of the expansion of air during exposure to elevated molding temperatures. If sand is used, it is removed through the holes drilled into the preform during the finishing operation.

Prior to molding them, all components are assembled onto a template, and the final prepreg pieces are added to strategic areas. The main tube is bent around the template, and the ends are pressed together and wrapped with a prepreg layer, forming the handle. The PZT fibers are incorporated as the outer layer in the racquet in the throat area, and the printed electrodes are connected to the Chipsystem®. The racquet is then placed into the mold, internally pressurized, and allowed to cure. Note that this operation is essentially an internally pressurized, pressure-bag molding process (see Fig. 19.24b). Figure 19.29a shows a racquet as it appears directly after molding.

The racquet then undergoes a number of finishing operations, including flash removal, drilling of holes to accommodate strings, and finishing of the handle, including wrapping it with a special grip material. A completed Head Protector® racquet is shown in Fig. 19.29b. This design has been found to reduce racquet vibrations by up to 50%, resulting in clinically proven reductions in tennis elbow, without any compromise in performance.

Source: Courtesy of J. Kotze and R. Schwenger, Head Sport AG.

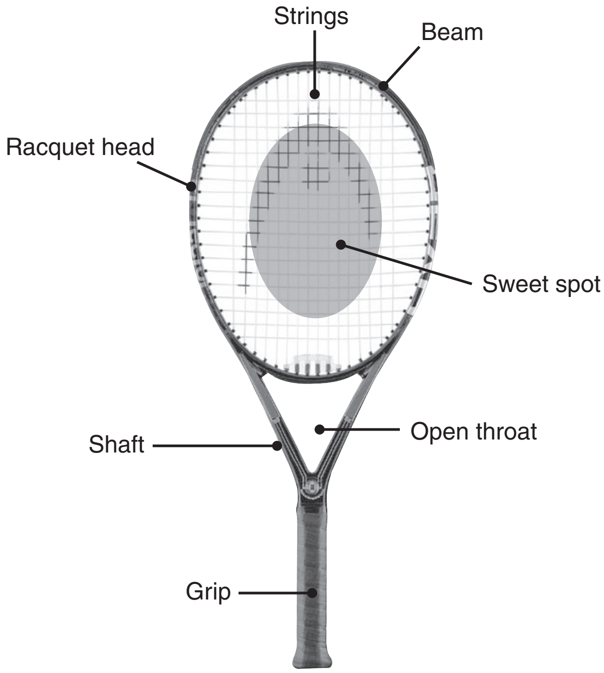


Figure 19.28: A Head Protector® tennis racquet. *Source:* Courtesy of Head Sport AG.

19.14 Processing Metal-matrix and Ceramic-matrix Composites

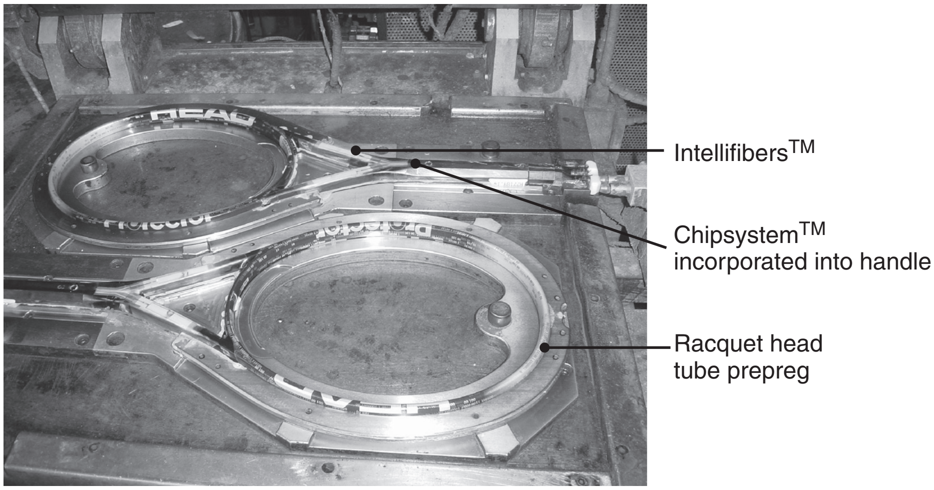
Metal-matrix composites can be made into near-net shaped parts by the following processes:

- **Liquid-phase processing** basically consists of casting together the liquid-matrix material, such as aluminum or titanium, and the solid reinforcement, such as graphite, aluminum oxide, or silicon carbide by conventional casting processes or by pressure-infiltration casting. In the latter process, pressurized gas forces the liquid-metal matrix into a preform, usually shaped out of wire or sheet and made of reinforcing fibers.

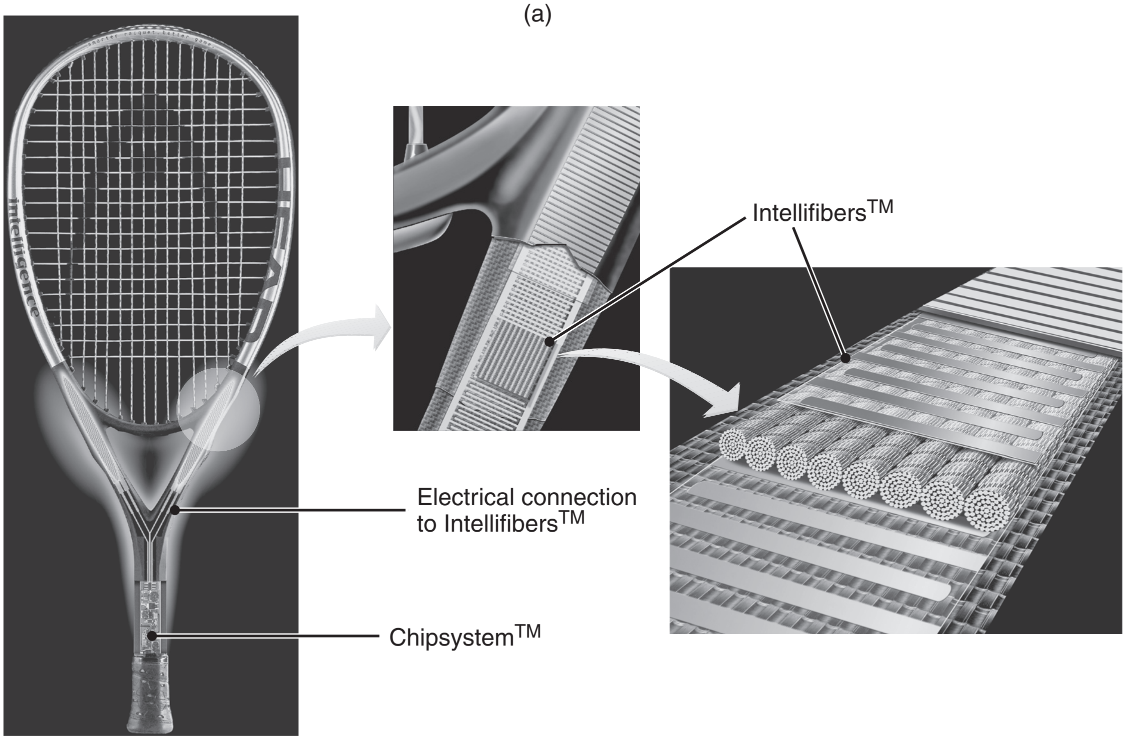
- **Solid-phase processing** utilizes powder-metallurgy techniques (Chapter 17), including cold and hot isostatic pressing. Proper mixing is important for homogeneous distribution of the fibers throughout the part. An example is the production of tungsten-carbide tools and dies, with cobalt as the matrix material.

- **Two-phase (liquid–solid) processing** involves technologies that consist of rheocasting (Section 11.4.7) and the techniques of *spray atomization* and *deposition*. In the latter two processes, the reinforcing fibers are mixed with a matrix that contains both liquid and solid phases of the metal.

In making complex metal-matrix composite parts with whisker or fiber reinforcement, die geometry and control of process variables are very important for ensuring the proper distribution and orientation of the fibers within the part. MMC parts made by PM techniques generally are heat treated for optimum properties.



(a)



(b)

Figure 19.29: (a) The composite Head Protector® racquet immediately after molding; (b) a completed Head Protector® racquet, highlighting the incorporation of piezoelectric Intellifibers®. *Source:* Courtesy of Head Sport AG.

Case Study 19.3 Metal-matrix Composite Brake Rotors and Cylinder Liners

Some brake rotors are made of composites consisting of an aluminum-based matrix, reinforced with 20% silicon-carbide particles. First, the particles are stirred into molten aluminum alloys, and the mixture is cast into ingots. The ingots are then remelted and cast into shapes, by such processes as green-sand, bonded-sand, investment, permanent-mold, and squeeze casting. The rotors (a) are about one-half the weight of those made of gray cast iron, (b) conduct heat three times faster, (c) add stiffness and wear-resistance characteristics of ceramics, and (d) reduce noise and vibration, because of internal damping in the rotors.

To improve the wear- and heat resistance of cast-iron cylinder liners in aluminum engine blocks, aluminum-matrix liners are also available. The metal-matrix layer consists of 12% aluminum-oxide fiber and 9% graphite fiber, and has a thickness that ranges from 1.5 to 2.5 mm.

19.14.1 Processing Ceramic-matrix Composites

Several processes, including such techniques as melt infiltration, controlled oxidation, and hot-press sintering, are used to make ceramic-matrix composites.

- **Slurry infiltration** is the most common process for making ceramic-matrix composites. It involves the preparation of a fiber preform, which is first hot pressed and then impregnated with a combination of slurry (containing the matrix powder), a carrier liquid, and an organic binder. This process imparts high strength, toughness, and uniform structure, but the product has limited high-temperature properties. A further improvement on the process is **reaction bonding** or **reaction sintering** of the slurry.

- **Chemical-synthesis** processes involve the sol-gel and the polymer-precursor techniques. In the **sol-gel process**, a *sol* (a colloidal fluid having the liquid as its continuous phase) that contains fibers is converted to a *gel*, which is then subjected to heat treatment to produce a ceramic-matrix composite. The **polymer-precursor method** is analogous to the process used in making ceramic fibers with aluminum oxide, silicon nitride, and silicon carbide.

- In **chemical-vapor infiltration**, a porous fiber preform is infiltrated with the matrix phase, using the chemical vapor deposition technique (Section 34.6). The product has very good high-temperature properties, but the process is time consuming and costly.

19.15 Design Considerations

Design considerations in forming and shaping plastics are similar to those for casting metals (Section 12.2). The selection of appropriate materials from an extensive list requires considerations of (a) service requirements, (b) possible long-range effects on properties and behavior, such as dimensional stability and wear, and (c) ultimate disposal of the product following its life cycle. Some of these issues are described in Sections I.4 and I.6 in the General Introduction, and Section 7.8.

Outlined below are the general design guidelines for the production of plastic and composite-material parts:

1. The processes for plastics have inherent flexibility, thus a wide variety of part shapes and sizes can be produced. Complex parts, with internal and external features, can be produced with relative ease and at high production rates. Consequently, a process such as injection molding competes well with powder-injection molding and die casting. All are capable of producing complex shapes and having thin walls.

2. In process substitutions, it is essential to consider the materials involved and that their characteristics may be very different, each having its own properties suitable to a particular method of production.

3. Compared with metals, plastics have much lower stiffness and strength; thus, section size, shape, and thickness must be selected accordingly. Depending on the application, a high section modulus can be achieved on the basis of design principles common to I-beams and tubes.

4. Large, flat surfaces can be stiffened by such simple means as specifying curvatures on parts. For example, observe the stiffness of very thin but gently curved slats in venetian blinds. Reinforcement with fibers or particles also are effective in achieving stiff and lightweight designs.

5. The overall part shape and thickness often determine the particular shaping or molding process to be selected. Even after a particular process is chosen, the design of parts and the dies should be appropriate for a particular shape generation (Fig. 19.30), dimensional control, and surface finish.

6. Because of low stiffness and thermal effects, dimensional tolerances, especially for thermoplastics, are not as small as in metalworking processes. For example, dimensional tolerances are much smaller in injection molding than they are in thermoforming.

7. As in casting metals and alloys, the control of material flow in the mold cavities is essential. The effects of molecular orientation during the processing of the polymer also must be considered, especially in extrusion, thermoforming, and blow molding.

8. Large variations in cross-sectional areas and section thicknesses, as well as abrupt changes in geometry, should be avoided. Note, for example, that the *sink marks* (pull-in) shown in the top piece in Fig. 19.30c are due to the fact that thick sections in a part solidify last.

9. Contractions in larger cross sections during cooling tend to cause *porosity* in plastic parts, as they do in metal casting (see Fig. 12.2), thus affecting product integrity and quality. By contrast, a lack of stiffness may make it more difficult to remove thin parts from molds after shaping them.

10. Low elastic moduli of plastics further requires that shapes be selected properly for improved stiffness of the component (Fig. 19.30b), particularly when saving material is an important factor. Note that these considerations are similar to those applicable to the design of metal castings and forgings, as is the need for drafts (typically less than 1 degree for polymers) to enable removal of the part from

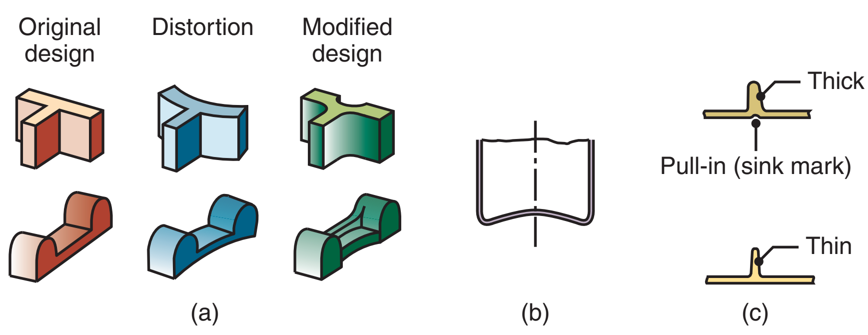


Figure 19.30: Examples of design modifications to eliminate or minimize distortion in plastic parts: (a) suggested design changes to minimize distortion, (b) stiffening the bottoms of thin plastic containers by doming (a technique similar to the process used to shape the bottoms of aluminum beverage cans, see Fig. 16.31), and (c) design change in a rib to minimize pull-in (*sink mark*), caused by shrinkage during the cooling of thick sections in molded parts.

molds and dies. Generally, the recommended part thickness ranges from about 1 mm for small parts to about 3 mm for large parts.

11. Physical properties, especially high coefficient of thermal expansion, are important factors. Improper part design can lead to uneven shrinking (Fig. 19.30a) and distortion (warping). Plastics can easily be molded around metallic parts and inserts; however, their interfacial strength and compatibility with metals, when so assembled, is an important consideration.

12. The properties of the final product depend on the original material and its processing history. For example, the cold working of polymers improves their strength and toughness. On the other hand, because of the nonuniformity of deformation, even in simple rolling, residual stresses develop in polymers, just as they do in metals. These stresses also can be due to thermal cycling of the part during processing.

13. However they are produced, the magnitude and direction of residual stresses are important factors, such as in stress cracking over time. Furthermore, these stresses can relax over a period of time and cause distortion of the part during its service life.

14. A major design advantage of reinforced plastics is the directional nature of the strength and stiffness of the composite (see, for example, Fig. 9.7). External forces applied to the part are transferred by the matrix to the fibers, which are much stronger and stiffer than the matrix. When all of the fibers are oriented in one direction, the resulting composite material is exceptionally strong in the fiber direction, but weak in the transverse direction. To achieve strength in two principal directions, individual unidirectional layers are laid at the controlled angles to each other, as is done in tape layup or filament winding. If strength in the third (thickness) direction is required, a different type of reinforcement, such as woven fiber, is used to form a sandwich structure.

19.16 Economics of Processing Plastics and Composite Materials

General characteristics of processing of plastics and composite materials are given in Table 19.2. Note the wide range of equipment and tooling costs, and the economic production quantities. As described throughout this chapter, there is some relationship between equipment costs and tool and die costs.

Table 19.2: Comparative Production Characteristics of Various Molding Methods.

Molding method	Equipment and tooling cost	Production rate	Economical production quantity
Extrusion	M–L	VH–H	VH
Injection molding	VH	VH	VH
Rotational molding	M	M–L	M
Blow molding	M	H–M	H
Compression molding	H–M	M	H–M
Transfer molding	H	M	VH
Thermoforming	M–L	M–L	H–M
Casting	M–L	M–L	L
Centrifugal casting	H–M	M–L	M–L
Pultrusion	H–M	H	H
Filament winding	H–M	L	L
Spray layup and hand layup	L–VL	L–VL	L

VH = very high; H = high; M = medium; L = low; VL = very low.

The most expensive machines are for injection molding, followed by compression molding and transfer molding; tool and die costs also are high for these operations. Thus, in an operation like injection molding, the size of the die and the optimum number of cavities in the die for producing more and more parts in one cycle are important considerations, as they are in an operation like die casting. Larger dies may be considered in order to accommodate several cavities, with runners to each cavity, but at the expense of increasing die cost even further. On the other hand, more parts will be produced per machine cycle, thus the production rate will increase.

A detailed analysis is thus required in order to determine the overall die size, the number of cavities in the die, and the machine capacity required to optimize the total operation, and to produce parts at minimum cost. Similar considerations also apply to all other processing methods described throughout this chapter.

Summary

- Thermoplastics can be shaped by a wide variety of processes, including extrusion, molding, casting, and thermoforming, as well as by some of the processes used in processing metals. The raw material usually is in the form of pellets, granules, and powders.
- The high strain-rate sensitivity of thermoplastics allows extensive stretching in forming operations; thus, complex and deep parts can be produced easily. Thermosetting plastics generally are molded or cast, and they have better dimensional accuracy than forming thermoplastics.
- Fiber-reinforced plastics are processed into structural components using liquid plastics, prepregs, and bulk- and sheet-molding compounds. Fabricating techniques include various molding methods, filament winding, pultrusion, and pulforming. The type and orientation of the fibers and the strength of the bond between fibers and matrix and between layers of materials are important considerations.
- The design of plastic parts must take into account their low strength and stiffness, and such physical properties as high thermal expansion and generally low resistance to temperature. Inspection techniques have been developed to determine the integrity of plastic products.
- Processing of metal-matrix and ceramic-matrix composites continues to undergo developments to ensure product integrity, reliability, and reduced costs. Metal-matrix composites are processed by liquid-phase, solid-phase, and two-phase processes. Ceramic-matrix composites can be processed by slurry infiltration, chemical synthesis, or chemical-vapor infiltration.
- Relevant factors in the economics of the operations described include the costs of the machinery, level of controls, tooling, cycle times, and production rate and volume.

Key Terms

Blow molding
Blow ratio
Bulk-molding compound
Calendering
Casting
Chemical synthesis
Chemical-vapor infiltration
Coat-hanger die
Coextrusion
Cold forming
Compression molding
Contact molding
Encapsulation
Extrusion
Extrusion blow molding
Filament winding
Foam molding
Hand layup

Ice-cold forming
Injection molding
Insert molding
Liquid-phase processing
Masticated
Melt spinning
Open-die processing
Overmolding
Parison
Pellets
Plastisols
Potting
Prepregs
Pulforming
Pultrusion
Reaction-injection molding
Resin transfer molding
Rotational molding
Sheet-molding compound
Sink marks
Sizing
Slurry infiltration
Slush molding
Solid-phase forming
Solid-phase processing
Spinneret
Spinning
Spray layup
Structural foam molding
Swell
Thermoforming
Thick-molding compound
Transfer molding
Two-phase processing
Vacuum-bag molding

Bibliography

Advani, S.G., and Hsiao, K.-T. (eds.), **Manufacturing Techniques for Polymer Matrix Materials**, Woodhead Publishing, 2012.
Baird, D.G., and Collias, D.I., **Polymer Processing: Principles and Design**, 2nd ed., Wiley, 2014.
Boisse, P. (ed.), **Advances in Composites in Manufacturing and Process Design**, Woodhead Publishing, 2015.
Brinson, G., **Handbook of Advanced Composites and Polymer Manufacturing**, Auris Reference Ltd., 2013.
Campbell, F., **Manufacturing Processes for Advanced Composites**, Elsevier, 2004.
Chanda, J.M., and Roy, S.K., **Plastics Technology Handbook**, 5th ed., CRC Press, 2017.
Giles Jr., H.F., and Wagner, J.R., Jr., **Extrusion**, 2nd ed., Woodhead Publishing, 2013.
Hopmann, C., **Extrusion Dies for Plastics and Rubbers**, 4th ed., Hanser, 2016.
Johnson, P.S., **Rubber Processing: An Introduction**, Hanser, 2001.
Kutz, M. (ed.), **Applied Plastics Engineering Handbook: Processing and Materials**, 2nd ed., William Andrew, 2016.
Mallick, P.K., **Processing of Polymer Matrix Composites**, CRC Press, 2017.
Mastro, P.F., **Plastics Products Design**, Wiley, 2016.
Menning, G., **Mold-Making Handbook**, 3rd ed., Hanser, 2013.
Osswald, T.A., **Injection Molding Handbook**, 2nd ed., Hanser, 2007.
Rauwendaal, C., **Polymer Extrusion**, 5th ed., Hanser, 2014.
Rosato, D., **Plastics Technology Handbook**, Momentum Press, 2010.
Rotheiser, J., **Joining of Plastics**, 3rd ed., Hanser, 2009.
Strong, A.B., **Plastics: Materials and Processing**, 3rd ed., Prentice Hall, 2005.

Strong, A.B., **Fundamentals of Composites Manufacturing**, 2nd ed., Society of Manufacturing Engineers, 2007.
Tadmore, Z., and Gogos, C.G., **Principles of Polymer Processing**, 2nd ed., Wiley, 2006.
Troughton, M.J., **Handbook of Plastics Joining**, 2nd ed., William Andrew, 2008.
Wagner, J.R. (ed.), **Handbook of Troubleshooting Plastics Processes**, Wiley-Scriver, 2012.

Review Questions

19.1. What are the forms of raw materials for processing plastics into products?

19.2. What is extrusion? What products are produced by polymer extrusion?

19.3. Describe the features of an extruder screw and their functions.

19.4. How are injection-molding machines rated?

19.5. What is (a) a parison, (b) a plastisol, and (c) a prepreg?

19.6. How is thin plastic film produced?

19.7. List several common products that can be made by thermoforming.

19.8. What similarities and differences are there between compression molding and closed-die forging?

19.9. Explain the difference between potting and encapsulation.

19.10. What is thermoforming?

19.11. Describe runner, gate, sprue and well.

19.12. Describe the advantages of cold-forming plastics over other plastic-processing methods.

19.13. What are the characteristics of filament-wound products? Explain why they are desirable.

19.14. Describe the methods that can be used to make tubular plastic products.

19.15. What is pultrusion? Pulforming?

19.16. How are plastic sheet and plastic film produced?

19.17. What process is used to make foam drinking cups?

19.18. If a polymer is in the form of a thin sheet, is it a thermoplastic or thermoset? Why?

19.19. How are polymer fibers made? Why are they much stronger than bulk forms of the polymer?

19.20. What are the advantages of coextrusion?

19.21. Explain how latex rubber gloves are made.

Qualitative Problems

19.22. Describe the features of a screw extruder and its functions.

19.23. Explain why injection molding is capable of producing parts with complex shapes and fine detail.

19.24. Describe the advantages of applying the traditional metal-forming techniques, described in Chapters 13 through 16, to making (a) thermoplastic and (b) thermoset products.

19.25. Explain the reasons that some plastic-forming processes are more suitable for certain polymers than for others. Give examples.

19.26. Describe the problems involved in recycling products made from reinforced plastics.

19.27. Can thermosetting plastics be used in injection molding? Explain.

19.28. Inspect some plastic containers, such as those containing talcum powder, and note that the integral lettering on them is raised rather than depressed. Explain.

19.29. An injection-molded nylon gear is found to contain small pores. It is recommended that the material be dried before molding it. Explain why drying will solve this problem.

19.30. Explain why operations such as blow molding and film-bag making are performed vertically.

19.31. Comment on the principle of operation of the tape-laying machine.

19.32. Typical production rates are given in Table 19.2. Comment on your observations and explain why there is such a wide range.

19.33. What determines the cycle time for (a) injection molding, (b) thermoforming, and (c) compression molding? Explain.

19.34. Does the pull-in defect (sink marks) shown in Fig. 19.30c also occur in metal-forming and casting processes? Explain.

19.35. What determines the intervals at which the indexing head in Fig. 19.14c rotates from station to station?

19.36. Identify processes that would be suitable for small production runs on plastic parts, of, say, 100.

19.37. Identify processes that are capable of producing parts with the following fiber orientations in each: (a) uniaxial, (b) cross-ply, (c) in-plane random, and (d) three-dimensional random.

19.38. Inspect several electrical components, such as light switches, outlets, and circuit breakers, and describe the process or processes used in making them.

19.39. Inspect several similar products that are made of metals and plastics, such as a metal bucket and a plastic bucket of similar shape and size. Comment on their respective thicknesses, and explain the reasons for their differences, if any.

19.40. What are the advantages of using whiskers as a reinforcing material?

19.41. Construct a table that lists the main manufacturing processes described in this chapter. Indicate those that can be used for (a) thermoplastics; (b) thermosets; (c) composite materials.

Quantitative Problems

19.42. Estimate the die-clamping force required for injection molding five identical 200 mm diameter disks in one die. Include the runners of appropriate length and diameter.

19.43. A 2-L plastic beverage bottle is made by blow molding a parison 125 mm long and with a diameter that is the same as that of the threaded neck of the bottle. Assuming uniform deformation during molding, estimate the wall thickness of the tubular portion of the parison.

19.44. Consider a Styrofoam® drinking cup. Measure the volume of the cup and its weight. From this information, estimate the percent increase in volume that the polystyrene beads have undergone.

19.45. In extrusion, what flight angle should be used on a screw so that a flight translates a distance equal to the barrel diameter with every revolution?

19.46. Consider the part in Problem 17.41. If this part is to be produced in injection molding, with four parts produced with each shot (see Fig. 19.10b), estimate the clamping force required. Sketch the layout of the part, sprue and runners.

19.47. Assume that you are asked to give a quiz to students on the contents of this chapter. Prepare five quantitative problems and five qualitative questions, and supply the answers.

Synthesis, Design, and Projects

19.48. Make a survey of a variety of sports equipment, such as bicycles, tennis racquets, golf clubs, and baseball bats, and identify the components made of composite materials. Explain the reasons for and advantages of using composites for these specific applications.

19.49. Explain the design considerations involved in replacing a metal beverage can with one made completely of plastic.

19.50. Give examples of several parts suitable for insert molding. How would you manufacture these parts if insert molding were not available?

19.51. Give other examples of design modifications in addition to those shown in Fig. 19.30.

19.52. With specific examples, discuss the design issues involved in making products out of plastics vs. reinforced plastics.

19.53. Die swell in extrusion is radially uniform for circular cross-sections, but is not uniform for other cross-sections. Recognizing this fact, make a qualitative sketch of a die profile that will produce (a) square, (b) triangular, (c) elliptical, and (d) gear-shaped cross-sections of extruded polymer.

19.54. Inspect various plastic components in a typical automobile, and identify the processes that could have been used in making them.

19.55. Inspect several similar products that are made either from metals or from plastics, such as a metal bucket and a plastic bucket of similar shape and size. Comment on their respective shapes and thicknesses and explain the reasons for their differences.

19.56. Write a brief paper on how plastic coatings are applied to (a) electrical wiring, (b) sheet-metal panels, (c) wire baskets, racks, and similar structures, and (d) handles for electricians tools, such as wire cutters and pliers requiring electrical insulation.

19.57. It is well-known that plastic forks, spoons, and knives are not particularly rigid. What suggestions would you have to make them better? Describe processes that could be used for producing them.

19.58. Some plastic products have lids with integral hinges; that is, no other material or part is used at the junction of the two parts. Identify such products, and describe a method for making them.

19.59. Make a survey of the technical literature, and describe how different types of (a) pneumatic tires, (b) automotive hoses, and (c) garden hoses are manufactured.

19.60. Obtain a boxed kit for assembling a model car or airplane. Examine the injection-molded parts provided, and describe your thoughts on the layout of the molds to produce these parts.

19.61. In injection-molding operations, it is common practice to remove the part from its runner, place the runner in a shredder, and recycle the runner by producing pellets. List the concerns you may have in using such recycled pellets for products, as against “virgin” pellets.

19.62. An increasing environmental concern is the very long period required for the degradation of polymers in landfills. Noting the information given in Section 7.8 on biodegradable plastics, conduct a literature search on the trends and developments in the production of these plastics.

19.63. Examine some common and colorful plastic poker chips and give an opinion on how they were manufactured.

19.64. Obtain different styles of toothpaste tubes, carefully cut them across, and comment on your observations regarding (a) the type of materials used and (b) how the tubes were produced.

19.65. By incorporating small amounts of blowing agent, it is possible to manufacture polymer fibers with gas cores. List some applications for such fibers.

Chapter 20

Additive Manufacturing

Case Studies:

- This chapter describes the technologies associated with additive manufacturing (AM), sharing the characteristics of computer integration, production without the use of traditional tools and dies, and the ability to rapidly produce a single part or small batches of parts on demand. All have the basic characteristics of producing individual parts layer by layer.
- Classes of processes used in additive manufacturing are reviewed, which include extrusion-based methods, photo polymerization, powder bed processes, sprayed powder approaches, and lamination-based methods.

- The practice of applying additive manufacturing techniques to the production of tooling that can be used in other manufacturing processes is described.
- The chapter closes with a summary of additive manufacturing design, opportunities, and economics.

Typical parts made: A wide variety of metallic and nonmetallic parts for product design analysis, evaluation and finished products.

Alternative processes: Machining, casting, molding, powder metallurgy, forging, and fabricating.

20.1 Introduction

Making a *prototype*, the first full-scale model of a product, has traditionally involved flexible manufacturing processes, and often required several weeks or months. Prototypes can now be quickly produced through **subtractive processes** (basically involving computer-controlled machining operations, described in Chapters 21–25) or by **virtual prototyping** (involving advanced graphics and software). An important advance is **additive manufacturing**, by which a solid physical model of a part is made directly from a three-dimensional CAD drawing without the use of tools, and allowing for extremely complex geometries (Fig. 20.1). Additive manufacturing is a suite of processes using different approaches, including photo polymerization, robot controlled extrusion, selective sintering, etc.

This chapter describes **additive manufacturing**, formerly called rapid prototyping, whereby parts are *built in layers*. Developments in additive manufacturing began in the mid-1980s. The advantages of this technology include:

- Physical models of parts, produced from CAD data files, can be manufactured in a matter of minutes to hours, and thus allow the rapid evaluation of manufacturability and design effectiveness. In this way, additive manufacturing serves as an important tool in the product development process.
- A wide variety of materials are available, ranging from compliant rubber-like polymers to stiff polymers, metals, and ceramics.

(a) (b) (c)

Figure 20.1: Examples of parts made by additive manufacturing processes: (a) a selection of parts from fused-deposition modeling; (b) full-color model of an anatomical model; and (c) a speaker cover produced by the CLIP process. *Source:* (a) and (b) Courtesy of Stratasys, Inc., (c) Courtesy of Carbon, Inc.

- Additive manufacturing operations can be used in some applications to produce actual tooling for manufacturing operations (**rapid tooling**, see Section 20.10). Thus, one can make tooling in a matter of a few days.

Case Study 20.1 Functional Prototyping

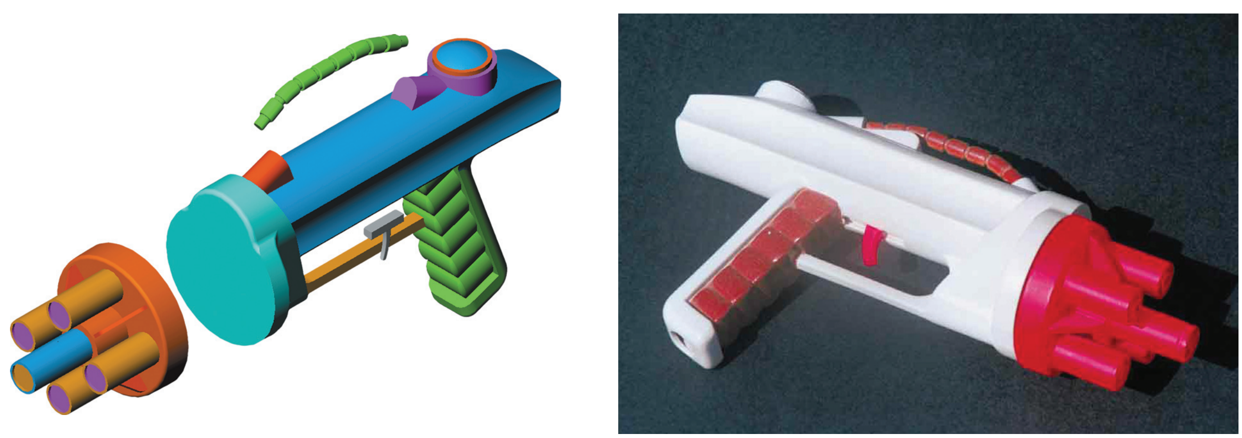
Toys are examples of mass-produced items with universal appeal. Because some toys are actually complex, the function and benefits of a computer-aided design (CAD) cannot be ensured until prototypes have been made. Fig. 20.2 shows a CAD model and a rapid-prototyped version of a water squirt gun (Super Soaker Power Pack Back Pack® water gun), which was produced on a fused-deposition modeling machine. Each component was produced separately and assembled into the squirt gun; the prototype could actually hold and squirt water. The alternative would be to produce components on CNC milling machines or fabricate them in some fashion, but this can be done only at a much higher cost.

By producing a prototype, interference issues and assembly problems can be assessed and, if necessary, corrected. Moreover, from an aesthetic standpoint, the elaborate decorations on such a toy can be more effectively evaluated from a prototype than from a CAD file. Also, they can be adjusted to improve the toy's appeal. Each component, having its design verified, then has its associated tooling produced, with better certainty that the tooling, as ordered, will produce the parts desired.

Additive manufacturing has now been transformed from a prototyping technology to a viable strategy for product production. In addition to traditional approaches to manufacturing, the use of additive manufacturing introduces opportunities, including the following:

1. A part produced from additive manufacturing can itself be used in subsequent manufacturing operations to produce the final parts. Also called **direct prototyping**, this approach can serve as an important manufacturing technology.

2. **Mass customization** can be achieved, where every part can be tailored to a particular user or an application. For example, it is possible to create prosthetic devices that are tailored to individual patients, based on scanned measurements of the person to produce an optimum fit and function. These prosthetics have the ability to allow bathing and are more comfortable, and have fewer complications than cast supports.



(a) (b)

Figure 20.2: Additive manufacturing of a Super Soaker® squirt gun. (a) original CAD description of a toy; (b) fully functional toy produced through fused-deposition modeling. *Source:* (b) Courtesy of Rapid Models and Prototypes, Inc., and Stratasys, Inc.

Table 20.1: Characteristics of Additive Manufacturing Technologies.

Process	Supply phase	Layer creation technique	Type of phase change	Materials
Stereolithography	Liquid	Liquid layer curing	Photopolymerization	Photopolymers (acrylates, epoxies, colorable resins, and filled resins)
CLIP	Liquid	Liquid layer curing	Photopolymerization	Similar to stereolithography
Multijet/PolyJet	Liquid	Liquid layer curing	Photopolymerization	similar to stereolithography
Material jetting	Liquid	Droplet deposition	Solidification	Polymers and wax
Fused-deposition modeling	Solid	Extrusion of melted polymer	Solidification	Thermoplastics such as ABS, polycarbonate, and polysulfone
Binder jetting	Powder	Binder-droplet deposition onto powder layer	No phase change	Ceramic, polymer, or metal powder; sand
Selective laser sintering	Powder	Layer of powder	Sintering	Polymer powder such as nylon
Selective Laser melting	Powder	Layer of powder	Solidification	Metal powders such as stainless steel, titanium, copper, and aluminum
Electron-beam melting	Powder	Layer of powder	Solidification	Titanium and titanium alloys, cobalt chrome
Laminated-object manufacturing	Solid	Deposition of sheet material	No phase change	Paper and polymers
Laser-engineered net shaping	Powder	Injection of powder stream	Solidification	Titanium, stainless steel, aluminum

3. Widespread application of additive manufacturing allows **distributed manufacturing**, so that parts can be produced anywhere and not only in factories. For example, prosthetics and braces can be produced at a hospital, avoiding time associated with orders and shipping. Thus, a child with a broken arm can be brought to a hospital and fixed with an optimum brace essentially as quickly as a typical conventional treatment.

Almost all materials can be used through one or more additive manufacturing approaches, as outlined in Table 20.1. However, because their properties are more suitable for these unique operations, polymers are the most commonly used material today, followed by metals and ceramics (see Table 20.2); still, new processes are being introduced continually. The rest of this chapter serves as a general introduction to the most common additive manufacturing operations, describes their advantages and limitations, and explores the present and future applications of these processes.

20.2 Additive Manufacturing Methodology

Additive manufacturing operations all build parts in *layers*, as summarized in Table 20.1. These processes use various physics to achieve a desired part, and a wide variety of materials can be used. All of the processes described in this section build parts *layer by layer*. In order to visualize the methodology employed, it is beneficial to think of the construction of a loaf of bread by stacking and bonding individual slices of bread on top of each other (hence the term *additive*). The main difference between the various additive processes lies in the method of producing the individual slices, which are typically 0.03 to 0.5 mm thick, although they can be thicker or thinner in some systems.

All additive operations require dedicated software. Note as an example, the solid part shown in Fig. 20.3a. The first step is to develop a CAD file description of the part; the computer then constructs slices of the three-dimensional part (Fig. 20.3b). Each slice is analyzed separately, and a set of instructions is compiled in order to provide the AM machine with detailed information regarding the manufacture of the part. A trajectory often has to be planned in order to produce the slice. For example, Fig. 20.3d shows the

Table 20.2: Mechanical Properties of Selected Materials for Additive Manufacturing.

Process	Material	Tensile strength (MPa)	Elastic modulus (GPa)	Elongation in 50 mm (%)	Characteristics
Stereolithography	Accura 60	68	3.10	5	Transparent; good general-purpose material for additive manufacturing
	Somos 9920	32	1.35–1.81	15–26	Transparent amber; good chemical resistance; good fatigue properties; used for producing patterns in rubber molding
	WaterClear Ultra	56	2.9	6–9	Optically clear resin with ABS-like properties
	WaterShed 11122	47.1–53.6	2.65–2.88	11–20	Optically clear with a slight green tinge; mechanical properties similar to those of ABS; used for rapid tooling
	DMX-SL 100	32	2.2–2.6	12–28	Opaque beige; good general-purpose material for additive manufacturing
PolyJet	FC720	60.3	2.87	20	Transparent amber; good impact strength, good paint adsorption and machinability
	FC830	49.8	2.49	20	White, blue, or black; good humidity resistance; suitable for general-purpose applications
	FC 930	1.4	0.185	218	Semiopaque, gray, or black; highly flexible material used for prototyping of soft polymers or rubber
Fused-deposition modeling	Polycarbonate	52	2.0	3	White; high-strength polymer suitable for additive manufacturing and general use
	Ultem 9085	71.64	2.2	5.9	Opaque tan, high-strength FDM material, good flame, smoke and toxicity rating.
	ABS-M30i	36	2.4	4	Available in multiple colors, most commonly white; a strong and durable material suitable for general use; biocompatible
	PC	68	2.28	4.8	White; good combination of mechanical properties and heat resistance
CLIP	Rigid polyurethane	45	1.9	100	Wide variety of colors
	Flexible polyurethane	29	0.86	280	Similar to rubber band
	Epoxy	88	3.14	5.2	
	Urethane methacrylate	46	2	17	
Selective laser sintering	Aluminum AlSi12 alloy	480	—	5.5	Common aluminum alloy for AM
	17-4 stainless steel	1300	—	16	Properties are after heat treatment
	316L stainless steel	600	190	40	After stress relief
	Titanium GR.5	1100	120	30	Properties are after stress relief
	WindForm XT	77.85	7.32	2.6	Opaque black polymide and carbon; produces durable heat- and chemical-resistant parts; high wear resistance.
	Polyamide PA 3200GF	45	3.3	6	White; glass-filled polyamide has increased stiffness and is suitable for higher temperature applications
	SOMOS 201	–	0.015	110	Multiple colors available; mimics mechanical properties of rubber
	ST-100c	305	137	10	Bronze-infiltrated steel powder
Electron-beam melting	Ti-6Al-4V	970–1030	120	12–16	Can be heat treated by HIP to obtain up to 600 MPa fatigue strength

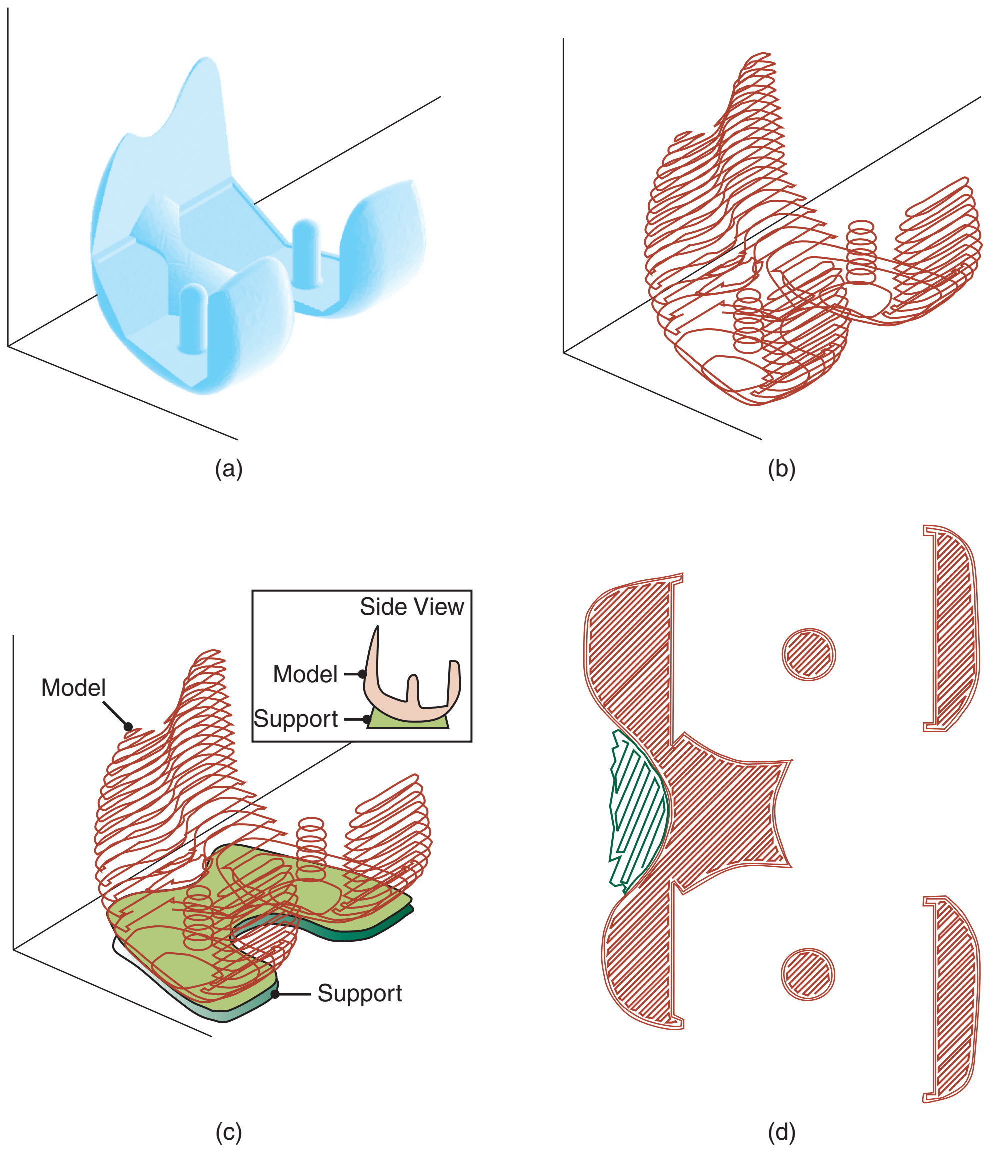


Figure 20.3: The computational steps in producing a stereolithography (STL) file. (a) Three-dimensional description of part. (b) The part is divided into slices; only 1 in 10 is shown. (c) Support material is planned. (d) A set of tool directions is determined to manufacture each slice. Also shown is the extruder path at section A–A from (c) for a fused-deposition-modeling operation.

path of the extruder in one slice, using the fused-deposition-modeling operation (Section 20.3.1). Similar paths will also be planned for the traverse of a laser in a powder bed process (Section 20.6). Other processes, such as binder jetting or CLIPS, do not require a path to be generated, but still need a definition of the desired slice. Triangular tessellation of surfaces has become an industry standard and is widely used for the geometry definition (see Section 38.4.2).

The production of a path requires operator input, both in the setup of the proper computer files and in the initiation of the production process. Following this stage, the machines generally operate unattended and produce a rough part after a few hours, or a few minutes for smaller parts. The part is then subjected to a series of finishing operations, such as sanding and painting, in order to complete the process.

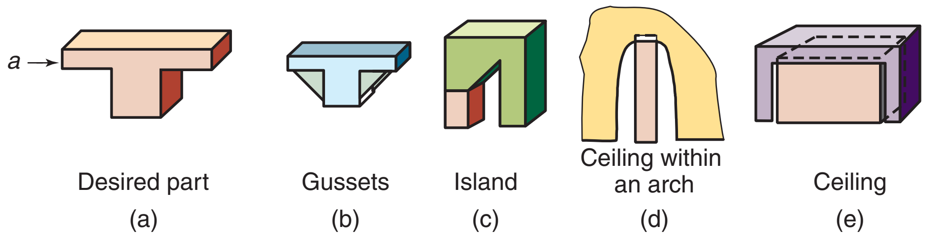


Figure 20.4: (a) A part with a protruding section that requires support material. (b)–(e) Common support structures used in additive manufacturing machines. *Source:* After P.F. Jacobs.

The setup and finishing operations are very labor intensive and the production time is only a portion of the time required to make a *functional* prototype. In general, additive processes are much faster than subtractive processes for limited production runs, taking as little as a few minutes to a few hours to produce a part.

Supports. Complex parts, such as that shown in Fig. 20.4a, may be difficult to build directly. For such processes as fused deposition modeling or stereolithography, a common difficulty is encountered once the part has been constructed up to height a. The next slice would require the filament to be placed at a location where no material exists to support it. The solution is to produce a support material separately from the modeling material, as shown in Fig. 20.4b. Note that the use of such structures allows all of the layers to be supported by the material directly beneath them. The support is made of a less dense, less strong, or a soluble material, so that it can be removed after the part is completed.

20.3 Extrusion-based Processes

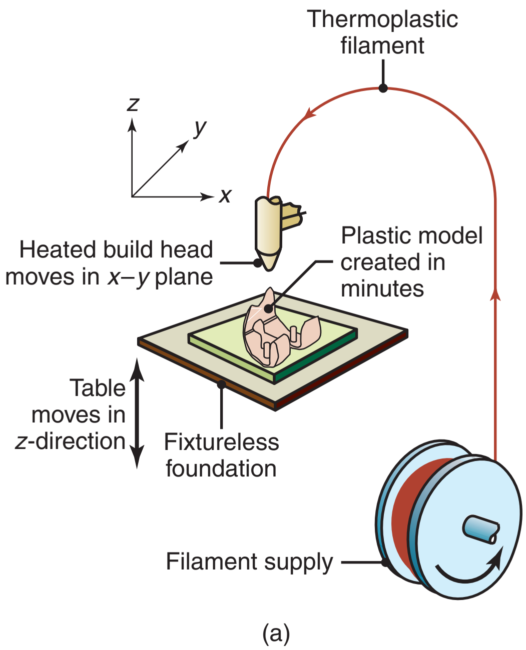
20.3.1 Fused-deposition Modeling

In the *fused-deposition-modeling* (FDM) or *fused filament fabrication* (FFF) process (Fig. 20.5), a gantry-robot controlled extruder head moves in two principal directions over a table, which can be raised and lowered as required. The extruder head is heated and extrudes a (usually thermoplastic) polymer filament through a small orifice at a constant rate. The head follows a predetermined path (see Fig. 20.3d); the extruded polymer bonds to the previously deposited layer. The initial layer is placed on a foam foundation or other base. When a layer is completed, the table is lowered so that the next layer can be superimposed over the previous one. When the part is finished, it can easily be removed.

In the FDM process, the extruded layer's thickness is typically 125–325 μm; this thickness limits the best achievable dimensional tolerance in the vertical direction. In the x-y plane, however, dimensional accuracy can be as fine as 0.025 mm, as long as a filament can be extruded into the feature. Close examination of an FDM-produced part will indicate that a *stepped* surface exists on oblique exterior planes. If the roughness of this surface is unacceptable, subsequent polishing or smoothing with a heated tool can be performed. Also, a coating can be applied, often in the form of a polishing wax. Unless care is taken in applying these finishing operations, the overall dimensional tolerances may be compromised.

An extreme application of FDM is **big area additive manufacturing** (BAAM), which can produce parts as large as 6 m $\times$ 2.3 m $\times$ 1.8 m, with a positioning accuracy of 25 μm. The feedstock in this process is injection molding compound (pellets, sometimes with carbon fiber reinforcement) instead of a filament, so that material costs are significantly lower than in other additive manufacturing processes. Even so, the filament in FDM can be a low cost material, often around \$20–\$40 per kilogram of spooled filament.

Upon expiration of the initial patents for fused deposition modeling, a large number of machines based on FDM have been developed. Some do-it-yourself machines are now freely available as plans that can be downloaded from the Internet. Alternatively, some very inexpensive **desktop machines** have been marketed, based on these crowd-sourced designs, such as the system shown in Fig. 20.6.



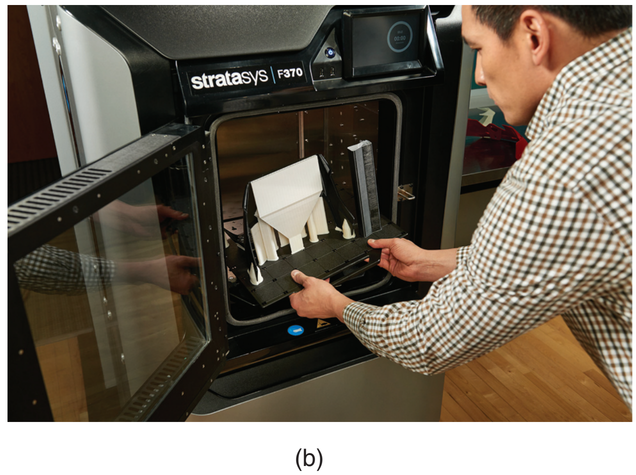


Figure 20.5: (a) Schematic illustration of the fused-deposition-modeling process. (b) Removing a part from an F370, a popular fused-deposition-modeling machine. *Source:* Courtesy of Stratasys, Inc.

Figure 20.6: Low-cost additive manufacturing machine. The F1000, based on digital light printing stereolithography (see Section 20.4.1). The maximum build space is 125 mm × 70 mm × 120 mm. *Source:* Courtesy of 3D Systems.

(a) (b)

Figure 20.7: Continuous fiber fabrication (CFF). (a) Motorcycle brake lever produced with CFF using continuous carbon fiber reinforcement. (b) The Mark X CFF machine. *Source:* Courtesy of Markforged, Inc.

The general trend for FDM materials is that the higher strength polymers require a higher processing temperature. Thus, stronger materials are more difficult to process, and warpage will be a greater concern. One of the differences between desktop systems and industrial FDM machines is the ability to process materials with better mechanical properties in the latter. Acrylonitrile butadiene styrene (ABS) and polylactic acid (PLA) are commonly used for prototyping and in desktop machines; nylon, polycarbonate (PC), and polyetheretherketone (PEEK) are common for high-strength components. FDM materials are available in a wide variety of colors.

A recent development with FDM is continuous fiber fabrication (CFF), where a first extruder prints nylon in the desired pattern. A second head extrudes a continuous carbon, kevlar, or fiberglass fiber inside the part (Fig. 20.7). Control software allows placement of fiber in locations and orientations desired.

Metal parts can be produced through two main methods:

- A plastic filament impregnated with metal powder can be used to produce the desired part. Once completed, the part is sintered to burn off the polymer and fuse the metal, as with powder injection molding (see Section 17.3.3).
- A metal paste can be extruded. This is commonly combined with a second print head that deposited a thermoplastic, allowing for direct inclusion of conductors inside a polymer part.

Low-cost machines have enabled the development of *maker spaces*, where individual designers (typically high school students) are given access to FDM equipment, sometimes for a nominal fee. Along with Internet-based services that accept CAD files, this trend has brought additive manufacturing capabilities to the general public. Moreover, because of the low cost and availability of these machines, researchers are now able to apply new and innovative materials to rapid prototyping machines. Recent novel approaches include printing of food or biological materials for making medical implants, printing of artificial organs (*bioprinting*), clothing, and shoes (Section 20.9).

20.4 Photopolymerization

20.4.1 Stereolithography

A common additive manufacturing process, one that actually was developed prior to fused-deposition modeling, is *stereolithography* (STL), a term coined by Charles W. Hull in 1986. This process (Fig. 20.8) is based on the principle of *curing* (hardening) of a liquid **photopolymer** into a specific shape. A vat, containing a mechanism whereby a platform can be lowered and raised, is filled with a photocurable liquid-acrylate polymer. The liquid is a mixture of acrylic monomers, oligomers (polymer intermediates), and a photoinitiator (a compound that undergoes a reaction upon absorbing light).

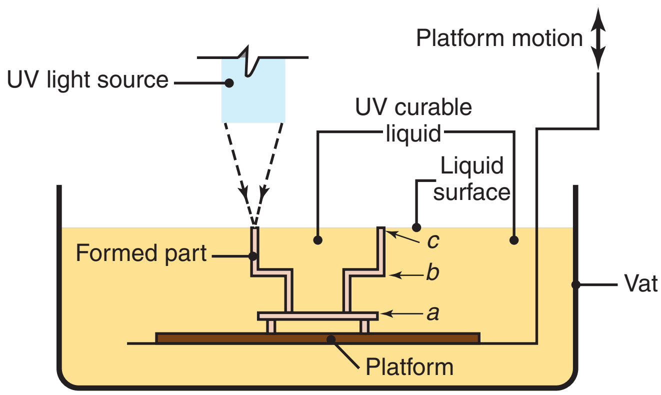


Figure 20.8: Schematic illustration of the stereolithography process.

At its highest position (depth a in Fig. 20.8), a shallow layer of liquid exists above the platform. A *laser*, generating an ultraviolet (UV) beam, is focused upon a selected surface area of the photopolymer, and then moved around in the x–y plane. The beam cures that portion of the photopolymer (say, a ring-shaped portion), and thereby producing a layer of solid body. The platform is then lowered sufficiently to cover the cured polymer with another layer of liquid polymer; the sequence is then repeated until level b in Fig. 20.8 is reached. A cylindrical part, with a constant wall thickness, has thus been generated. Note that the platform has now been lowered by a vertical distance ab.

At level b, the x–y movements of the beam define a wider geometry, thus a flange-shaped portion is being produced over the previously formed segment. After the desired thickness of the liquid has been cured, the process is repeated, producing another cylindrical section between levels b and c. Note that the surrounding liquid polymer is still fluid (because it has not been exposed to the ultraviolet beam), and that the part has been produced from the bottom up in individual *slices*. The unused portion of the liquid polymer can be used again to make another part or another prototype.

Note that the term stereolithography as used to describe this particular process comes from the observations that the movements are three dimensional (hence the word stereo) and the process is similar to lithography (see Section 28.7). Note also that, as in FDM, stereolithography will sometimes require a support material, depending on geometry. In stereolithography, this support often takes the form of porous structures.

After its completion, the part is removed from the platform, blotted, and cleaned ultrasonically with an alcohol bath. The support structure is then removed, and the part is subjected to a final curing cycle in an oven. The smallest tolerance that can be achieved in stereolithography depends on the sharpness of the focus of the laser, typically being around 0.0125 mm. Oblique surfaces also can be produced, with high quality.

Solid parts can be made by applying special laser-scanning patterns to speed up production. For example, by spacing the scan lines in stereolithography, volumes or pockets of uncured polymer can be formed within cured solid shells. When the part is later placed in a postprocessing oven, the pockets are cured and a solid part is produced. Similarly, parts that are to be investment cast (Section 11.3.2) will have a drainable honeycomb structure, which permits a significant fraction of the part to remain uncured.

Total cycle times in stereolithography range from a few hours to one day, without requiring post-processing steps, such as sanding and painting. Depending on their capacity, the cost of the machines is in the range from \$100,000 to \$400,000. The cost of the liquid polymer is on the order of \$80 per liter. The maximum part size that can be produced is 0.5 m $\times$ 0.5 m $\times$ 0.6 m. The layer height in STL is 25–100 μm, depending on the machine, and use a laser spot size of 50–150 μm.

Stereolithography has been used with highly focused lasers to produce parts with micrometer-sized features. The use of optics required to produce such features necessitates the use of thinner layers and lower volumetric cure rates. When used to fabricate micromechanical systems (Chapter 29), this process is called **microstereolithography**.

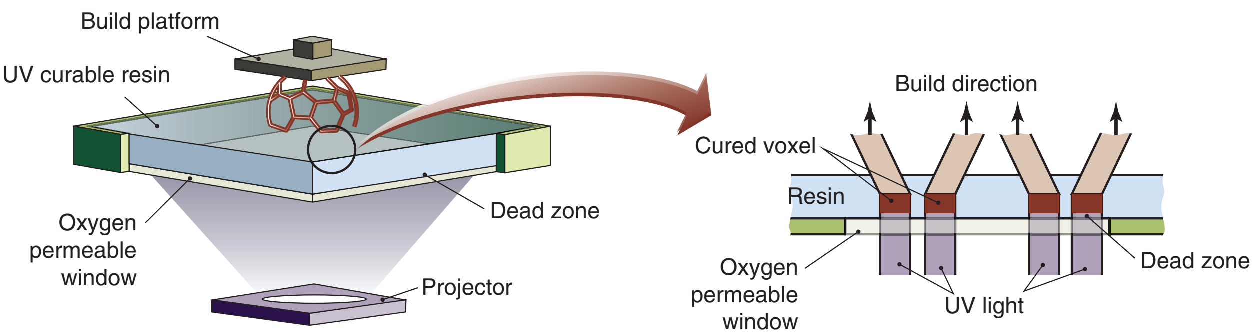


Figure 20.9: The CLIP process. An array of micro mirrors directs light to cure a layer, but the photopolymer next to the oxygen-permeable window does not cure because oxygen is a curing inhibitor. The cured polymer can be pulled out of the liquid photopolymer, with liquid photopolymer flowing into the interface to replenish the build layer.

Another form of stereolithography is *mask projection stereolithography* or *direct light processing* (DLP) with the advantage of much higher rate of part production. In this process, a DLP device made up of millions of microscopic mirrors is used to direct UV light from a lamp or light emitting diode to expose the entire layer at once.

20.4.2 Continuous Liquid Interphase Production

The **continuous liquid interphase production** (CLIP) process is illustrated in Fig. 20.9. CLIP uses a special window that is transparent to light and is permeable to oxygen, much like a contact lens. By controlling the oxygen diffusing through the window, a "dead zone" is created in the resin pool just tens of microns thick where photopolymerization cannot occur, as oxygen acts as an inhibiter. This ensures that a liquid layer will persist adjacent to the optics, regardless of light exposure. The projector transmits light in the desired pattern into the resin pool from underneath, curing the polymer above the dead zone. The build plate pulls the printed physical object out of the vat, at a speed low enough so that the cured material maintains contact with the uncured liquid and new liquid flows into the curing zone. The CLIP process is continuous, but does require discretization of layers from a CAD file and exposure of layers.

The light is projected an entire layer at a time, not in raster fashion as in selective laser sintering or conventional stereolithography. The layer is produced through digital light processing (DLP) hardware that is also common in projector systems and some televisions. A DLP device consists of an array of micromirrors, each of which can direct light towards the build chamber if activated; by activating selected mirrors, "pixels" in the build space are activated, curing the polymer into voxels. A typical voxel dimension is 75 μm. Parts produced through the CLIP process must undergo secondary operations, consisting of, at least, cleaning and either a secondary UV flood cure or thermal cure in an oven. A variety of polymeric material chemistries are now available, and CLIPs can achieve production rates two orders of magnitude higher than other additive manufacturing processes.

Case Study 20.2 Production of Athletic Shoes

CLIP represents a breakthrough additive manufacturing process in that it allows manufacture of parts at high quantities, and provides a strategy for mass production for certain parts (see Section 37.2.2).

Carbon3D, the developer of CLIP, has also developed designs and software to produce a metamaterial (Section 6.16) that has similar mechanical properties as polymer foam, but is easier to clean (Fig. 20.10).

By placing more material or changing the metamaterial design where a higher stiffness is desired, it is possible to produce a shoe insole with tuned stiffness, that can also be varied by location.

Adidas, well known for its athletic products and materials, partnered with Carbon3D to produce high performance footware, called the Futurecraft 4D (Fig. 20.11). Adidas had collected and maintained athlete data, which was used for the development of a stiffness-tuned midsole for the new shoe design.

Adidas and Carbon3D have developed a digitized footwear-component creation process that eliminates the need for traditional prototyping or molding. CLIP also allowed Adidas to create a monolithic midsole that addresses precise needs related to movement, cushioning, stability, and comfort. Further, over the course of product development, CLIP enabled Adidas to evaluate more than 50 design iterations, a substantial increase when compared with what is achievable with traditional injection molding in the same amount of time. Moreover, engineers from both companies collaborated closely and tested nearly 150 resin iterations.

The final midsole material is made of a dual cure resin that generates a polyurethane upon thermal cure. It is a stiff elastomer printed in a lattice structure to develop a high-performance midsole that also offers excellent durability and is aesthetically pleasing. Adidas expects to produce 100,000 such annually.

Source: Courtesy Steven Pollack, Carbon and Adidas.

Figure 20.10: The use of a metamaterial shoe sole. Note that the pattern in the metamaterial changes with location to obtain a desired mechanical performance. *Source:* Courtesy of Carbon, Inc.

20.5 Material Jetting

Material jetting (MJ) is a class of additive manufacturing processes that include *drop on demand* (DOD) and the *Polyjet* related processes, also known as *multijet modeling* (MJM). The main difference is that PolyJet uses photopolymer feedstocks; DOD uses thermoplastics or wax. In both cases, a low viscosity is needed to produce droplet jets, which may require preheating of the build material.

The **PolyJet process** is a form of material jetting where print heads deposit a photopolymer on the build tray. Ultraviolet bulbs, alongside the jets, instantly cure and harden each layer, thus eliminating the need for any postmodeling curing that is required in stereolithography. PolyJet results in a smooth surface with layers as thin as 16 μm that can be handled immediately after the process is completed.

Two different materials are used: the material for the actual model, and a gel-like resin for support, such as shown in Fig. 20.4. Each material is simultaneously jetted and cured, layer by layer. When completed,

Figure 20.11: The Adidas Futurecraft 4D shoe, using a CLIP-produced metamaterial sole. *Source:* Courtesy of Carbon, Inc.

the support material is removed by soaking in an aqueous solution. Build sizes have an envelope of up to 500 mm $\times$ 400 mm $\times$ 200 mm. The PolyJet process has capabilities similar to those of stereolithography and uses similar resins (Table 20.2). The main advantages of this process are the capabilities of avoiding part cleanup and lengthy post-process curing operations and the much thinner layers produced, thus allowing for better resolution.

In DOD, a stream of a material droplets are ejected through a small orifice and deposited on a surface (target), using an ink-jet type mechanism. A second print head deposits a support material that is soluble in water or related solvent. DOD is commonly used for producing investment casting patterns (see Section 11.3.2). DOD is considered the most accurate form of 3D printing because of the absence of thermal stresses and because it is mainly used for small parts. Generally, a tolerance of ± 0.1 mm can be achieved.

A recent innovation is the **nano particle jetting** (NPJ) process, which uses a suspension of nanoparticles in a liquid carrier as the liquid printed in material jetting. It was noted previously that particles smaller than around 20 μm are difficult to spread onto a build chamber because they can easily become airborne and interfere with optics and lasers. The liquid carrier prevents the entrainment of small particles into air, and therefore allows the incorporation of much smaller particles than other processes.

The build chamber is heated sufficiently to evaporate the carrier and bond the nanoparticles. Once completed, the particles are sintered to create fully dense parts. The use of small particles allows printing of detailed features and the development of superior mechanical properties.

20.6 Powder Bed Processes

Powder Bed Processes involve a number of approaches that utilize powder as the workpiece material, and where the powder is deposited layer-by-layer in a *bed* or *build chamber*. Several powder application systems are used, but they typically involve a counter-rotating roller or a wiping mechanism. The deposited powder has limited green strength but can serve as a support for complicated parts.

Powder spreading is a critical step in these processes. Some of the considerations associated with powder spreading are as follows:

1. The powder must be capable of spreading into thin layers. Because this can be compromised by moisture, polymer powders may need to be dried before they can be effectively spread.

2. Mean particle sizes of the powder are around the layer thickness, or slightly smaller; there is a range of powder sizes that can be used. Particles larger than the layer thickness will be pushed ahead of the wiper or the roller and are less likely to be part of a spread layer; particles that are too small are likely to become airborne and adhere to exposed surfaces.

3. The powder may be preheated to reduce laser or electron beam power required for melting.

4. Powder explosions or fire can result from static electric discharge; thus, safety protocols regarding equipment and worker grounding, oxygen-free shielding gases, and increased humidity must be carefully followed.

20.6.1 Selective Laser Sintering

Selective Laser Sintering (SLS) is a process based on sintering (Section 17.4) of nonmetallic powders selectively into an individual object. *Direct Metal Laser Sintering* (DMLS) or *Selective laser melting* (SLM) is a related process used with metals; in both cases, material is always at least partially melted. The basic elements in this process are shown in Fig. 20.12. Note that the bottom of the processing chamber is equipped with two cylinders:

1. A powder-feed cylinder, which is raised incrementally to supply powder to the part-build cylinder.

2. A part-build cylinder, which is lowered incrementally as the part is being shaped.

In the SLS process, a thin layer of powder is first deposited in the part-build chamber. Then a laser beam, guided by a process-control computer using instructions generated by the three-dimensional CAD program of the desired part, is focused on that layer, tracing and sintering a particular cross section into a solid mass. The powder in other areas remains loose, but this powder can support the sintered portion; with some processes, separate support structures may still be needed. Another layer of powder is then deposited, and the cycle is repeated continuously until the entire three-dimensional part has been produced.

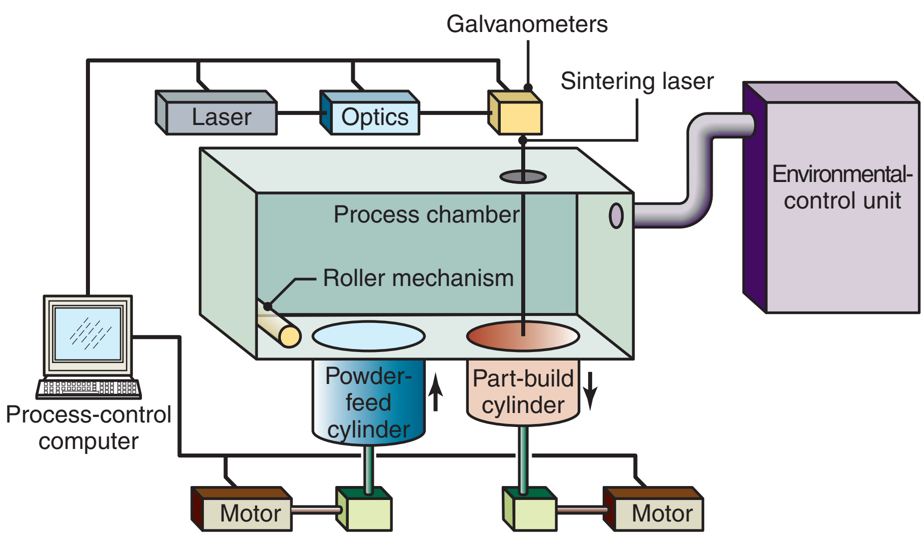


Figure 20.12: Schematic illustration of the selective-laser-sintering process.

The feed chamber contains the desired part supported by unfused powder, called the *cake*, which has very low green strength. The loose particles are shaken or brushed off, and the part is recovered. The part may not require further curing to develop strength. The supports are removed or machined, and the part is then post-processed as required.

A variety of materials can be used in this process, including polymers (such as acrylonitrile butadiene styrene, polyvinyl chloride, nylon, polyester, polystyrene, and epoxy), wax, metals, and ceramics (with appropriate binders).

Generally the feed and build chambers are preheated. A major concern is the thermal management in the build chamber, as is the control of the powder during spreading and lasing. Powder that becomes airborne in the build chamber could coat sensors or optics and compromise productivity. Modern SLS machines therefore use a constant flow of shielding gas that directs the powder away from sensitive machine elements. Shielding gases vary with the powder material; argon and nitrogen are common options.

SLS parts are susceptible to shrinkage and warpage due to thermal stresses. Each layer that is built is produced on a previous layer; as the new layer cools, it shrinks, which can cause a part to curl upwards. In extreme cases, the part can collide with the wiper or roller depositing a new powder layer, necessitating a build to be aborted. Despite thermal effects, dimensional tolerances of ± 0.1 mm can be achieved with well-designed parts. Layer thickness range from around 30–100 μm depending on the material. Stainless steel and titanium alloys generally produce the best part fidelity because of their lower thermal conductivities.

SLS has a number of advantages over other AM processes. The material is generally isotropic and accurate (although not as good as stereolithography or material jetting) with very good mechanical properties. SLS does not usually need support materials, except for metal parts that are bonded to a build plate by support material to prevent part curl. The main drawback is the high cost of machines; metal-capable SLS machines cost around $300,000 for low-end machines, and can cost almost $1 million.

20.6.2 Electron-beam Melting

A process similar to selective laser sintering and *electron-beam welding* (Section 30.6), **electron-beam melting** (EBM) uses the energy source associated with an electron beam to melt titanium or cobalt-chrome powder to make metal prototypes. The workpiece is produced in a vacuum, making the part build size limited to around $200 \times 200 \times 180$ mm. Electron-beam melting is up to 95% efficient from an energy standpoint, as compared with 10–20% efficiency for selective laser sintering.

In EBM, the supply powder and the build chamber are heated to near the material's melting point, significantly reducing the energy needed to melt the metal and also reducing thermal stresses. Because the build chamber is at an elevated temperature, the melted metal solidifies more slowly and results in more fully dense parts. A volume build rate of up to 60 cm^3/hr can be obtained, with individual layer thicknesses of 0.050–0.200 mm. Parts may also be subjected to hot isostatic pressing (Section 17.3.2) to improve their fatigue strength. Although applied mainly to titanium and cobalt-chrome alloys to date, the process is being developed also for stainless steels, aluminum, and copper alloys.

20.6.3 Binder-jet Printing

In the *Binder-jet Printing* (BJP) process, also known as *Binder Jetting* or *Three-Dimensional Printing*, a print head deposits an inorganic binder material onto a layer of sand, polymer, ceramic, or metallic powder, as shown in Fig. 20.13. A piston, supporting the powder bed, is lowered incrementally and with each step a layer is deposited and then fused by the binder.

Binder-jet printing allows considerable flexibility in the choice of materials and binders used. Common powder materials are polymers (sometimes blended with fibers), metals, and foundry sand. Since multiple binder print heads can be incorporated into one machine, it is possible to produce full-color prototypes by having different color binders (Fig. 20.14). The effect is a three-dimensional analog to printing photographs using three ink colors (red, cyan, and blue) in an ink-jet printer.

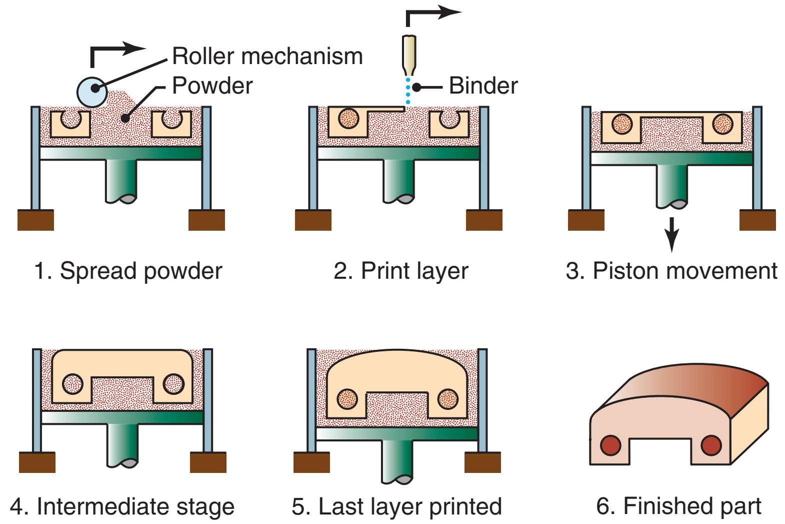


Figure 20.13: Schematic illustration of the binder-jet printing process.

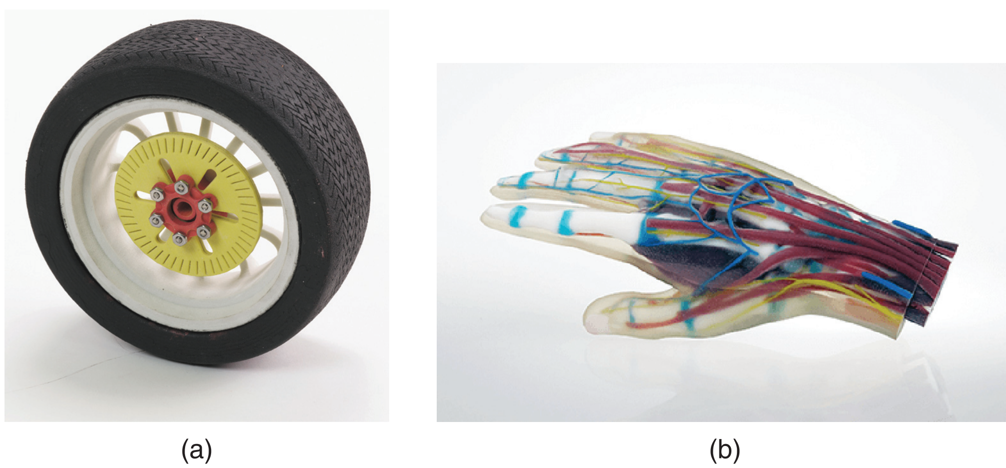


Figure 20.14: Full color parts produced by binder jet printing. (a) A simple model of toy truck wheel; (b) a more detailed model of a human hand, with transparent and colored components. *Source:* Courtesy of Stratasys.

A typical part produced by BJP from ceramic powder is a ceramic-casting shell (see Section 11.2.4), in which aluminum-oxide or aluminum-silica powder is fused with a silica binder. The molds are post-processed in two steps: (a) curing at around 150°C and (b) firing at 1000° to 1500°C. Printing of sand molds is a common practice; molds can be produced with blind risers and with cores, thus avoiding the complicated assembly operations associated with copes and drags (see Section 11.2.1).

The parts produced through the BJP process are somewhat porous, and thus may lack strength. Three-dimensional printing of metal powders can also be combined with sintering and metal infiltration (see Section 17.4) to produce fully-dense parts, using the sequence shown in Fig. 20.15. Here, the part is produced as before by directing the binder onto powders. However, the build sequence is then followed by sintering in order to burn off the binder and partially fuse the metal powders, just as is done in powder injection molding (Section 17.3.3). Common metals used in 3DP are stainless steels, aluminum, and titanium.

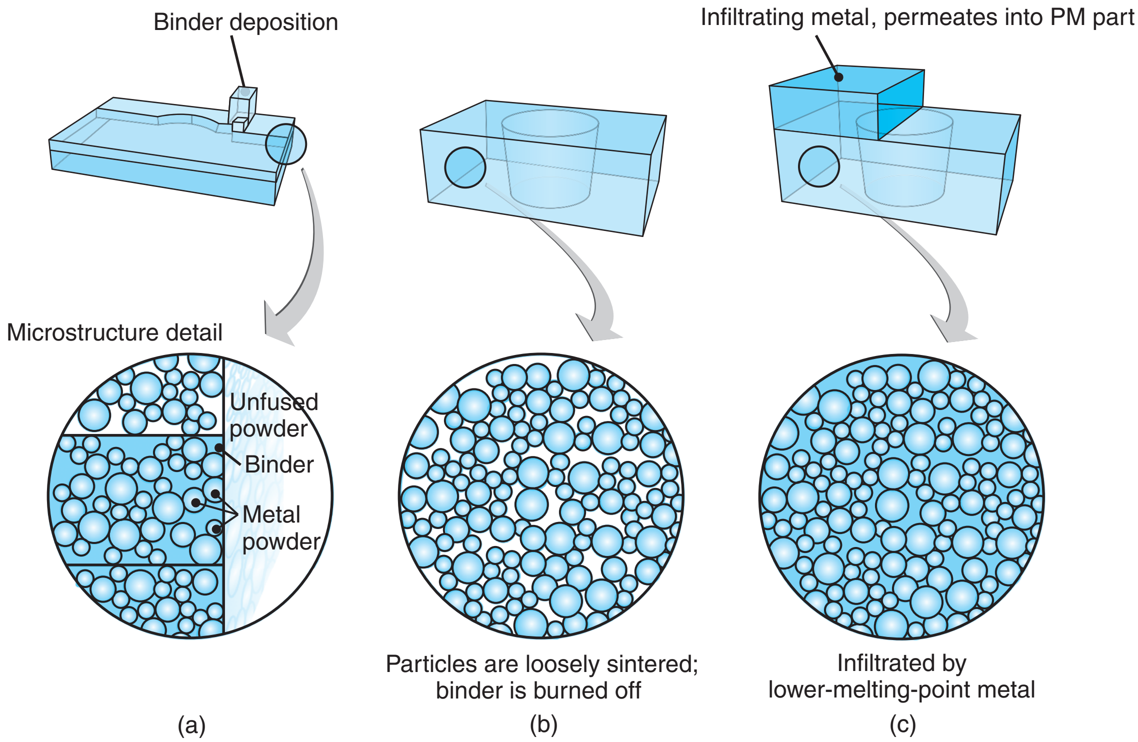


Figure 20.15: Three-dimensional printing using (a) part-build, (b) sinter, and (c) infiltration steps to produce metal parts. *Source:* Courtesy of Kennametal Extrude Hone.

The infiltrating materials typically are copper and bronze, which provide good heat-transfer capabilities as well as wear resistance. This approach represents an efficient strategy for rapid tooling (Section 20.10).

Dimensional tolerances vary widely by machine manufacturer and feedstock in BJP. Sand molds and cores are commonly produced with layer thicknesses of 240–380 μm, but layers may be as low as 50 μm with some materials.

A more recently developed process, known by its trade name of **jet fusion**, is based on BJP, but with a number of unique features. In jet fusion,

- A powder of polymer is spread in a build chamber.
- Binder is jetted onto the polymer to fuse the powder as desired in the layer.
- A *detailing agent* is jetted adjacent to the regions where the binder had been applied.
- The layer is then subjected to a heat source that cures the polymer containing binder.

The function of the detailing agent needs some clarification. In any thermal curing approach, temperatures are difficult to control and can lead to poor part resolution. The detailing agent prevents curing, so that the boundary between the cured part and the unaffected polymer has very good definition with sharp and smooth edges.

In addition, the jet fusion process uses an array of sensors to determine the temperature distribution in the build chamber. If the sensors determine an area of the bed has too high or too low of a temperature compared to the optimum, the intensity of the UV light over the build chamber is varied accordingly, leading to improved mechanical properties and dimensional accuracy.

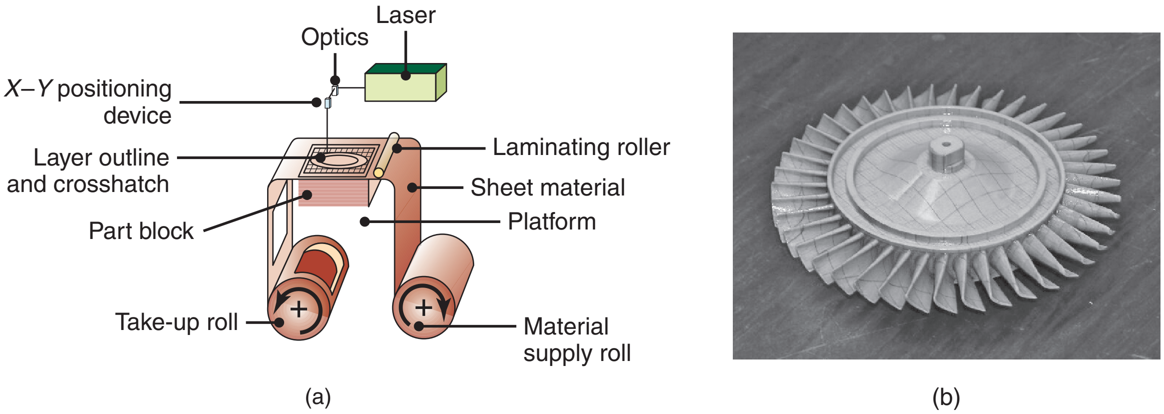


Figure 20.16: (a) Schematic illustration of the laminated-object-manufacturing process. (b) Turbine prototype made by LOM. *Source:* Courtesy of M. Feygin, Cubic Technologies, Inc.

20.7 Laminated-object Manufacturing

Lamination involves laying down layers bonded adhesively to one another. Several variations of *laminated-object manufacturing* (LOM) are now available.

Producing parts by LOM systems can be elaborate, where the more advanced systems use layers of paper or plastic with a heat-activated glue on one side. The desired shapes are burned into the sheet with a laser, and the parts are built layer by layer (Fig. 20.16). On some systems, the excess material must be removed manually after the part is made; the removal is simplified by programming the laser to burn perforations in crisscrossed patterns. The resulting grid lines make the part appear as if it had been constructed from gridded paper, similar to graph paper.

20.8 Miscellaneous Processes

20.8.1 Laser-engineered Net Shaping

Laser-engineered net shaping (LENS), also known as *laser powder forming* (LPF) involves the principle of using a laser beam to melt and deposit metal powder or wire, layer by layer, over a previously deposited layer (Fig. 20.17). The heat input and cooling have to be controlled precisely to develop a favorable microstructure.

The deposition process is carried out inside an enclosed volume and in an argon environment, to avoid the adverse effects of oxidation, particularly on aluminum. It is suitable for a wide variety of metals and specialty alloys for the direct manufacturing of parts, including fully-dense tools and molds. The process can also be used for repairing thin and delicate components. There are other, similar processing methods, including *controlled-metal buildup* (CMB) and *precision-metal deposition* (PMD, a trade name).

LENS has been found suitable for incorporation into *hybrid machines* that have both additive and subtractive (machining) manufacturing capabilities. The advantages are that complex shapes can be quickly produced without refixturing, with high dimensional tolerance and surface finish, and with little scrap. Usually, this operation involves the incorporation of a LENS deposition head in combination with a machining or turning center (see Section 25.2). This is a compelling combination, since LENS on its own does not maintain tight tolerances; ± 1 mm is typical of the process limitations.

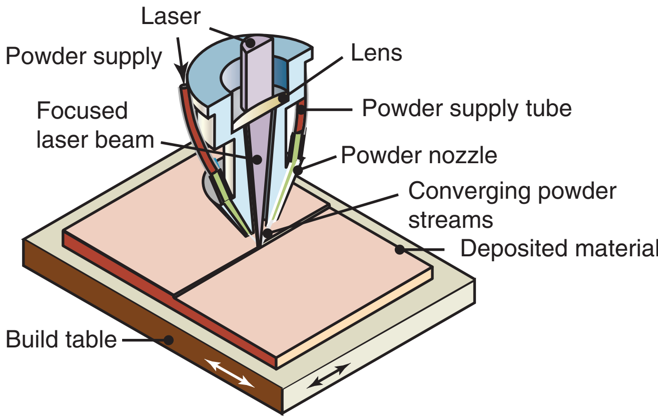


Figure 20.17: Schematic illustration of the laser-engineered net shaping (LENS) process.

20.8.2 Friction Stir Modeling

The **friction stir modeling** (FSM) process shares several similarities with friction stir welding (Section 31.4). In this process, powder is delivered to a build location by pushing it into a rotating tube. The friction between the powders and the substrate are sufficiently high to densify the powder and develop a solid material. Because the process is at solid state, there is no appreciable heat-affected zone.

Friction stir modeling has been successfully applied to magnesium, aluminum, and titanium, and has the advantage of being able to change the deposited material during a build. For example, a lightweight aluminum part can possess an integral hardened surface for wear resistance.

The equipment used for FSM involves conventional CNC milling machines (Section 24.2), modified to deliver the desired powder. Typical layer thickness is around 100 μm, and surface finish is generally poor, requiring subsequent machining to achieve smooth surfaces; a machining allowance is therefore essential.

20.8.3 Wire and Arc Additive Manufacturing

As described in Section 30.4.3, gas metal arc welding and gas tungsten arc welding are commonly incorporated into robot welding systems; they involve material transfer from an electrode or filler material into a weld joint. The same approach can be used to deposit material in a controlled manner; a welding end-effector on a robot provides a platform for large volume, large deposition rate additive manufacturing. This arrangement is known as *wire and arc additive manufacturing* (WAAM), with a unique feature of the ability to produce designs that are not based on layers; the robot can follow any trajectory.

20.8.4 Hybrid Approaches

Additive manufacturing has specific advantages in certain applications, but one of the drawbacks compared to machining is the inability to hold tight tolerances or to achieve a desired surface finish. One solution is to combine additive and subtractive processes in the same machine.

To date, the most common **hybrid approaches** involve combining either laser engineered net shaping or selective laser sintering with a CNC machining center. With LENS, a part will generally be produced with a generous machining allowance, since the process can rarely hold tolerances better than a millimeter over 50 mm. The constructed part is then machined without refixturing.

With SLS, the machining operations are performed after a layer or a group of layers is produced. Even though the build chamber can be disturbed by machining, fresh powder fills in machined areas while the wiper spreads the powder into smooth layers.

A wide range of capabilities have recently been developed into the machinery, including the ability to combine materials to produce composites or to tailor mechanical properties, the incorporation of sensors to detect defects during printing, and automated handling of completed parts.

20.9 Emerging AM Applications

The low cost and high reliability of additive manufacturing has directly led to its widespread application in a number of areas that are far removed from industrial production.

Bioprinting. The production of medical devices is well established and can also be done by additive manufacturing. An emerging area is *bioprinting*, involving the printing of living cells. Using processes related to fused-deposition modeling or binder-jet printing, cells are suspended in a liquid carrier or *bioink*, producing a construct. This approach allows printing of cells in desired structures and concentrations.

While applications are emerging, bioprinting has the potential and aggressive goal of producing living functional tissue, such as organ transplants or tissue that can be used in drug studies. Current limitations are the cell survival during and after the printing process.

Architectural applications. Processes that are based mainly on fused deposition modeling have been used to build buildings or various structures from extruded concrete. The approaches generally use tower robots that place concrete along the periphery of the desired structure. Trowels can be located near the extruder head to build near-vertical walls that match from layer to layer, a process variant called *contour crafting*. Alternative approaches use binder-jet printing with sand to produce structures. Permanent habitats on the moon or on Mars are now expected to be produced through additive manufacturing, using as much native soil as possible, given the cost of transport of materials.

20.10 Direct Manufacturing and Rapid Tooling

While extremely beneficial as a demonstration and visualization tool, additive manufacturing processes also have been used to produce functional parts. There are two basic methodologies involved:

1. Direct production of engineering metals, ceramics, and polymer components or parts.
2. Production of tooling or patterns by additive manufacturing, for use in various manufacturing operations.

Additive manufacturing operations can be also used to manufacture parts directly, referred to as **direct manufacturing**. This approach includes the case where a part involves a machining or grinding allowance or requires further finishing operations. Thus, the component is generated directly to a near-net shape, from a computer file containing part geometry. The main limitations to the widespread use of additive manufacturing for direct manufacturing, or *rapid manufacturing*, are as follow:

- Raw-material costs are high, and the time required to produce each part is too long to be viable for large production runs. However, there are many applications in which production runs are sufficiently small to justify direct manufacturing through additive manufacturing technologies, or where the required material properties are attainable.
- The long-term and consistent performance of rapidly manufactured parts (as compared with the more traditional methods of manufacturing them) should be considered, especially with respect to fatigue, wear, and life cycle.

Much progress is being made to address and respond to these concerns in order to make rapid manufacturing a more competitive and viable option in manufacturing.

Several methods have been devised for the rapid production of tooling (RT) by means of additive manufacturing processes. The advantages to rapid tooling include the following:

1. The high cost of labor and the shortening supply of skilled patternmakers can be overcome.
2. There is a major reduction in lead time.
3. The integral use of CAD technologies allows the use of modular dies, with base-mold tooling (match plates) and specially fabricated inserts.
4. Chill- and cooling-channel placement in molds can be optimized more easily, leading to reduced cycle times. *Conformal cooling* is a strategy for producing cooling channels that are located in a way to maximize heat extraction from a mold or die, while preserving mechanical strength (Fig. 20.18).
5. Shrinkage due to solidification or to thermal contraction can be compensated for automatically, through software, to produce tooling of the proper size and, in turn, to produce the desired parts. Large flat parts should be oriented at an angle or vertically to minimize the cross-sectional area of each layer, thereby minimizing warpage.

The main shortcoming of rapid tooling is the potentially reduced tool or pattern life, as compared to those obtained from machined tool and die materials, such as tool steels and tungsten carbides (Chapter 21).

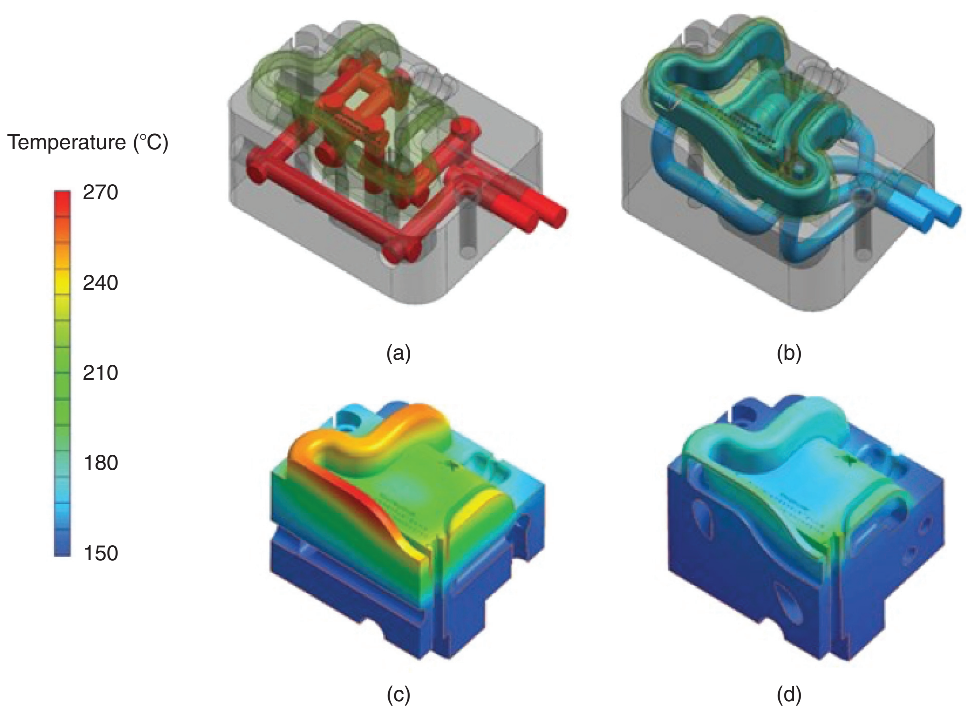


Figure 20.18: The benefit of conformal cooling in molds produced by additive manufacturing. The images on the left show conventional (machined or drilled) cooling channels, and those on the right show conformal cooling channels that can be produced in additive manufactured molds. The top images depict the channel layout; the bottom images the temperature distributions in the mold during production. Note that the temperature distribution is more uniform on the molds with conformal cooling, leading to less warpage and higher production rates. *Source:* Copyright image provided courtesy of Milacron product brand DME Company showing their TruCoolTM technology.

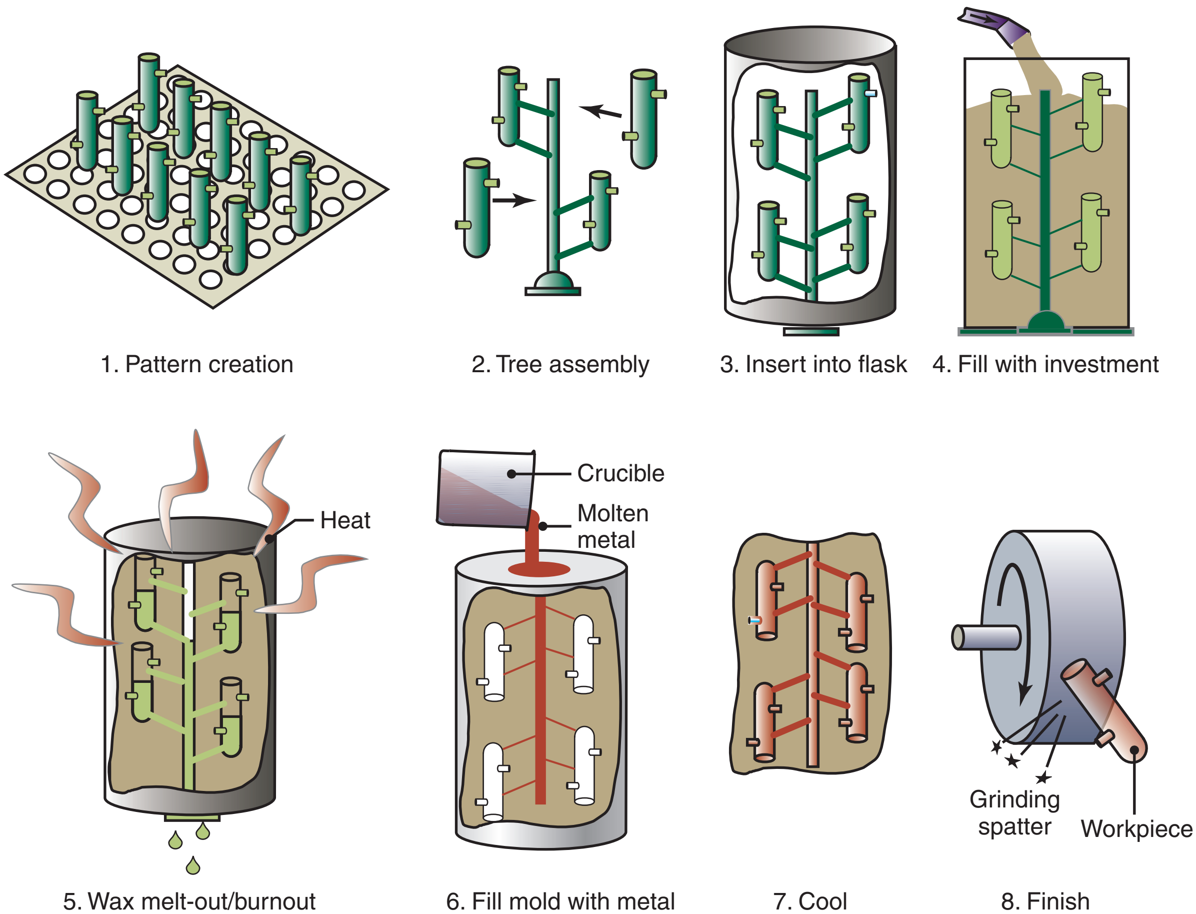


Figure 20.19: Manufacturing steps for investment casting with rapid-prototyped wax parts as blanks. This method uses a flask for the investment, but a shell method also can be used. *Source:* Courtesy of 3D Systems, Inc.

A number of strategies have developed to incorporate additive manufacturing into mold and die production. As an example, Fig. 20.19 shows an approach for investment casting. Here, the individual patterns are made in an AM operation (in this case, stereolithography), and then used as patterns in assembling a tree for investment casting (Fig. 11.14). Note that this approach will require a polymer that completely melts and burns from the ceramic mold; such polymers are available for all forms of polymer AM operations. Furthermore, as drawn in CAD programs, the parts are usually software modified to account for shrinkage, and it is the modified part that is produced in the additive manufacturing machinery.

Binder-jet printing also can easily produce a ceramic-mold casting shell (Section 11.2.2) or a sand mold (Section 11.2.1), in which an aluminum-oxide or aluminum-silica powder is fused with a silica binder. The molds have to be postprocessed in two steps: curing at around 150°C, and then firing at 1000°–1500°C.

Another common application of rapid tooling is injection molding of polymers (Section 19.3), in which the mold or, more typically, a *mold insert* is manufactured by additive manufacturing. Molds for slip casting of ceramics (Section 18.2.1) also can be produced in this manner. To produce individual molds, AM processes are used directly, and the molds will be shaped with the desired permeability. For example, in fused-deposition modeling, this requirement mandates that the filaments be placed onto the individual slices, with a small gap between adjacent filaments; the filaments are then positioned at right angles in adjacent layers.

The advantage of rapid tooling is the capability to produce a mold or a mold insert that can be used to manufacture components without the time lag (typically several months) traditionally required for the procurement of tooling. Moreover, the design is simplified, because the designer needs to analyze only a CAD file of the desired part; software then produces the tool geometry and automatically compensates for shrinkage.

In addition to the straightforward application of additive manufacturing technology to tool or pattern production, other rapid-tooling approaches, based on AM technologies, have been developed.

Room-temperature vulcanizing (RTV) **molding/urethane casting** can be performed by preparing a pattern of a part by any AM operation, which is then used to produce an RTV mold. The pattern is first coated with a parting agent, and may or may not be modified to define mold parting lines. Liquid RTV rubber is then poured over the pattern, and cures (usually within a few hours) to produce mold halves. The mold is then used with liquid urethanes in injection molding or reaction-injection molding operations (Section 19.3). One main limitation of this approach is a lower mold life, because the polyurethane present in the mold causes progressive damage and the mold may be suitable only for as few as 25 parts.

Epoxy and *aluminum-filled epoxy* molds also can be produced, but mold design requires special care. With room temperature vulcanizing (RTV) rubber, the flexibility of the mold allows it to be peeled off the cured part. With epoxy molds, their high stiffness precludes this method of part removal, and mold design is more complicated. Thus, for example, drafts are required, and undercuts and other design features that can be produced by RTV molding must be avoided.

Acetal clear epoxy solid (ACES) *injection molding*, also known as *direct AIM*, refers to the use of additive manufacturing, usually stereolithography, to directly produce molds suitable for injection molding. The molds are shells, with an open end to allow filling with a material such as epoxy, aluminum-filled epoxy, or a low-melting-point metal. Depending on the polymer used, mold life may be as few as 10 parts, although hundred parts per mold are possible.

Sprayed-metal tooling. In this process, shown in Fig. 20.20, a pattern is first created through AM. A metal spray operation (Section 34.5) then coats the pattern surface with a zinc-aluminum alloy. The metal coating is placed in a flask, and potted with an epoxy or an aluminum-filled epoxy material. In some applications, cooling lines can be incorporated into the mold before the epoxy is applied. The pattern is removed, and two such mold halves are used as in injection-molding operations. Mold life is highly dependent on the materials used and the temperatures involved, and can vary from a few to thousands of parts.

Keltool process. In the *Keltool process*, an RTV rubber mold is first produced, based on a rapid-prototyped pattern, as described earlier. The mold is then filled with a mixture of powdered A6 tool steel (Section 5.7), tungsten carbide, and polymer binder, and is allowed to cure. The so-called *green* tool (green, as in ceramics and powder metallurgy) is fired to burn off the polymer and fuse the steel and the tungsten-carbide powders. The tool is then infiltrated with copper in a furnace to produce the final mold.

The mold can subsequently be machined or polished to impart a superior surface finish and good dimensional tolerances. Keltool molds are limited in size to around $150 \times 150 \times 150$ mm. Thus, a mold insert, suitable for high-volume molding operations, is made and installed. Depending on the material and processing conditions, mold life can range from 100,000 to 10 million parts.

Case Study 20.3 Casting of Plumbing Fixtures

A global manufacturer of plumbing fixtures and accessories for baths and kitchens used rapid tooling to transform its development process. One of the company's major product lines is decorative water faucets, made from brass castings that are subsequently polished to achieve the desired surface finish. The ability to produce prototypes from brass is essential for quickly evaluating designs and identifying processing difficulties that may occur.

A new faucet design was prepared in a CAD program; the finished product is shown in Fig. 20.21. As part of the product development cycle, it was decided to produce prototypes of the faucet to confirm the aesthetics of the design. Since such faucets are typically produced by sand casting, it was also essential to validate the design through a sand-casting operation, followed by polishing. This approach allowed evaluation of the cast parts in terms of porosity and various other casting defects, and also would identify processing difficulties that might arise in the finishing stages.

A sand mold was first produced, as shown in Fig. 20.22. The mold material was a blend of foundry sand, plaster, and other additives that were combined to provide strong molds with good surface finish (see also Section 11.2.1). A binder was printed onto the sand mixture to produce the mold. The mold could be produced as one piece, with an integral core (see Figs. 11.3 and 11.6), but in practice, it is often desired to smoothen the core and assemble it later onto core prints. In addition, slender cores may become damaged, as support powder is being removed from the mold, especially for complex casting designs. Therefore, the core for this design is produced separately and then assembled into the two-part mold.

Using 3D printing, the operation produced brass prototypes of the faucets in five days, which included the time required for mold design, printing, metal casting, and finishing. The actual print time of the mold was just under three hours, and the material cost was approximately $280. The production of pattern plates for sand casting is, in general, too expensive for producing prototypes, and would cost over $10,000 and add several months to the lead time. The incorporation of 3D printing into the design process thus provided new capabilities that confirmed the design aesthetics and function, as well as manufacturing robustness and reliability.

Source: Courtesy of 3D Systems.

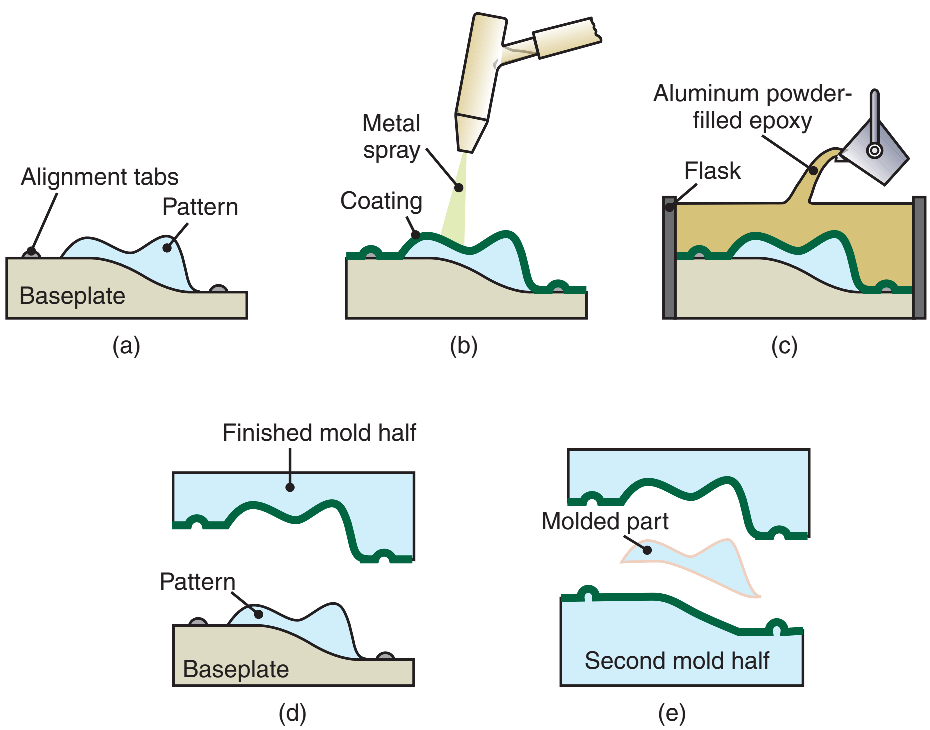


Figure 20.20: Production of tooling for injection molding by the sprayed-metal tooling process: (a) A pattern and baseplate are prepared through a additive manufacturing operation; (b) a zinc–aluminum alloy is sprayed onto the pattern (see Section 34.5); (c) the coated baseplate and pattern assembly are placed together in a flask and backfilled with aluminum-impregnated epoxy; (d) after curing, the baseplate is removed from the finished mold; and (e) a second mold half suitable for injection molding is prepared.

Figure 20.21: A new faucet design, produced by casting from rapid-prototyped sand molds. *Source:* Courtesy of 3D Systems.

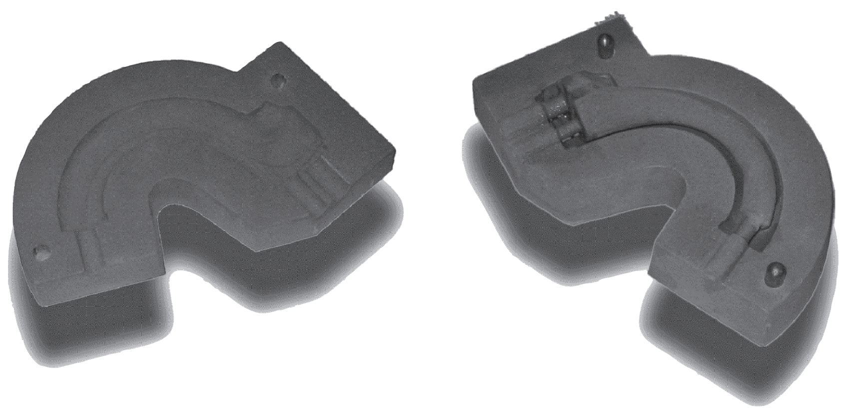


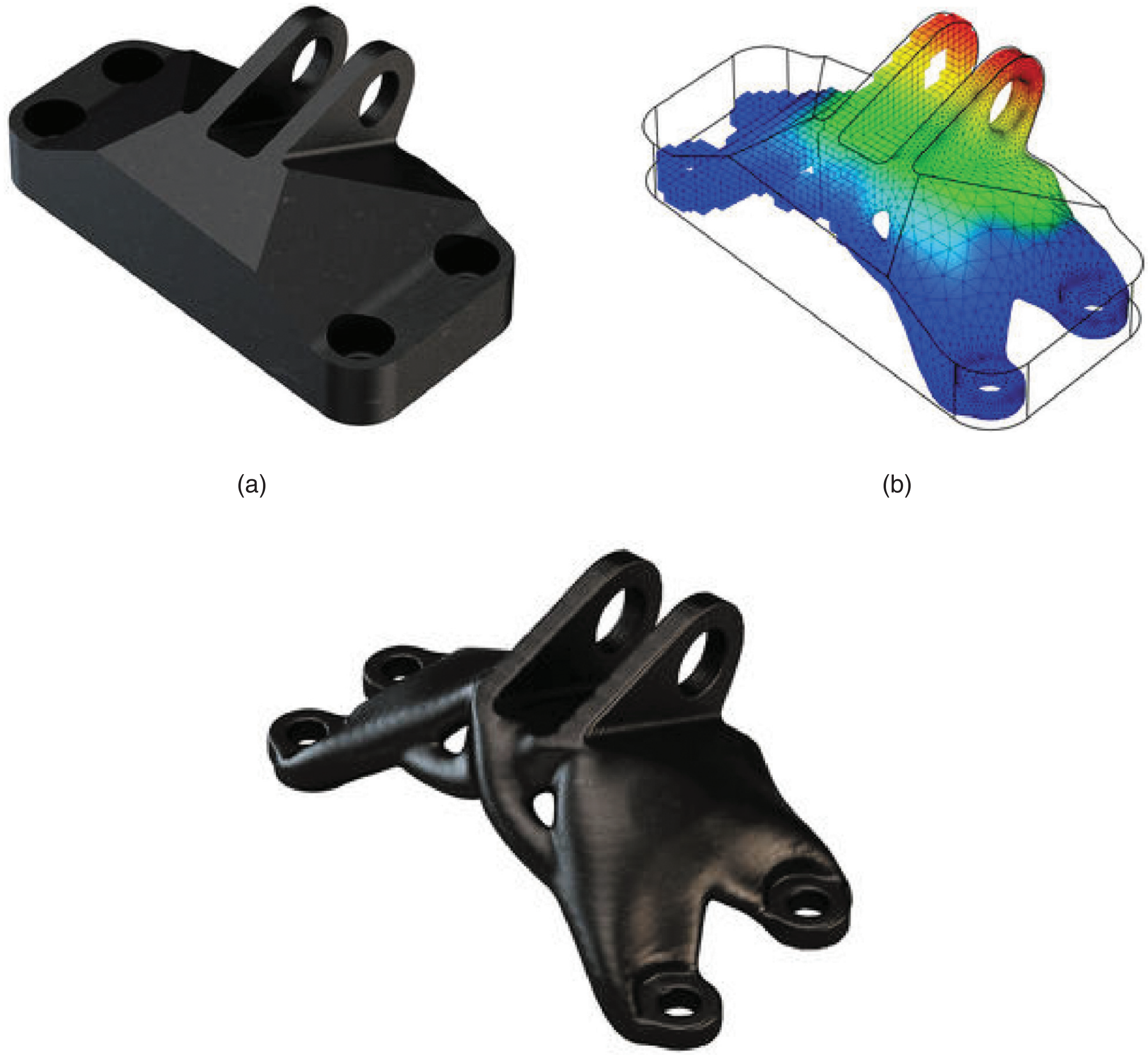
Figure 20.22: Sand molds produced through three-dimensional printing. *Source:* Courtesy of 3D Systems.

20.11 Design for Additive Manufacturing

Additive manufacturing is attractive because designers are able to easily produce complex geometries. Often, this has been expressed as the notion that *complexity is free*, which has led to the development of design optimization software. This approach has led to, for example, minimum-weight parts given shape constraints and the loads applied (see Fig. 20.23), as well as the production of parts with inherent aesthetic aims.

There are, however, limits to the shapes that can be produced by AM. Several design rules have been developed that are unique for additive manufacturing. Since machines are now available in a wide variety of capacities and capabilities, detailed design recommendations are manufacturer-specific. The following considerations are generic and are considered to be good design practice:

1. Additive manufacturing processes tend to warp the part, because of thermal stresses and shrinkage encountered during production. In general, the design guidelines for plastic parts, given in Section 19.15, are also applicable to parts produced through additive manufacturing.



(a) (b) (c)

Figure 20.23: Topology optimization to reduce the weight of a bracket. (a) Original bracket design; (b) predicted minimum-weight bracket from topology optimization software; and (c) final bracket, representing a 70% weight loss from the original design, as produced through selective laser sintering. *Source:* 3D Systems.

2. The dimensional tolerance standard used (see Section 35.8) should involve symmetric tolerances in order to be applied easily to additive manufacturing.

3. The tolerances within a plane can be much higher than those outside of a plane. Therefore, the part should be oriented to place the critical dimension in the plane of a build, not in its thickness direction.

4. Dimensional tolerances and surface finish depend on the particular machine, the material, and part size and its orientation. In stereolithography, tolerances of ± 0.05 to 0.1 mm are achievable, or ± 0.001 mm/mm for well-designed parts that do not warp excessively. Typical selective laser sintering of

polymers yields tolerances of ± 0.4 mm, or 0.1 mm/mm, whichever is greater. For metal selective laser systems, tolerances of 0.05–0.125 mm are generally achievable, with roughnesses in the range of 5–40 μm. For better tolerances, a *machining allowance* of 0.5–1 mm should be provided for post-processing.

5. Steps can be noticeable in an inclined plane; generally, the use of flat planes or planes inclined at not less than 20° are producible without noticeable steps.
6. The same considerations as stated for powder injection molding are valid for binder jetting (see Section 17.6).
7. In selective laser sintering of polymers, it is recommended to have a clearance of 0.3–0.5 mm within the plane for surfaces that are not joined together; up to 0.6 mm is required in the build direction.
8. The thinnest wall that can be produced depends on the material and the aspect ratio; common ranges are 0.5–1.5 mm for polymers in selective laser sintering. In fused deposition modeling, it is generally recommended that a wall be at least four times wider than the thickness of the layer.
9. Recognizing that the powder in the build chamber may not be reusable, as well as to maximize production, it is beneficial to fill a build space with as many parts as possible, and nestable (Section 16.14) parts be used when possible.
10. To reduce costs, the height in the build direction should be low, and stackable parts should be used to increase the amount of powder that is fused in a build chamber.
11. Consideration must be given to the removal of the uncured photopolymer or powder when the parts being made are hollow.
12. Large parts are especially susceptible to warpage; it may be a good strategy to produce a part in components that can be assembled after printing, or else design parts to use as little mass as possible.
13. Plan the part to allow for powder or liquid photopolymer removal when appropriate.
14. Build time depends on the volume of the material that is to be fused in a process. It is therefore beneficial to model an object with solid surfaces, but supported by porous structures or struts, instead of a solid bulk. This approach produces designs that can be optimized to minimize weight by carefully designing the supporting structure.
15. *Complexity is free*. That is, there is no need to restrict designs to geometries that are easy to manufacture for casting, forging or machining operations. Corner radii, draft angles, accommodations for parting lines, etc., do not need to be included in AM part design. With binder jetting, color can be incorporated into designs easily.

20.12 Additive Manufacturing Economics

As in all processes, design and manufacturing decisions are ultimately based on performance and cost, including the costs of equipment, tooling, and production. The final selection of a process or processes also depends greatly on production volume. High costs of equipment and tooling in plastics processing can be acceptable only if the production run is large, as is also the case in casting and forging. However, using additive manufacturing operations makes these processes economical for limited production runs by applying rapid tooling approaches (see Section 20.10), though the tools and molds have limited life.

Additive manufacturing operations are suitable for prototypes and limited production runs, but they require expensive consumables, and thus are unsuitable for moderate to high production runs. This situation is complicated by the fact that some processes (such as selective laser sintering and electron beam melting) may require the unfused powder in the build chamber to be discarded. Thus, if only 10% of the build volume is reused, the material cost for the part is ten times the nominal material cost. This is a significant concern; titanium (Ti-6Al-4V), for example, costs over \$400/kg for the raw powder.

The cost of a part produced by additive manufacturing can be generalized as

$$C_p = C_m + C_s + C_t + C_f, \tag{20.1}$$

where C_m is the material cost, C_s is the setup cost, C_t is the cost of machinery or tooling per part, and C_f is the cost of finishing operations. The material cost is high compared to conventionally produced polymers (such as in injection molding) or metals (such as extrusions). However, C_t in conventional processing is generally much higher, since no tools are required in additive manufacturing. Finishing operations may or may not be necessary, so that C_f may often be ignored.

When deciding if additive manufacturing is suitable for production, the cost compared to the conventional alternative has to be justified. Consider the case where a part is being considered for either injection molding or selective laser sintering. For low production runs, the high cost of tooling associated with injection molding dominates the part cost. However, for mass production, the tooling cost is amortized over many parts. The higher cost of the materials in additive manufacturing makes mass production a less economical option.

Case Study 20.4 Implications of Powder Reuse

Powder for additive manufacturing processes can be expensive; for example, \$400 per kg of titanium powder is not unusual, and \$100 per kg for high quality polymers is common. This high cost is especially important if a powder-bed process is used, as the volume fused in the build chamber may be as little as 10% of the total volume.

A common concern is whether or not unfused powder can be reused, that is, taken from the build chamber and then placed into the feed chamber. Often, the unfused powder (or *cake*) is loosely adhering and has to be sieved or otherwise treated to break up clumps of powder. If powder is not reused, then the cost embedded into the material can be several times the powder cost, thus making AM uneconomical for almost all commercial applications.

There are several strategies that can be applied in powder reuse:

1. The unfused powder can be taken as is from the build chamber, then sieved, examined, and placed in the feed chamber. Evaluations are generally associated with powder size distributions and the ability of the powder to flow or spread itself into a continuous and smooth layer.
2. The powder from the build chamber can be *blended* with virgin powder, often in a 1:1 mixture.

Some of the concerns associated with powder reuse are:

1. Additive manufacturing takes place under a controlled atmosphere, generally argon or nitrogen, in order to prevent powder oxidation and also to control any fire or explosion hazards associated with powders. However, when the powder is removed from the build chamber and is sieved, it is exposed to air, and therefore has the potential of oxidation.
2. The additive manufacturing process is rather complex. Videos of selective laser sintering have shown that particles that have been exposed to laser energy jump off the powder layer, the melt pool is highly turbulent, and there is a contraction as the powder melts and then solidifies.
3. The size distribution of powders (see also Section 17.2.2) can change over time.
4. Careful examination of powder size distributions have found subtle changes. When the feed chamber piston moves upwards, and the wiper or cylinder moves across the build chamber to create a fresh powder layer, there is always a slight surplus of powder in order to ensure that the build chamber layer is fully developed. The larger particles are pushed by the wiper or cylinder into an

overflow trough, leaving the smaller particles in the build chamber layer. These smaller particles are consumed during additive manufacturing, but the overfeed trough is blended with the powder reintroduced into the feed chamber. The result is that the mean particle size tends to increase slightly as the powder is reused.

5. There is a major concern that the high temperatures associated with the melt pool could cause particles near the melt pool to fuse. However, sieving eliminates such particles from being introduced into the build chamber.

6. Selective laser sintering and electron beam melting involve preheating the build chamber, in order to have a more robust process and to reduce the power required in the laser. With selective laser sintering, this preheat is much lower than with electron beam melting; still, there is a concern that this preheat can alter the microstructure or the chemistry of the powder.

Figure 20.24 shows the effect of reuse on ultimate tensile strength. Note that there is no noticeable reduction in mechanical properties associated with the first powder reuse for any of the metals considered. There is a drop in strength when a nylon powder is reused four times, but there is no further reduction through eight reuses. This is a main justification for the practice of blending virgin nylon powder with reclaimed powder from the build chamber.

It should also be noted that the observations regarding Fig. 20.24 may not hold for all materials. It has been suggested that alloys that are especially sensitive to oxygen and water vapor (such as magnesium alloys) may undergo a degradation in mechanical properties associated with reuse, because of their exposure to humidity during reclamation and sieving.

Regardless, the reuse of powders is now seen to be a plausible strategy for cost reduction in AM, and that it could greatly accelerate additive manufacturing application to actual production.

Source: Courtesy of the National Center for Defense Manufacturing and Machining, America Makes, and the Air Force Research Laboratory.

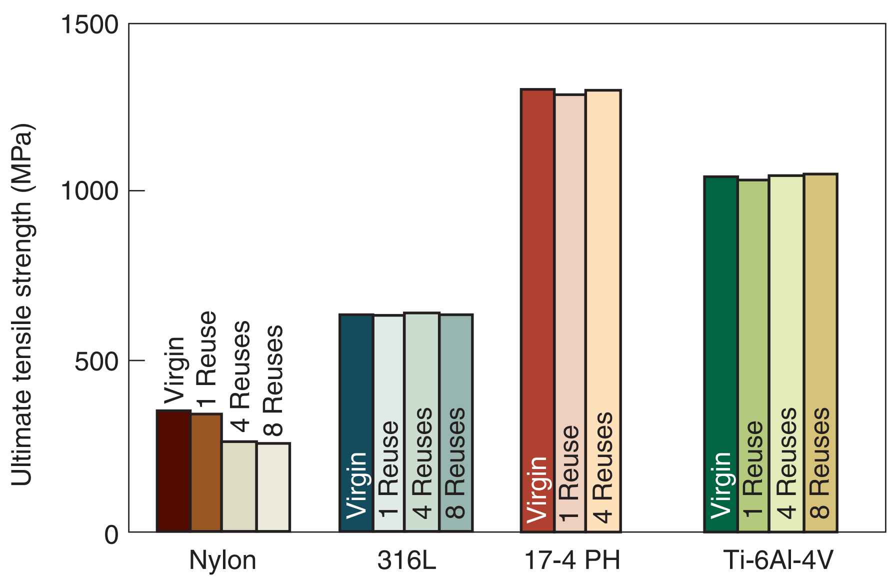


Figure 20.24: Effect of powder reuse on mean ultimate tensile strength for nylon, 316L stainless steel, 17-4 precipitation hardening stainless steel and titanium alloy Ti-6Al-4V.

Summary

- Additive manufacturing techniques have made possible much faster product development times, and they are having a major effect on other manufacturing processes. When appropriate materials are used, additive manufacturing machinery can produce blanks for investment casting or similar processes, so that metallic parts can now be obtained quickly and inexpensively, even for lot sizes as small as one part. Such technologies also can be applied to producing molds for operations (such as injection molding, sand and shell mold casting, and forging), thereby significantly reducing the lead time between design and manufacture.
- Additive manufacturing continues to grow into a valuable new manufacturing discipline. It is a useful technique for identifying and correcting design errors. Several techniques have been developed for producing parts through AM.
- Fused-deposition modeling consists of a computer-controlled extruder, through which a polymer filament is deposited to produce a part slice by slice.
- Stereolithography involves a computer-controlled laser-focusing system, that cures a liquid thermosetting polymer containing a photosensitive curing agent.
- Multijet and PolyJet modeling use mechanisms similar to ink-jet printer heads to eject photopolymers to directly build prototypes.
- Laminated-object manufacturing uses a laser beam or vinyl cutter to first cut the slices on paper or plastic sheets (laminations); then it applies an adhesive layer, if necessary, and finally stacks the sheets to produce the part.
- Three-dimensional printing uses an ink-jet mechanism to deposit liquid droplets of the liquid binder onto polymer, metal, or ceramic powders. The related process of material jetting directly deposits the build material. Using multiple printheads, three-dimensional printing can also produce full-color prototypes.
- Selective laser sintering uses a high-powered laser beam to sinter powders or coatings on the powders in a desired pattern. Selective laser sintering has been applied to polymers, sand, ceramics, and metals.
- Electron-beam melting uses the power of an electron beam to melt powders and form fully-dense functional parts.

Key Terms

ACES
Additive manufacturing
Big area additive manufacturing
Binder jet printing
Bioprinting
CLIP
Continuous liquid interphase production
Contour crafting
Desktop machines
Direct AIM
Direct manufacturing
Distributed manufacturing
Direct prototyping
Electron-beam melting
Friction stir modeling
Fused-deposition modeling
Hybrid approaches
JetFusion
Keltool
Laminated-object manufacturing

Laser-engineered net shaping
Mask projection stereolithography
Mass customization
Material jetting
Multijet modeling
Photopolymer
PolyJet
Powder bed
Prototype
Rapid tooling
RTV molding/urethane casting
Selective laser sintering
Sprayed metal tooling
Stereolithography
Subtractive processes
Three-dimensional printing
Virtual prototyping
Wire and arc additive manufacturing

Bibliography

Chua, C.K., and Leong, K.F., **Additive manufacturing: Principles and Applications**, 5th ed., World Scientific Publishing Co., 2017.
Gebhardt, A., **Understanding Rapid Prototyping, Rapid Tooling, Rapid Manufacturing**, Hanser, 2011.
Gibson, I., Rosen, D.W., and Stucker, B., **Additive Manufacturing Technologies: 3D Printing, Rapid Prototyping, and Direct Digital Manufacturing**, 2nd ed., Springer 2015.
Kamrani, K., and Nasr, E.A. (eds.), **Rapid Prototyping: Theory and Practice**, Springer, 2005.
Liou, F.W., **Rapid Prototyping and Engineering Applications: A Toolbox of Prototype Development**, CRC Press, 2007.
Milewski, J.O., **Additive Manufacturing of Metals**, Springer, 2017.
Noorani, R.I., **Rapid Prototyping: Principles and Applications**, Wiley, 2006.
Redwood, B., Schöffer, F., and Garrett, B., **The 3D Printing Handbook,** 3D Hubs, 2017.
Wang, W., Stoll, H., and Conley, J.G., **Rapid Tooling Guidelines for Sand Casting**, Springer, 2010.

Review Questions

20.1. What is the basic difference between additive manufacturing and rapid prototyping?

20.2. What is stereolithography?

20.3. What is virtual prototyping, and how does it differ from additive methods?

20.4. What is fused-deposition modeling?

20.5. Explain what is meant by rapid tooling.

20.6. Why are photopolymers essential for stereolithography?

20.7. Explain what each of the following means: (a) 3DP, (b) LOM, (c) STL, (d) SLS, (e) FDM, and (f) LENS.

20.8. What starting materials can be used in fused-deposition modeling and in three-dimensional printing?

20.9. What are the cleaning and finishing operations in additive manufacturing processes? Why are they necessary?

20.10. Which additive manufacturing technologies do not require a laser?

20.11. What are the advantages of electron beam melting?

20.12. What is the Keltool process?

20.13. What is CLIP?

20.14. What is unique about BAAM?

20.15. Why are supports needed with some parts?

20.16. Which materials can be processed with binder jet printing?

20.17. Which additive manufacturing operations can produce transparent workpieces?

20.18. Which additive manufacturing operations can produce multi-colored workpieces?

Qualitative Problems

20.19. How can a mold for sand casting be produced using additive manufacturing techniques? Explain.

20.20. Examine a ceramic coffee cup and determine in which orientation you would choose to produce the part if you were using (a) fused-deposition manufacturing or (b) laminated-object manufacturing.

20.21. How would you rapidly manufacture tooling for injection molding? Explain any difficulties that may be encountered.

20.22. Explain the significance of rapid tooling in manufacturing.

20.23. List the processes described in this chapter that are best suited for the production of ceramic parts. Explain.

20.24. Few parts in commercial products today are directly manufactured through additive manufacturing operations. Explain.

20.25. Can rapid-prototyped parts be made of paper? Explain.

20.26. Careful analysis of a rapid-prototyped part indicates that it is made up of layers with a distinct filament outline visible on each layer. Is the material a thermoset or a thermoplastic? Explain.

20.27. Why are the metal parts in three-dimensional printing often infiltrated by another metal?

20.28. Make a list of the advantages and limitations of each of the additive manufacturing operations described in this chapter.

20.29. In making a prototype of a toy automobile, list the post-additive manufacturing finishing operations that you think would be necessary. Explain.

20.30. List approaches for quickly manufacturing tooling for injection molding.

20.31. What are the similarities and differences between stereolithography and CLIP?

20.32. List the additive manufacturing approaches that are suitable for metals.

20.33. Explain why part orientation in rapid prototyping is important.

20.34. Do you expect that materials produced from additive manufacturing will be isotropic? Explain.

Quantitative Problems

20.35. Using an approximate cost of \$1500 per liter for the liquid polymer, estimate the material cost of a rapid-prototyped rendering of a pen.

20.36. The extruder head in a fused-deposition modeling setup has a diameter of 1.27 mm and produces layers that are 0.28 mm thick. If the extruder head and polymer extrudate velocities are both 45 mm/s, estimate the production time for the generation of a 40-mm solid cube. Assume that there is a six-second delay between layers as the extruder head is moved over a wire brush for cleaning.

20.37. Using the data for Problem 20.36 and assuming that the porosity for the support material is 50%, calculate the production rate for making a 120-mm high cup with an outside diameter of 100 mm and a wall thickness of 5 mm. Consider the cases (a) with the closed end up and (b) with the closed end down.

20.38. Inspect Table 20.2 and compare the numerical values given with those for metals and other materials, as can be found in Part I of this text. Comment on your observations.

Synthesis, Design, and Projects

20.39. Additive manufacturing machines represent a large capital investment; consequently, many companies cannot justify the purchase of their own system. Thus, service companies that produce parts based on their customers' drawings have become common. Conduct an informal survey of such service companies, identify the classes of additive manufacturing machines that they use, and determine the percentage use of each class.

20.40. One of the major advantages of stereolithography is that it can use transparent polymers, so that internal details of parts can readily be discerned. List and describe several parts in which this feature is valuable.

20.41. A manufacturing technique is being proposed that uses a variation of fused-deposition modeling in which there are two polymer filaments that are melted and mixed prior to being extruded to make the part. What advantages does this method have?

20.42. Identify the additive manufacturing processes described in this chapter that can be performed with materials available in your home or that you can purchase easily at low cost. Explain how you would go about it. Consider materials such as thin plywood, thick paper, glue, and butter, as well as the use of various tools and energy sources.

20.43. Design a machine that uses additive manufacturing technologies to produce ice sculptures. Describe its basic features, commenting on the effect of size and shape complexity on your design.

20.44. Because of relief of residual stresses during curing, long unsupported overhangs in parts made by stereolithography tend to curl. Suggest methods of controlling or eliminating this problem.

20.45. Describe methods that would allow the use of reinforced polymers to be used in additive manufacturing.

20.46. Conduct an Internet and literature study and write a two-page paper on developments of producing artificial organs through additive manufacturing related processes.

20.47. A current topic of research involves producing parts from additive manufacturing operations and then using them in experimental stress analysis, in order to infer the strength of final parts produced by means of conventional manufacturing operations. List your concerns with this approach, and outline means of addressing these concerns.

20.48. Outline the approach you would use to produce prototypes of metal gears from plastic. Assume the gears are 100 mm in diameter, 25 mm thick, and have 25 teeth. Explain how your preferred method of production would change if you needed to produce (a) one gear; (b) 100 per month; (c) 100 per day; (d) 100 per hour.

20.49. There is a great desire to increase the speed of additive manufacturing approaches. List three strategies for increasing the speed of a process, along with the advantages and disadvantages of each method. Write a one-page paper on the approach you think is best.

PART IV
Machining Processes and Machine Tools

Parts made by the casting, forming, and shaping processes described in Parts II and III often require further operations before they are ready for use. Consider, for example, the following features and whether they could be produced by the processes described thus far:

- Smooth and shiny surfaces, such as the bearing surfaces of the crankshaft shown in Fig. IV.1.
- Small-diameter and deep holes in a part, such as the injector nozzle shown in Fig. IV.2.
- Parts with sharp features, a threaded section, or specified close dimensional tolerances, such as the part shown in Fig. IV.3.
- A threaded hole or holes on different surfaces of a part, for assembly with other components.
- Complex geometries, often in hard or high-performance materials that cannot be easily or economically produced and in the quantities desired through the processes described earlier in the book (see Fig. 25.1).
- Special surface finish and texture for functional purposes or for appearance.

It soon will become clear that none of the processes described in the preceding chapters is capable of producing the specific characteristics outlined above, thus the parts will require further processing, generally referred to as *secondary* or *finishing operations*. **Machining** is a general term describing a group of processes that consist of the **removal** of material and **modification** of a workpiece surfaces after it has been made. The very wide variety of shapes produced by machining can be seen in an automobile, as shown in Fig. IV.4.

In reviewing the contents of Parts II and III of this text, it will be recalled that some parts may indeed be produced to final shape (net shape) and in large quantities. However, machining processes may be preferable or even necessary for the following reasons:

1. Closer **dimensional accuracy** may be required than can be achieved by metalworking or casting processes alone. For example, the bearing surfaces in a crankshaft cannot be produced with good dimensional accuracy and surface finish through forging or sand casting alone.
2. Parts may require external and/or internal **geometric features**, such as sharp corners and internal threads, that cannot be produced by other processes.

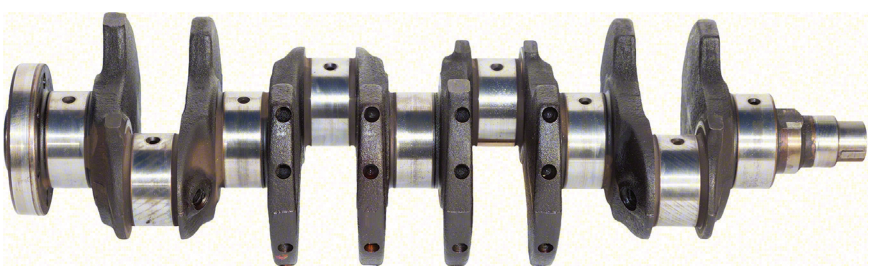


Figure IV.1: A forged crankshaft, highlighting the smooth and shiny machined and ground bearing surfaces. The shiny bearing surfaces cannot be made to their final dimensions and surface finish by any of the processes described in previous chapters. *Source:* Shutterstock/AleksandrN

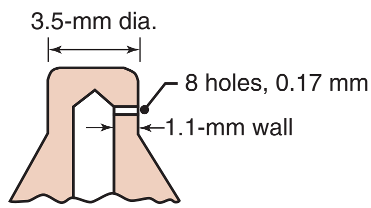


Figure IV.2: Cross section of a fuel-injection nozzle, showing a small hole made by the electrical-discharge machining process (Section 27.5). The material is heat-treated steel.

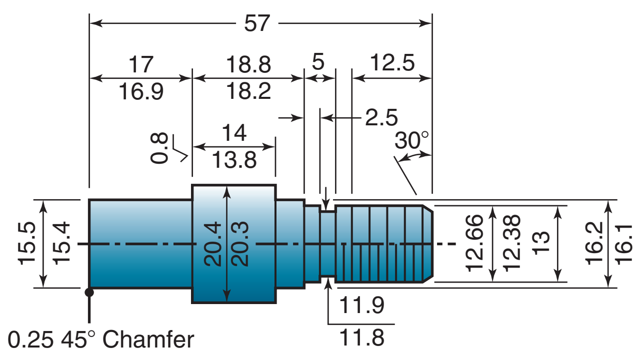


Figure IV.3: A machined and threaded part, showing various dimensions and tolerances; all dimensions are in mm. Note that some tolerances are only a few tenths of an mm.

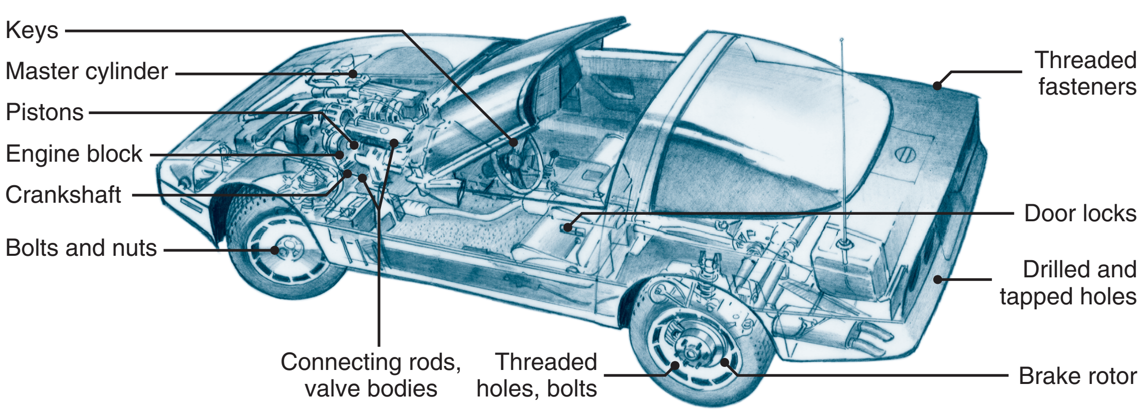


Figure IV.4: Typical parts on an automobile that require machining operations to impart desirable shapes, surface characteristics, dimensions, and tolerances.

3. Some parts are heat treated for improved hardness and wear resistance. However, because heat-treated parts may undergo distortion and surface discoloration, they may require additional **finishing operations**.

4. Special **surface characteristics** or **textures** may be required that cannot be produced by other means. As an example, copper mirrors with very high reflectivity are typically made by machining with a diamond cutting tool.

5. Some parts may be more **economical** to machine than to make by other processes, particularly if the number of parts required is relatively small. Recall that metalworking processes typically require expensive dies and tooling; the cost of these can only be justified if the production volume is sufficiently high.

In spite of their advantages, material-machining processes have certain limitations:

- They *waste material*, even though the amount may be relatively small.
- They generally require *more energy* than do forming and shaping operations.
- They can have *adverse effects* on the surface quality and properties of the product.

As outlined in Fig. I.6e in the General Introduction, machining consists of several major types of material-removal processes:

- **Cutting**, typically involving single-point or multipoint cutting tools, each with a clearly defined shape (Chapters 23 through 25).
- **Abrasive processes**, such as grinding and various related operations (Chapter 26).
- **Advanced machining processes**, typically utilizing electrical, chemical, laser, thermal, and hydrodynamic methods (Chapter 27).

The machines on which these operations are carried out are called **machine tools**. As can be noted in Table I.2 in the General Introduction, the first primitive tools, dating back several millennia, were made for the main purpose of chipping away and cutting wood, stone, vegetation, and livestock. It was not until the 1500s that developments began on making products by machining operations, particularly with the introduction of the lathe. Compared to the rather simple machinery and tools employed, a wide variety of computer-controlled machine tools and advanced techniques are now available, capable of making large parts as well as functional parts as small as tiny insects and with cross sections much smaller than a human hair.

As in all manufacturing operations, it is essential to view machining operations as a **system**, consisting of the (a) workpiece, (b) cutting tool, and (c) machine tool.

In the next seven chapters, the basic mechanics of chip formation in machining are described. These include tool forces, power requirements, temperature, tool wear, surface finish, integrity of the part machined, cutting tools, and cutting fluids. Specific machining processes are then described, including their capabilities, limitations, and typical applications, and important machine-tool characteristics for such basic operations as turning, milling, boring, drilling, and tapping.

The features of **machining centers**, which are versatile machine tools controlled by computers and capable of efficiently performing a variety of operations, are then presented. The next group of processes described are those in which the removal of material is carried out by **abrasive processes** and related operations. For technical and economic reasons, some parts cannot be machined satisfactorily by cutting or abrasive processes only. Since the 1940s, important developments have taken place in **advanced machining processes**, including chemical, electrochemical, electrical-discharge, laser-beam, electron-beam, abrasive-jet, and hydrodynamic machining.

Chapter 21

Fundamentals of Machining

Examples:

- This chapter is an introduction to the fundamentals of machining processes and presents the basic concepts relevant to all machining operations.
- The chapter opens with a description of the mechanics of chip formation, including the model typically used for studying the basic cutting operations, which allows the calculation of force and power in machining.
- Temperature rise and its importance on the workpiece and cutting tool, and the mechanisms of tool wear are then discussed.
- The chapter concludes with a description of surface finish, integrity of the parts produced by machining, and the factors involved in the machinability of materials.

21.1 Introduction

Machining processes remove material from the surfaces of a workpiece by producing **chips**. Some of the more common cutting processes, illustrated in Fig. 21.1 (see also Fig. I.6e), are:

- **Turning**, in which the workpiece is rotated and a cutting tool removes a layer of material as the tool moves along its length, as shown in Fig. 21.1a.
- **Cutting off**, in which the tool moves radially inward and separates a piece (on the right in Fig. 21.1b) from the blank.
- **Slab milling**, in which a rotating cutting tool removes a layer of material from the surface of the workpiece (Fig. 21.1c).
- **End milling**, in which a rotating cutter travels to a certain depth in the workpiece, producing a cavity (Fig. 21.1d).

In the turning process, illustrated in greater detail in Fig. 21.2, the cutting tool is set at a certain *depth of cut* (mm), and travels to the left with a certain *cutting speed* as the workpiece rotates. The *feed*, or *feed rate*, is

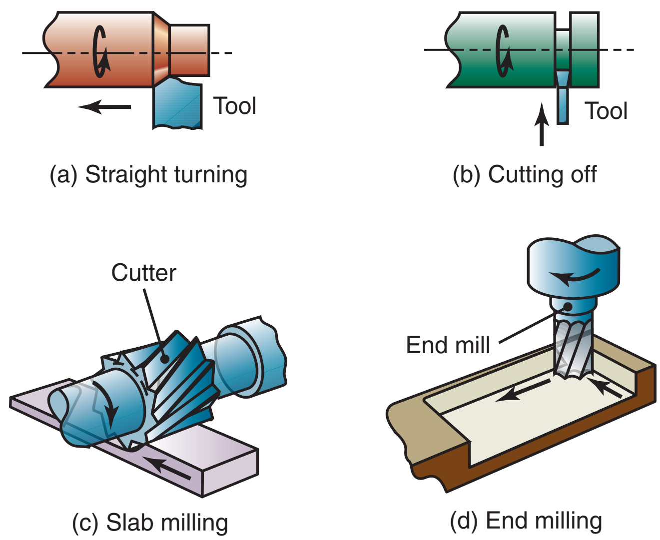


Figure 21.1: Some examples of common machining operations.

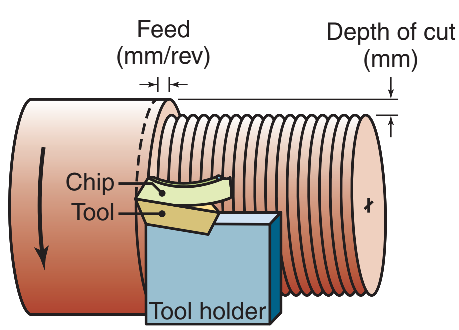


Figure 21.2: Schematic illustration of the turning operation, showing various features.

the distance the tool travels per revolution of the workpiece (mm/rev); this movement of the cutting tool produces a chip, which moves up the face of the tool.

In order to analyze this basic machining process in greater detail, a two-dimensional model of it is presented in Fig. 21.3a. In this *idealized* model, a cutting tool moves to the left along the workpiece at a constant velocity, V, and a depth of cut, t_o. Ahead of the tool, a chip is produced by plastic deformation, shearing the material continuously along the *shear plane*. This phenomenon can easily be demonstrated by slowly scraping the surface of a stick of butter lengthwise with a sharp knife, and observing how a chip is being produced. Chocolate shavings, used as decorations on cakes and pastries, are produced in a similar manner.

In comparing Figs. 21.2 and 21.3, note that the *feed* in turning is equivalent to t_o, and the *depth of cut* in turning is equivalent to the width of cut (the dimension perpendicular to the page). These dimensional

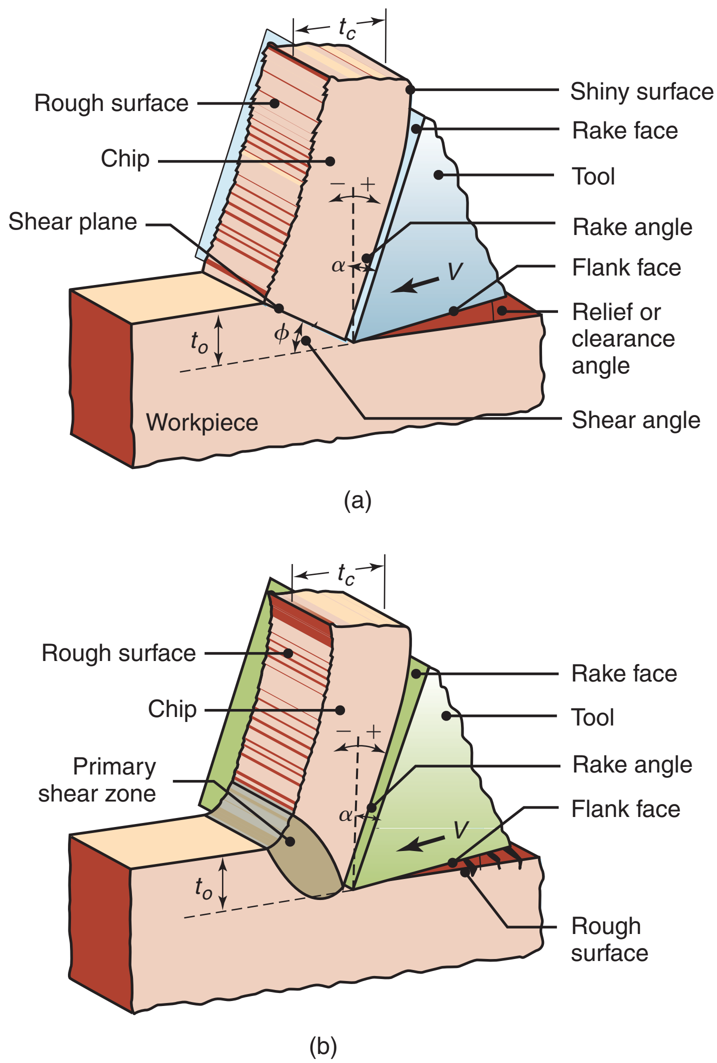


Figure 21.3: Schematic illustration of a two-dimensional cutting process, also called orthogonal cutting: (a) Orthogonal cutting with a well-defined shear plane, also known as the M.E. Merchant model. Note that the tool shape, the depth of cut, t_o, and the cutting speed, V, are all independent variables. (b) Orthogonal cutting without a well-defined shear plane.

relationships can be visualized by rotating Fig. 21.3 clockwise by 90°. With this brief introduction as a background, the cutting process will now be described in greater detail.

21.2 Mechanics of Cutting

The factors that influence the cutting operation are outlined in Table 21.1. In order to appreciate the contents of this table, consider the major *independent variables* in the basic cutting process: (a) tool material and coatings, if any; (b) tool shape, its surface finish and sharpness; (c) workpiece material and its processing history; (d) cutting speed, feed, and depth of cut; (e) cutting fluids, if any; (f) characteristics of the machine tool; and (g) the type of workholding device and fixturing.

Dependent variables in machining are those that are influenced by changes made in the independent variables listed above. They include: (a) type of chip produced, (b) force and energy dissipated during cutting, (c) temperature rise in the workpiece, the tool, and the chip, (d) tool wear and failure, and (e) surface finish and surface integrity of the workpiece.

The importance of establishing *quantitative relationships* among the independent and dependent variables in machining can best be appreciated by considering some typical questions to be posed: Which of the independent variables should be changed first and to what extent (a) if the surface finish of the workpiece being machined becomes unacceptable, (b) if the cutting tool wears rapidly and becomes dull, (c) if the workpiece becomes very hot, and (d) if the tool begins to vibrate and chatter.

In order to understand these phenomena and respond to the questions posed, consider that the mechanics of chip formation have been studied extensively since the early 1940s. Several models, with varying degrees of complexity, have been proposed to describe the basic cutting process. More advanced machining models are being developed, especially *computer simulation* of the mechanics of the basic machining process.

The simple model shown in Fig. 21.3a, and referred to as the M.E. Merchant model, developed in the early 1940s, is sufficient for the purposes of this introduction. This model is known as **orthogonal cutting**, because it is two-dimensional whereby the forces involved are perpendicular to each other. The cutting tool has a **rake angle**, α (positive as shown in the figure), and a **relief** or **clearance angle**.

Microscopic examination of chips produced in actual operations reveal that they are produced by *shearing* (as modeled in Fig. 21.4a), a phenomenon similar to the movement of cards in a deck that is being deformed (see also Fig. 1.6). Shearing takes place within a **shear zone** (usually along a well-defined plane referred to as the **shear plane**) and at an angle ϕ (called the **shear angle**). Below the shear plane, the workpiece remains undeformed; above it, the chip (which is already formed) moves up the rake face of the tool.

Table 21.1: Factors Influencing Machining Operations.

Parameter	Influence and interrelationship
Cutting speed, depth of cut, feed, cutting fluids	Forces, power, temperature rise, tool life, type of chip, surface finish and integrity
Tool angles	As above; influence on chip flow direction; resistance to tool wear and chipping
Continuous chip	Good surface finish; steady cutting forces; undesirable, especially in modern machine tools
Built-up edge chip	Poor surface finish and integrity; if thin and stable, edge can protect tool surfaces
Discontinuous chip	Desirable for ease of chip disposal; fluctuating cutting forces; can affect surface finish and cause vibration and chatter
Temperature rise	Influences tool life, particularly crater wear and dimensional accuracy of workpiece; may cause thermal damage to workpiece surface
Tool wear	Influences surface finish and integrity, dimensional accuracy, temperature rise, forces and power
Machinability	Related to tool life, surface finish, forces and power, and type of chip produced

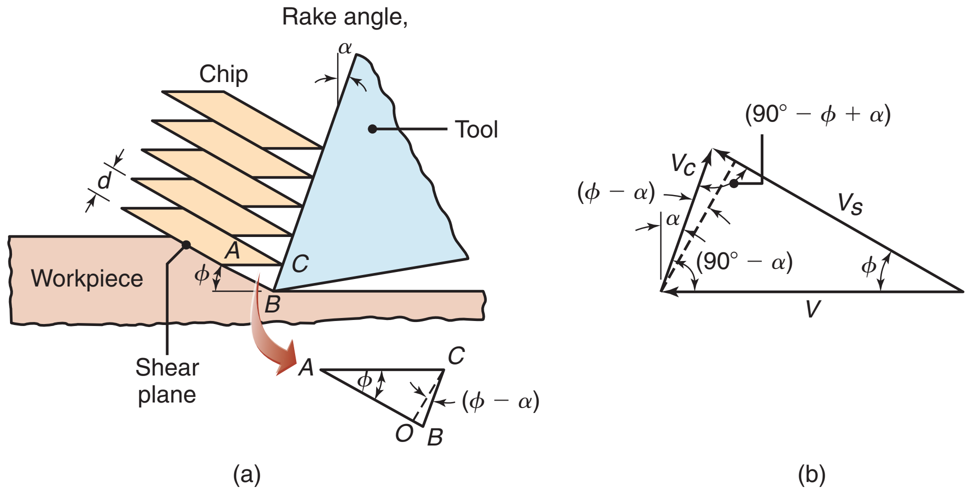


Figure 21.4: (a) Schematic illustration of the basic mechanism of chip formation by shearing. (b) Velocity diagram showing angular relationships among the three speeds in the cutting zone.

In this figure, the dimension d is highly exaggerated in order to show the mechanism involved in chip formation; this dimension has been found to be only on the order of 10^{-2} to 10^{-3} mm.

Some materials, notably cast irons machined at low speeds, do not undergo shearing along a well-defined plane, but instead within a *shear zone*, as shown in Fig. 21.3b. The shape and size of this zone is important in the machining operation, as will be described in Section 21.2.1.

Cutting Ratio. It can be seen from Fig. 21.3a that the chip thickness, t_c, can be determined from the depth of cut, t_o, the rake angle, α, and the shear angle, ϕ. The ratio of t_o/t_c is known as the **cutting ratio**, or chip-thickness ratio, r. It is related to the two angles by the following relationships:

$$\tan\phi = \frac{r\cos\alpha}{1 - r\sin\alpha} \tag{21.1}$$

and

$$r = \frac{t_o}{t_c} = \frac{\sin\phi}{\cos(\phi - \alpha)}. \tag{21.2}$$

Because the chip thickness is always greater than the depth of cut, the value of r is always less than unity. The reciprocal of r is known as the *chip-compression ratio* or *chip-compression factor*; it is a measure of how thick the chip has become as compared with the depth of cut. Thus, the chip-compression ratio always is greater than unity. As may be visualized by reviewing Fig. 21.3a, the depth of cut is also referred to as the *undeformed chip thickness*.

The cutting ratio is an important and useful parameter for evaluating cutting conditions. Since the undeformed chip thickness, t_o, is easily specified as a machine setting, and is therefore known, the cutting ratio can be calculated by measuring the chip thickness, using a micrometer. With the rake angle also known for a particular cutting operation (since it is a function of the tool and workpiece geometries), Eq. (21.1) allows calculation of the shear angle.

Although t_o is referred to as the *depth of cut*, note that in a machining process such as turning, shown in Fig. 21.2, this quantity is the *feed* or *feed rate*, expressed as the distance traveled per revolution of the workpiece. To visualize the situation, assume that the workpiece in Fig. 21.2 is a thin-walled tube, and that the width of cut is the same as the thickness of the tube. Then, by rotating Fig. 21.3 clockwise by 90°, the figure now becomes similar to the view in Fig. 21.2.

Shear Strain. Referring to Fig. 21.4a, it can be seen that the **shear strain**, γ, that the material undergoes can be expressed as

$$\gamma = \frac{AB}{OC} = \frac{AO}{OC} + \frac{OB}{OC},$$

or

$$\gamma = \cot\phi + \tan(\phi - \alpha). \tag{21.3}$$

Note that large shear strains are associated with (a) low shear angles and (b) with low or negative rake angles. Shear strains of 5 or higher have been observed in actual cutting operations. The material removed from the workpiece undergoes greater deformation during cutting than in forming and shaping processes, as is also seen in Table 2.4. Furthermore, deformation in machining generally takes place within a very narrow zone; in other words, the dimension $d = OC$ in Fig. 21.4a is very small. Thus, the rate at which shearing takes place in machining is high. The nature and size of the deformation zone is further described in Section 21.3.

The shear angle has a major significance in the mechanics of machining operations, as it influences force and power requirements, chip thickness, and temperature rise in machining. One of the earliest analyses was based on the assumption that the shear angle adjusts in order to minimize the cutting force, or that the shear plane is a plane of maximum shear stress. This analysis yields the expression

$$\phi = 45^\circ + \frac{\alpha}{2} - \frac{\beta}{2}, \tag{21.4}$$

where β is the **friction angle**, and is related to the *coefficient of friction*, μ, at the tool–chip interface by the expression $\mu = \tan\beta$.

Among several other shear-angle relationships that have been developed, another approximate but useful formula is

$$\phi = 45^\circ + \alpha - \beta. \tag{21.5}$$

The *coefficient of friction* in metal cutting has been found to generally range from about 0.5 to 2 (see also Section 33.4), indicating that the chip undergoes considerable frictional resistance as it moves up the rake face of the tool. Experiments have shown that μ varies considerably along the tool–chip interface, because of large variations in contact pressure and temperature. Consequently, μ is also called the *apparent mean coefficient of friction*.

Equation (21.4) indicates that (a) as the rake angle decreases or as the friction at the tool–chip interface increases, the shear angle decreases and the chip becomes thicker; (b) thicker chips indicate more energy dissipation, because the shear strain is higher, as can be noted from Eq. (21.2); and (c) because the work done during cutting is converted into heat, the temperature rise is also higher.

Velocities in the Cutting Zone. Note in Fig. 21.3 that since the chip thickness is greater than the depth of cut, the velocity of the chip V_c has to be lower than the cutting speed V. Because mass continuity has to be maintained,

$$Vt_o = V_c t_c \text{ or } V_c = Vr.$$

Hence,

$$V_c = \frac{V\sin\phi}{\cos(\phi - \alpha)}. \tag{21.6}$$

A *velocity diagram* also can be constructed, as shown in Fig. 21.4b. From trigonometric relationships,

$$\frac{V}{\cos(\phi - \alpha)} = \frac{V_s}{\cos\alpha} = \frac{V_c}{\sin\phi} \tag{21.7}$$

where V_s is the velocity at which shearing takes place in the shear plane. Note also that

$$r = \frac{t_o}{t_c} = \frac{V_c}{V}. \tag{21.8}$$

These relationships will be utilized later in Section 21.3, describing power requirements in machining operations.

21.2.1 Types of Chips Produced in Metal Cutting

The types of metal chips commonly observed in practice and their photomicrographs are shown in Fig. 21.5. The four main types are:

- **Continuous**
- **Built-up edge**
- **Serrated** or **segmented**
- **Discontinuous**.

Note that a chip has two surfaces:

1. A surface that has been in contact with the rake face of the tool and has a shiny and burnished appearance, caused by sliding as the chip moves up the tool face.
2. A surface that is the original surface of the workpiece; it has a rough, jagged appearance (as can be seen on the chips in Figs. 21.3 and 21.5) caused by the shearing mechanism shown in Fig. 21.4a.

Continuous Chips. *Continuous chips* are generally formed with ductile materials, machined at high cutting speeds and/or at high rake angles (Fig. 21.5a). Deformation of the material takes place along a narrow shear zone, called the *primary shear zone*. Continuous chips may develop a *secondary shear zone* (Fig. 21.5b) because of high friction at the tool–chip interface; this zone becomes wider as friction increases.

Deformation in continuous chips also may take place along a wide primary shear zone with *curved boundaries* (see Fig. 21.3b), unlike that shown in Fig. 21.5a. Note that the lower boundary of the deformation zone in Fig. 21.3b projects *below* the machined surface, subjecting it to distortion, as depicted by the distorted vertical lines within the machined subsurface. This situation generally occurs in machining soft metals, at low speeds, and low rake angles. It usually results in a poor surface finish and surface residual stresses, which may be detrimental to the properties of a machined part in its service life.

Although they generally produce a good surface finish, continuous chips are not necessarily desirable as they tend to become tangled around the toolholder, the fixturing, and the workpiece. They also interfere with chip-disposal systems, described in Section 23.3.7. This situation can be alleviated using **chip breakers** (see below), as well as by changing processing parameters, such as cutting speed, feed, and depth of cut, or by using appropriate cutting fluids.

Built-up Edge Chips. A *built-up edge* (BUE) consists of layers of material from the workpiece that gradually are deposited on the tool tip, hence the term *built-up* (Fig. 21.5c). As it grows larger, a BUE becomes unstable, and eventually breaks apart. A portion of the BUE material is carried away by the tool side or *rake face* of the chip; the rest is deposited randomly on the workpiece surface. Note that, in effect, a built-up edge changes the geometry of the cutting edge and dulls it, as can be seen in Fig. 21.6a. The cycle of BUE formation and destruction is repeated continuously during the cutting operation.

Built-up edge is a major factor that adversely affects surface finish, as can be seen in Figs. 21.5c and 21.6b and c. On the other hand, a thin, stable BUE is generally regarded as desirable, because it reduces tool wear by protecting its rake face. Cold-worked metals have a lower tendency to form BUE than those in their annealed condition. Because of work hardening and deposition of successive layers of material, the BUE hardness is significantly higher than that of the workpiece (Fig. 21.6a).

The tendency for BUE formation can be reduced by one or more of the following means:

- Increase the cutting speed
- Decrease the depth of cut

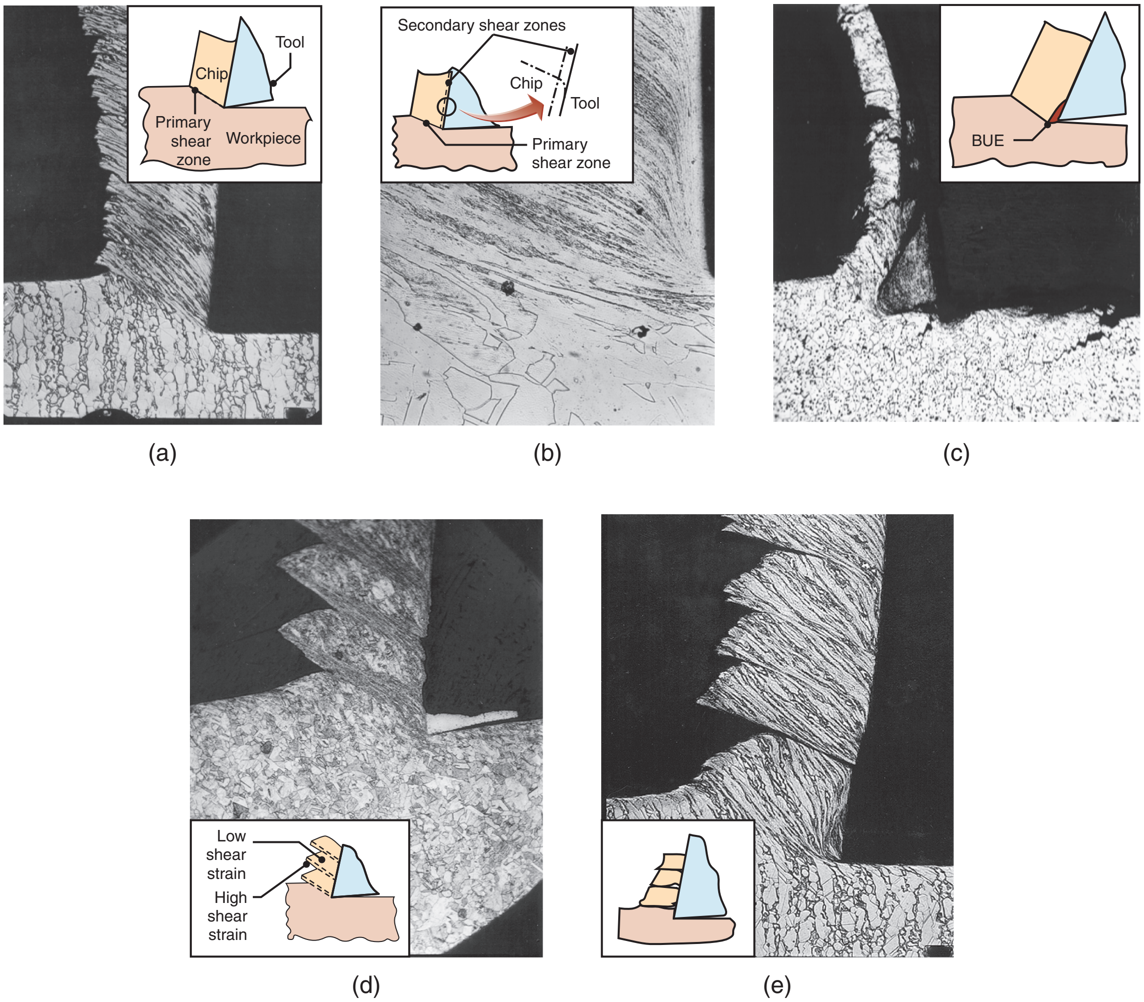


Figure 21.5: Basic types of chips produced in orthogonal metal cutting, their schematic representation, and photomicrographs of the cutting zone: (a) continuous chip, with narrow, straight, and primary shear zone; (b) continuous chip, with secondary shear zone at the chip–tool interface; (c) built-up edge; (d) segmented or nonhomogeneous chip; and (e) discontinuous chip. *Source:* After M.C. Shaw, P.K. Wright, and S. Kalpakjian.

- Increase the rake angle
- Use a sharper tool
- Use a cutting tool that has lower chemical affinity for the workpiece material
- Use an effective cutting fluid.

Serrated Chips. *Serrated chips*, also called *segmented* or *nonhomogeneous chips* (Fig. 21.5d), are semicontinuous chips with large zones of low shear strain and small zones of high shear strain (called *shear localization*). These chips have a sawtooth-like appearance (not be confused with the illustration in Fig. 21.4a, in which the dimension d is highly exaggerated). Metals that have low thermal conductivity and strength that decreases sharply with temperature (called *thermal softening*) exhibit this behavior, and is most notably observed with titanium.

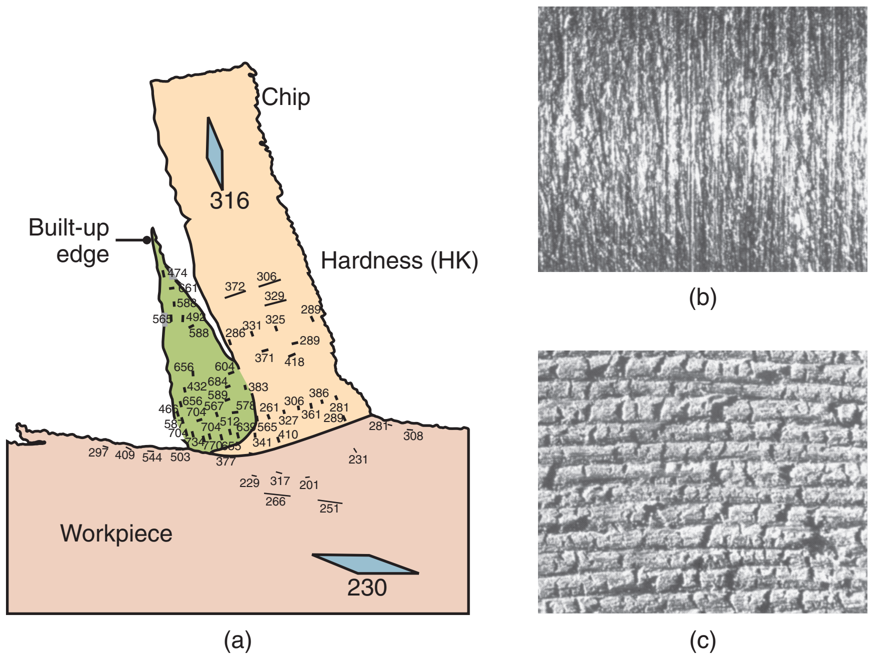


Figure 21.6: (a) Hardness distribution in a built-up edge in 3115 steel. Note that some regions within the built-up edge are as much as three times harder than the bulk metal being machined. (b) Surface finish produced in turning 5130 steel with a built-up edge. (c) Surface finish on 1018 steel in face milling. Magnifications: $15\times$. *Source:* (b) and (c) Courtesy of TechSolve, Inc.

Discontinuous Chips. *Discontinuous chips* consist of segments, either firmly or loosely attached to each other (Fig. 21.5e). Discontinuous chips generally develop under the following conditions:

- Brittle workpiece materials, because they do not have the capacity to undergo the high shear strains encountered in machining
- Workpiece materials containing hard inclusions and impurities, or have structures such as the graphite flakes in gray cast iron (see Fig. 4.11a)
- Very low or very high cutting speed, V
- Large depth of cut, d
- Tools with low rake angle, α
- Lack of an effective cutting fluid (Section 22.12)
- Low stiffness of the toolholder or the machine tool, thus allowing vibration and chatter to occur (Section 25.4).

Another factor in the formation of discontinuous chips is the magnitude of the compressive stresses on the shear plane. The maximum shear strain at fracture increases with increasing compressive stress.

Because of the discontinuous nature of chip formation, cutting forces continually vary during machining. Consequently, the stiffness or rigidity of the cutting-tool holder, the workholding devices, and the machine tool and its condition (see Chapters 23 through 25) are significant factors in machining with serrated or discontinuous chips. If not sufficiently rigid, the machine tool may begin to vibrate and chatter, as described in detail in Section 25.4. This condition, in turn, adversely affects the surface finish and dimensional accuracy of the machined part, and it may cause premature wear or damage to the cutting tool. Even the components of the machine tool may be damaged if the amplitude of the vibration is excessive.

Chip Curl. In all cutting operations performed on metals and nonmetallic materials, chips develop a curvature (*chip curl*) as they leave the workpiece surface (Fig. 21.5). Among the factors affecting chip curl are:

- The distribution of stresses in the primary and secondary shear zones
- Thermal effects in the cutting zone
- Work-hardening characteristics of the workpiece material
- The geometry of the cutting tool
- Process parameters
- Cutting fluids.

The first four items above are complex phenomena and beyond the scope of this text. As for the effects of process parameters: as the depth of cut decreases, the radius of curvature of the chip generally decreases (i.e., the chip becomes more curly). Also, cutting fluids can make chips become more curly, thus reducing the tool–chip contact area (see Fig. 21.7a) and concentrating the heat closer to the tip of the tool (Section 21.4). As a result, tool wear increases.

Chip Breakers. As stated above, continuous and long chips are undesirable in machining operations because they tend to become severely entangled, interfere with the machining operation, and can also become a potential safety hazard. The usual procedure employed to avoid such a situation is to break the chip intermittently with special features on cutting tools, called *chip-breakers*, as shown in Fig. 21.7.

The basic principles of a chip breaker on a tool's rake face is to bend and break the chip periodically. Cutting tools and inserts (see Fig. 22.2) now have built-in chip-breaker features of various designs (Fig. 21.7). Chips also can be broken by changing the tool geometry to control chip flow, as in the turning operations shown in Fig. 21.8. Experience indicates that the ideal chip size to be broken is in the shape of either the letter C or the number 9, and fits within a 25-mm square space.

Controlled Contact on Tools. Cutting tools can be designed such that the tool–chip contact length is deliberately reduced by recessing the rake face of the tool some distance away from its tip. The reduction in contact length then affects the chip-formation mechanics; primarily, it reduces the cutting forces and, thus, the energy and temperature in machining. Determining an optimum length is important, as too small a contact length would concentrate the heat at the tool tip, increasing tool wear.

Machining Nonmetallic Materials. The mechanics of cutting metals are generally applicable to polymers as well as metals. A variety of chips are encountered in cutting *thermoplastics* (Section 7.3), depending on the type of polymer and process parameters, such as depth of cut, tool geometry, and cutting speed. Because they are brittle, *thermosetting* plastics (Section 7.4) and ceramics (Chapter 8) generally produce discontinuous chips. The characteristics of other machined materials are described in Section 21.7.3.

21.2.2 Oblique Cutting

The majority of machining operations involve tool shapes that are three dimensional, whereby the cutting action is *oblique*. The basic difference between oblique and orthogonal cutting can be seen in Fig. 21.9a and c. In orthogonal cutting, the chip slides directly up the face of the tool and it becomes a spiral, whereas in *oblique cutting*, the chip becomes helical and leaves the workpiece surface at an angle i, called the **inclination angle** (Fig. 21.9b). Note the lateral direction of chip movement in oblique cutting is similar to the action of a snowplow blade, whereby the snow is thrown *sideways* as the plow travels straight forward.

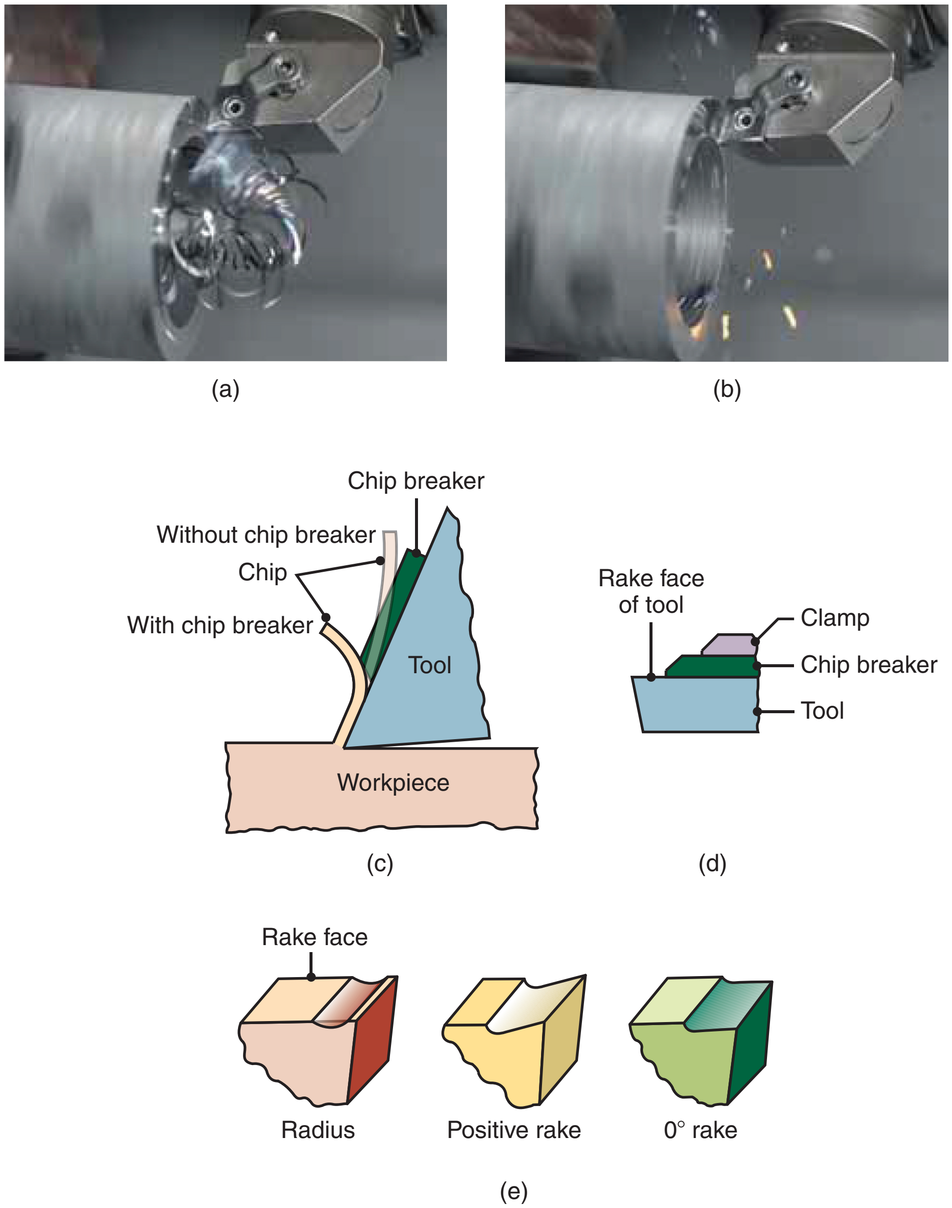


Figure 21.7: (a) Machining aluminum using an insert without a chip breaker; note the long chips that can interfere with the tool and present a safety hazard. (b) Machining aluminum with a chip breaker. (c) Schematic illustration of the action of a chip breaker; note that the chip breaker decreases the radius of curvature of the chip and eventually breaks it. (d) Chip breaker clamped on the rake face of a cutting tool. (e) Grooves in cutting tools acting as chip breakers; the majority of cutting tools are now inserts with built-in chip-breaker features. *Source:* (a) and (b) Courtesy of Kennametal, Inc.

Note in Fig. 21.9a that the chip moves up the rake face of the tool at an angle α_c (called the **chip flow angle**), measured in the plane of the tool face. Angle α_i is the **normal rake angle**, and is a basic geometric feature of the tool. It is the angle between line oz normal to the workpiece surface and line oa on the tool face.

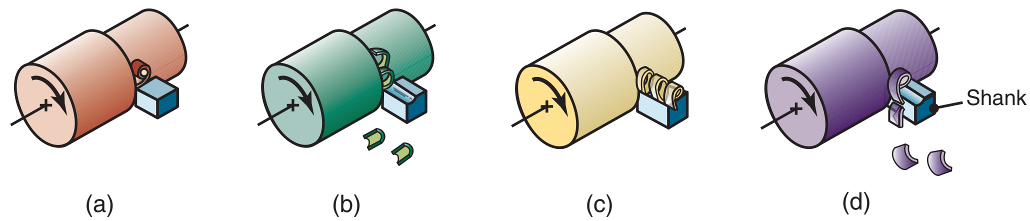


Figure 21.8: Chips produced in turning: (a) tightly curled chip; (b) chip hits workpiece and breaks; (c) continuous chip moving radially away from workpiece; and (d) chip hits tool shank and breaks off. *Source:* After G. Boothroyd.

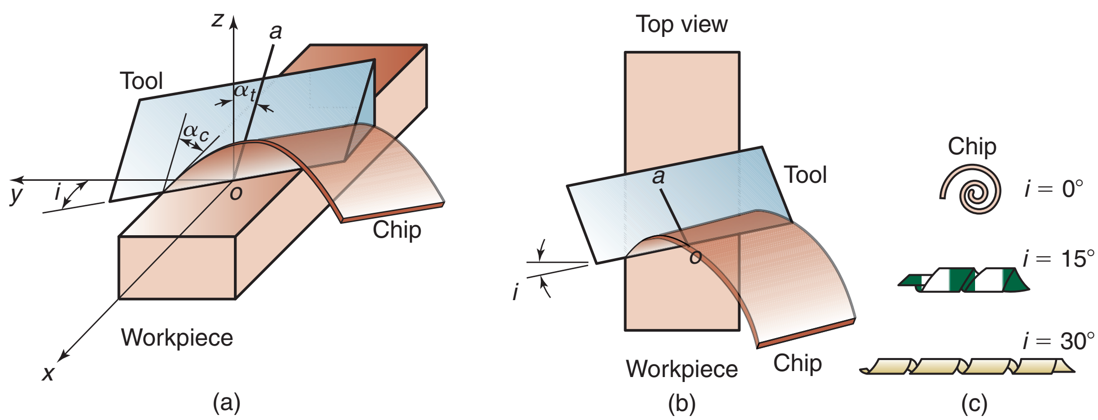


Figure 21.9: (a) Schematic illustration of cutting with an oblique tool; note the direction of chip movement. (b) Top view, showing the inclination angle, i. (c) Types of chips produced with tools at increasing inclination angles.

In oblique cutting, the workpiece material approaches the cutting tool at a velocity V and leaves the surface (as a chip) with a velocity V_c. The *effective rake angle*, α_e, is calculated in the plane of these two velocities. Assuming that the chip flow angle, α_c, is equal to the inclination angle (an assumption that has been verified experimentally), the effective rake angle, α_e, is

$$\alpha_e = \sin^{-1}\left(\sin^2 i + \cos^2 i \sin\alpha_n\right). \tag{21.9}$$

Since both i and α_n can be measured directly, the effective rake angle can now be calculated. Note that as i increases, the effective rake angle increases, the chip becomes thinner and longer and, as a consequence, the cutting force decreases. The influence of the inclination angle on chip shape is shown in Fig. 21.9c.

A typical single-point turning tool, used on a lathe, is shown in Fig. 21.10a; note the various angles involved, each of which has to be selected properly for efficient cutting. Although these angles have traditionally been produced by grinding (Chapter 26), the majority of cutting tools are now widely available as **inserts**, as shown in Fig. 21.10b and described in detail in Chapter 22. Various three-dimensional cutting tools, including those for drilling, tapping, milling, planing, shaping, broaching, sawing, and filing, are described in greater detail in Chapters 23 and 24.

Shaving and Skiving. Thin layers of material can be removed from straight or curved surfaces by a process similar to the use of a plane in shaving wood. *Shaving* is used particularly for improving the surface finish and dimensional accuracy of sheared sheet metals and punched holes, as shown in Fig. 16.9.

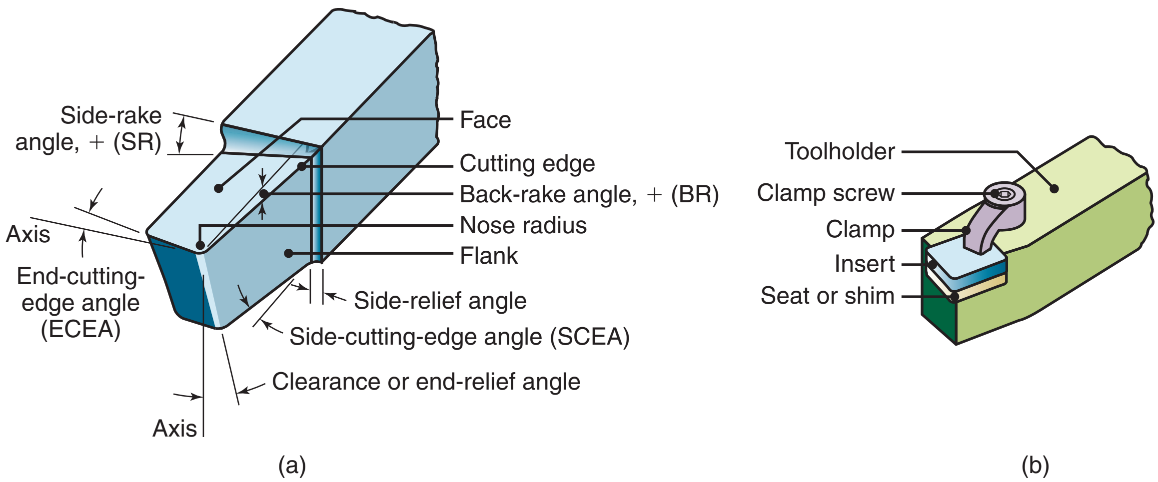


Figure 21.10: (a) Schematic illustration of a right-hand cutting tool. The various angles on these tools and their effects on machining are described in Section 23.2. Although these tools traditionally have been produced from solid-tool steel bars, they have been replaced largely with (b) inserts, typically made of carbides and other materials; they are available in a wide variety of shapes and sizes.

A common application of shaving is in finishing gears, using a cutter that has the shape of the gear tooth (see Section 24.7). Parts that are long or have complicated shapes are shaved by *skiving*, using a specially shaped cutting tool that moves tangentially across the length of the workpiece shaved.

21.3 Cutting Forces and Power

Studying the *cutting forces* and *power* involved in machining operations is important for the following reasons:

- Data on cutting forces is essential so that

  1. Machine tools can be designed to minimize distortion of their components, maintain the desired dimensional accuracy of the machined part, and help select appropriate toolholders and workholding devices.
  2. The workpiece, the workholding devices, and the fixtures are capable of withstanding these forces without excessive distortion.

- Power requirements must be known to enable the selection of a machine tool with sufficient capacity or to select process parameters that can be achieved by the machine selected.

The forces acting in orthogonal cutting are shown in Fig. 21.11a. The **cutting force**, F_c, acts in the direction of the cutting speed, V, and supplies the energy required for cutting. The ratio of the cutting force to the cross-sectional area being cut (i.e., the product of width of cut and depth of cut) is referred to as the *specific cutting force*.

The **thrust force**, F_t, acts in a direction normal to the cutting force. These two forces produce the **resultant force**, R, as can be seen from the force circle diagram shown in Fig. 21.11b. Note that the resultant force can be resolved into two components on the tool face: a **friction force**, F, along the tool–chip interface, and a **normal force**, N, perpendicular to it.

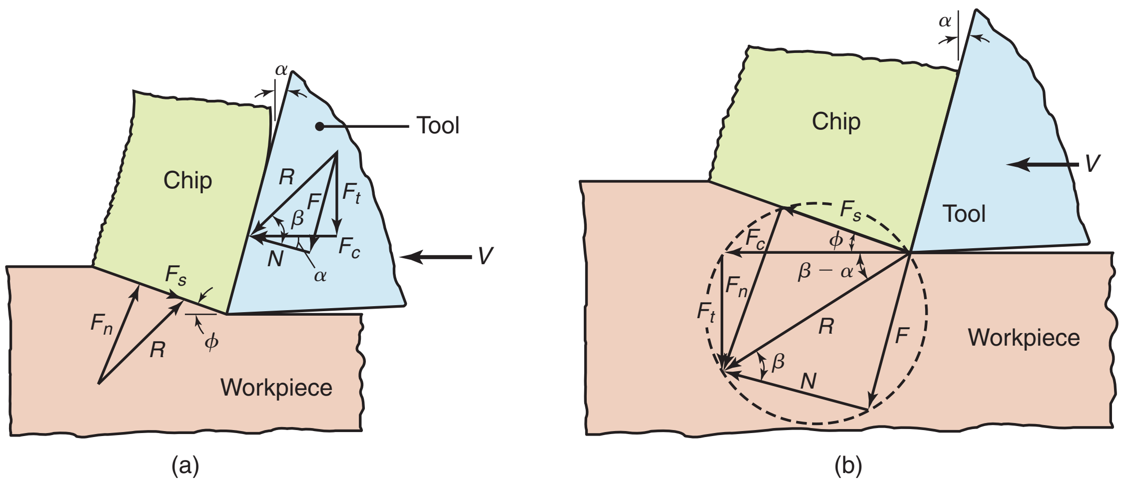


Figure 21.11: (a) Forces acting in the cutting zone during two-dimensional cutting. Note that the resultant force, R, must be colinear to balance the forces. (b) Force circle to determine various forces acting in the cutting zone.

It can also be shown that

$$F = R \sin \beta \tag{21.10}$$

and

$$N = R \cos \beta. \tag{21.11}$$

Note that the resultant force is balanced by an equal and opposite force along the shear plane, and is resolved into a **shear force**, F_s, and a **normal force**, F_n. These forces can be expressed as

$$F_s = F_c \cos \phi - F_t \sin \phi \tag{21.12}$$

and

$$F_n = F_c \sin \phi + F_t \cos \phi. \tag{21.13}$$

Because the area of the shear plane can be calculated by knowing the shear angle and the depth of cut, the shear and normal stresses in the shear plane can thus be determined.

The ratio of F to N is the **coefficient of friction** at the tool–chip interface, μ, and the angle β is the **friction angle** (as in Fig. 21.11). The magnitude of μ can be determined as

$$\mu = \frac{F}{N} = \frac{F_t + F_c \tan \alpha}{F_c - F_t \tan \alpha}. \tag{21.14}$$

Although the magnitude of forces in actual cutting operations is generally on the order of a few hundred newtons, the *local stresses* in the cutting zone and the *pressure* on the cutting tool are very high because the contact areas are very small. For example, the tool–chip contact length (see Fig. 21.3) is typically on the order of 1 mm. Consequently, the tool tip is subjected to very high stresses, which lead to wear as well as chipping and fracture of the tool.

Thrust Force. The *thrust force* in cutting is important because the toolholder, the work-holding devices, and the machine tool itself must be sufficiently stiff to support that force with minimal deflections. For example, if the thrust force is too high or if the machine tool is not sufficiently stiff, the tool will deflect away from the workpiece. This movement will, in turn, reduce the depth of cut, resulting in poor dimensional accuracy in the machined part.

The effect of rake angle and friction angle on the magnitude and direction of thrust force can be determined by noting, from Fig. 21.11b, that

$$F_t = R\sin(\beta - \alpha), \tag{21.15}$$

or

$$F_t = F_c\tan(\beta - \alpha). \tag{21.16}$$

The magnitude of the cutting force, F_c, is always positive, as shown in Fig. 21.11, because it is this force that supplies the work required in cutting. However, the sign of the thrust force, F_t, can be either positive or negative, depending on the values of β and α. Note that when $\beta > \alpha$, the sign of F_t is positive (*downward*), and when $\beta < \alpha$, the sign is negative (*upward*). It is therefore possible to have an upward thrust force under the conditions of (a) high rake angles, (b) low friction at the tool–chip interface, or (c) both. A negative thrust force can have important implications in the design of machine tools and workholders and in the stability of the cutting process.

Power. It can be seen from Fig. 21.11 that the power input in cutting is

$$\text{Power} = F_cV. \tag{21.17}$$

The power is dissipated mainly in the shear zone (due to the energy required to shear the material) and on the rake face of the tool (due to tool–chip interface friction). From Figs. 21.4b and 21.11, the power dissipated in the shear plane is

$$\text{Power for shearing} = F_sV_s. \tag{21.18}$$

Denoting the width of cut as w, the **specific energy for shearing**, u_s, is given by

$$u_s = \frac{F_sV_s}{wt_oV}. \tag{21.19}$$

Similarly, the power dissipated in friction is

$$\text{Power for friction} = FV_c, \tag{21.20}$$

and the **specific energy for friction**, u_f, is

$$u_f = \frac{FV_c}{wt_oV} = \frac{Fr}{wt_o}. \tag{21.21}$$

The **total specific energy**, u_t, is thus

$$u_t = u_s + u_f. \tag{21.22}$$

Because numerous factors are involved, reliable prediction of cutting forces and power still is based largely on experimental data, such as those given in Table 21.2. The wide range of values seen in the table can be attributed to differences in strength within each material group, and to other factors, such as friction, use of cutting fluids, the wide range in process parameters, and the sharpness of the tool tip. Dull tools require higher power and result in higher forces because the tip rubs against the machined surface and makes the deformation zone ahead of the tool larger.

Measuring Cutting Forces and Power. Cutting forces can be measured using a **force transducer** (typically with quartz piezoelectric sensors), a **dynamometer**, or a **load cell** (with resistance-wire strain gages placed on octagonal rings) mounted on the cutting-tool holder. It is also possible to *calculate* the cutting force from the **power consumption** during cutting, using Eq. (21.4).

It should be recognized that Eq. (21.4) represents the power in the machining process itself, and the machine tool will need additional power in order to overcome friction. Thus, to determine the cutting

Table 21.2: Approximate Range of Energy Requirements in Cutting Operations at the Drive Motor of the Machine Tool, Corrected for 80% Efficiency (for dull tools, multiply by 1.25).

Material	Specific energy W-s/mm^3
Aluminum alloys	0.4–1
Cast irons	1.1–5.4
Copper alloys	1.4–3.2
High-temperature alloys	3.2–8
Magnesium alloys	0.3–0.6
Nickel alloys	4.8–6.7
Refractory alloys	3–9
Stainless steels	2–5
Steels	2–9
Titanium alloys	2–5

force from the measured machine power consumption, the *mechanical efficiency* of the machine tool must be known. The *specific energy* in cutting, such as that shown in Table 21.2, also can be used to estimate cutting forces.

Example 21.1 Relative Energies in Cutting

Given: In an orthogonal cutting operation, $t_o = 0.1$ mm, $V = 2$ m/s, $\alpha = 10°$, and the width of cut is 5 mm. It is observed that $t_c = 0.20$ mm, $F_c = 500$ N, and $F_t = 200$ N.

Find: Calculate the percentage of the total energy that goes into overcoming friction at the tool–chip interface.

Solution: The percentage of the energy can be expressed as

$$\frac{\text{Friction energy}}{\text{Total energy}} = \frac{FV_c}{F_cV} = \frac{Fr}{F_c},$$

where

$$r = \frac{t_o}{t_c} = \frac{0.1}{0.20} = 0.50,$$

$$F = R\sin\beta,$$

$$F_c = R\cos(\beta - \alpha),$$

and

$$R = \sqrt{F_t^2 + F_c^2} = \sqrt{200^2 + 500^2} = 538 \text{ N}.$$

Thus,

$$500 = 538\cos(\beta - 10°),$$

so $\beta = 32°$ and

$$F = 538\sin 32° = 285 \text{ N}.$$

Hence,

$$\text{Percentage} = \frac{(285)(0.5)}{500} = 0.28, \quad \text{or } 28\%.$$

21.4 Temperatures in Cutting

As in all metalworking processes involving plastic deformation (Chapters 13 through 16), the energy dissipated in cutting is converted into *heat* which, in turn, raises the temperature in the cutting zone and the workpiece surface. *Temperature rise* is a major factor in machining because of its various adverse effects:

- Excessive temperature lowers the strength, hardness, stiffness, and wear resistance of the cutting tool; tools may also soften and undergo plastic deformation, thus the altering tool shape.
- Heat causes uneven dimensional changes in the part being machined, thus making it difficult to control its dimensional accuracy and tolerances.
- An excessive temperature rise can induce thermal damage and metallurgical changes (Chapter 4) in the machined surface, adversely affecting properties.

The main sources of heat in machining are: (a) work done in shearing in the primary shear zone, (b) energy dissipated as friction at the tool–chip interface, and (c) heat generated as the tool rubs against the machined surface, especially with dull or worn tools. Much effort has been expended in establishing relationships among temperature and various material and process variables in cutting. It can be shown that, in *orthogonal cutting*, the *mean temperature*, T_{mean}, in K is

$$T_{\text{mean}} = \frac{0.000665\sigma_f}{\rho c}\sqrt[3]{\frac{Vt_o}{K}}, \tag{21.23}$$

where σ_f is the flow stress (see Section 14.2), in MPa, ρc is the volumetric specific heat in kJ/m^3·K, and K is the thermal diffusivity (ratio of thermal conductivity to volumetric specific heat) in m^2/s. Because the material parameters in this equation also depend on temperature, it is important to use appropriate values that are applicable to the predicted temperature range. It can be seen from Eq. (21.23) that the *mean* cutting temperature increases with workpiece strength, cutting speed, and depth of cut, and decreases with increasing specific heat and thermal conductivity of the workpiece material.

A simple expression for the *mean temperature in turning* on a lathe is given by

$$T_{\text{mean}} \propto V^a f^b, \tag{21.24}$$

where V is the cutting speed and f is the feed of the tool, as shown in Fig. 21.2. Approximate values of the exponents a and b are $a = 0.2$ and $b = 0.125$ for *carbide* tools and $a = 0.5$ and $b = 0.375$ for *high-speed steel* tools.

Temperature Distribution. Because the sources of heat generation in machining are concentrated in the primary shear zone and at the tool–chip interface, it is to be expected that there will be severe *temperature gradients* within the cutting zone. A typical temperature distribution is shown in Fig. 21.12; note the presence of severe gradients, and that the *maximum* temperature is about halfway up the tool–chip interface.

The temperatures typically developed in a *turning* operation on 52100 steel are shown in Fig. 21.13. The temperature distribution along the *flank surface* of the tool is shown in Fig. 21.13a for $V = 60$, 90, and 170 m/min as a function of the distance from the tip of the tool. The distributions at the *tool–chip interface* for the same three cutting speeds are shown in Fig. 21.13b as a function of the fraction of the contact length. Thus, zero on the abscissa represents the tool tip, and 1.0 represents the end of the tool–chip contact length.

Note from Eq. (21.23) that the temperature increases with cutting speed and that the highest temperature is almost 1100°C. The presence of such high temperatures in machining can be verified simply by observing the dark-bluish color of the chips (caused by oxidation) typically produced at high cutting speeds. Chips can indeed become red hot, and thus create a safety hazard.

From Eq. (21.24) and the values for the exponent a, it can be seen that the cutting speed, V, greatly influences temperature. The explanation is that, as speed increases, the time for heat dissipation decreases,

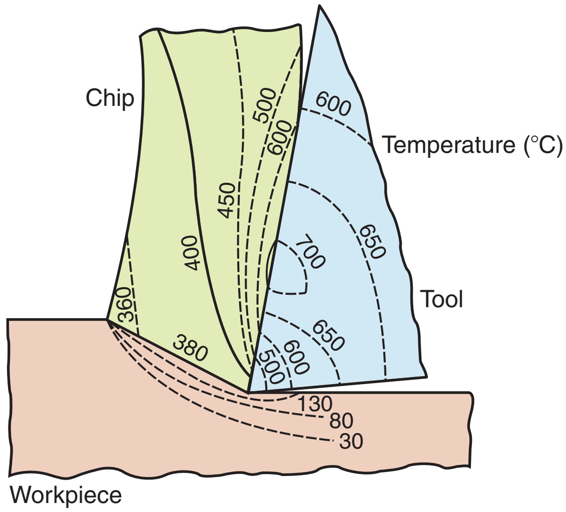


Figure 21.12: Typical temperature distribution in the cutting zone. Note the severe temperature gradients within the tool and the chip, and that the workpiece is relatively cool.

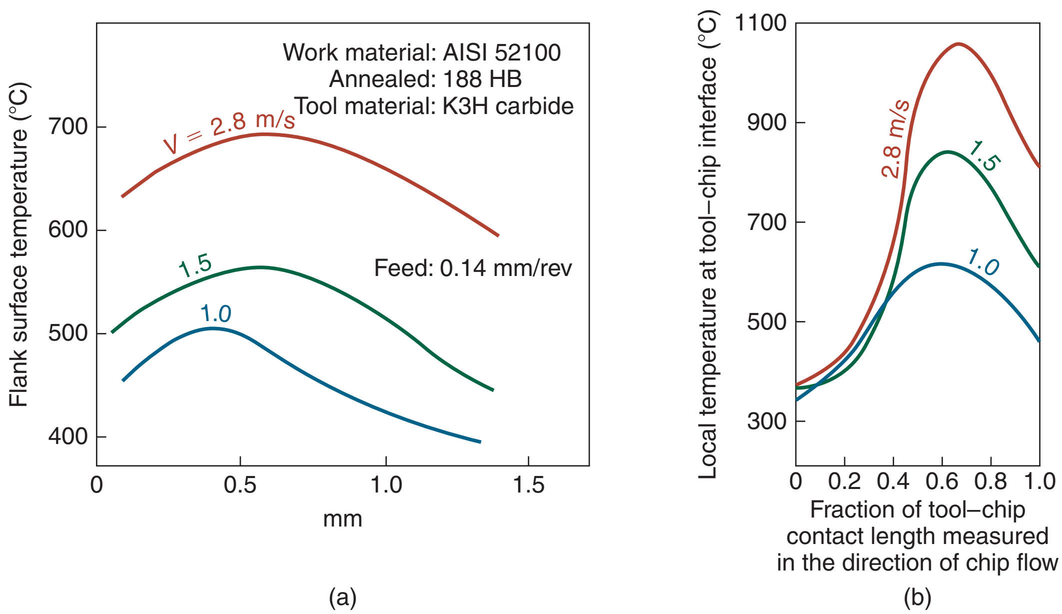


Figure 21.13: Temperatures developed in turning 52100 steel: (a) flank temperature distribution and (b) tool–chip interface temperature distribution. *Source:* After B.T. Chao and K.J. Trigger.

and hence the temperature rises, eventually becoming almost an *adiabatic* process. The effect of speed can be simulated easily by rubbing hands together faster and faster.

As can be seen in Fig. 21.14, the chip carries away most of the heat generated. In a typical machining operation, it has been estimated that 90% of the energy is removed by the chip, with the remainder taken by the tool and the workpiece. Note also that, as the cutting speed increases, a larger proportion of the total

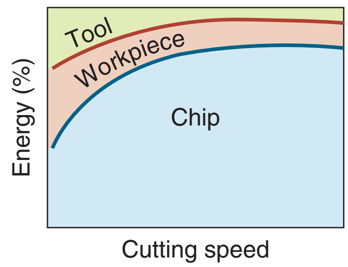


Figure 21.14: Proportion of the heat generated in cutting transferred to the tool, workpiece, and chip as a function of the cutting speed. Note that the chip removes most of the heat.

heat generated is carried away by the chip, and less heat is transferred elsewhere. This is one reason for the continued trend of increasing machining speeds (see *high-speed machining*, Section 25.5). The other main benefit of higher cutting speeds is associated with the favorable economics in reducing machining time (see Section 25.8).

Techniques for Measuring Temperature. Temperatures and their distribution in the cutting zone may be determined using **thermocouples**, embedded in the tool or the workpiece. The *mean* temperature can be determined using the **thermal emf** (electromotive force) at the tool–chip interface, which acts as a *hot junction* between two different materials (tool and chip). A third method is monitoring the **infrared radiation** from the cutting zone, using sensors; however, this technique indicates only surface temperatures, and its accuracy depends on the emissivity of the surfaces, which can be difficult to determine accurately.

21.5 Tool Life: Wear and Failure

It can be noted from the previous sections that cutting tools are subjected to (a) high localized stresses at the tip of the tool, (b) high temperatures, especially along the rake face, (c) sliding of the chip at relatively high speeds along the rake face, and (d) sliding of the tool along the newly machined workpiece surface. These conditions induce **tool wear**, a major consideration in all machining operations (as are mold and die wear in casting and metalworking processes). Tool wear, in turn, adversely affects tool life, the quality of the machined surface, its dimensional accuracy, and, consequently, the economics of machining operations.

Wear is a gradual process (see Section 33.5), much like the wear of the tip of an ordinary pencil. The *rate* of tool wear (that is, volume worn per unit time) depends on the workpiece material, tool material and its coatings, tool geometry, process parameters, cutting fluids, and characteristics of the machine tool. Tool wear and the resulting changes in tool geometry (Fig. 21.15) are generally classified as: *flank wear*, *crater wear*, *nose wear*, *notching*, *plastic deformation*, *chipping*, and *gross fracture*.

21.5.1 Flank Wear

Flank wear occurs on the relief (flank) face of the tool, as shown in Fig. 21.15a, b, and e. It generally is attributed to (a) rubbing of the tool along the machined surface, thereby causing adhesive or abrasive wear and (b) high temperatures, adversely affecting tool-material properties.

In a classic study by F.W. Taylor on the machining of steels conducted in the early 1890s, the following approximate relationship for tool life, known as the *Taylor tool life equation*, was established:

$$VT^n = C. \tag{21.25}$$

where V is the cutting speed, T is the time (in minutes) that it takes to develop a certain flank **wear land** (shown as VB in Fig. 21.15a), n is an exponent that depends on tool and workpiece materials and cutting

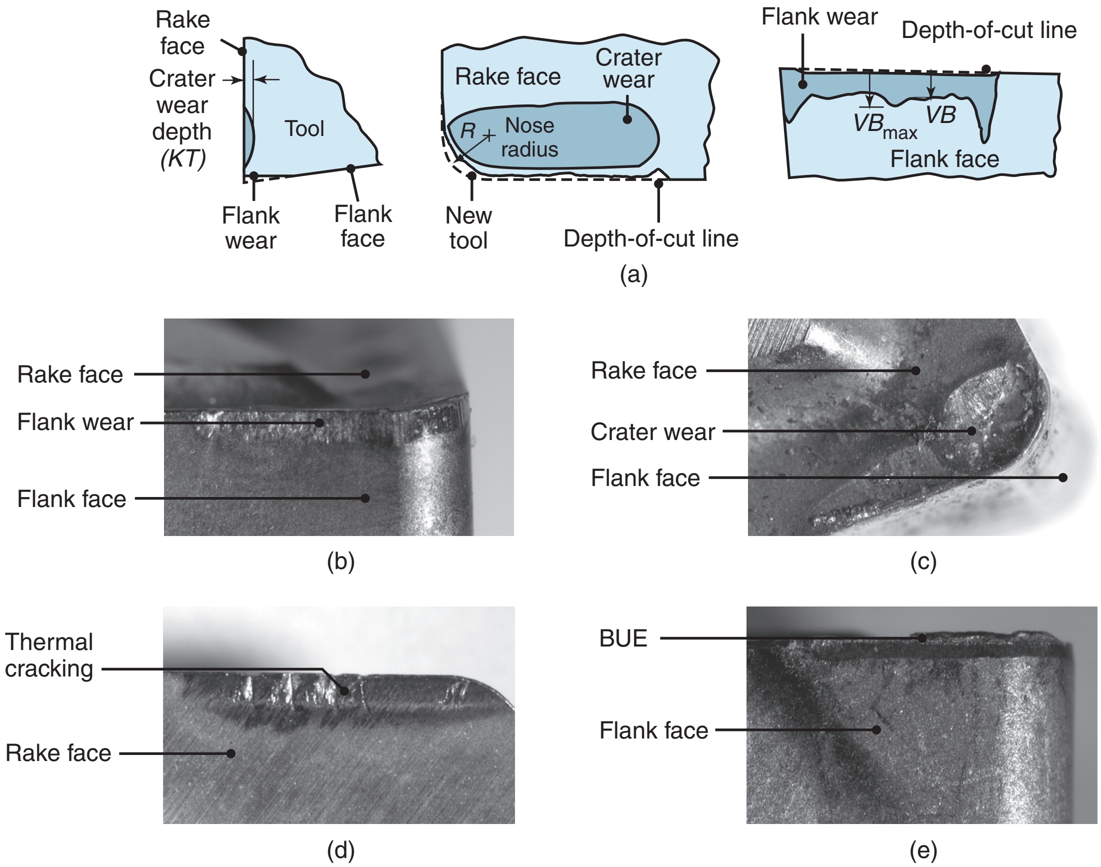


Figure 21.15: (a) Features of tool wear in a turning operation. The VB indicates average flank wear. (b)–(e) Examples of wear in cutting tools: (b) flank wear, (c) crater wear, (d) thermal cracking, and (e) flank wear and built-up edge. *Source:* (a) Terms and definitions reproduced with the permission of the International Organization for Standardization, ISO, copyright remains with ISO. (b)–(e) Courtesy of Kennametal Inc.

conditions, and C is a constant. Each combination of workpiece and tool materials and each cutting condition have their own n and C values, both of which are determined experimentally, often based on surface finish requirements. Moreover, the Taylor equation is often applied even when flank wear is not the dominant wear mode (see Fig. 21.15), or if a different criterion (such as the machining power required) is used to define C and n. Generally, n depends on the tool material, as shown in Table 21.3, and C on the workpiece material. Note that the magnitude of C is the cutting speed at $T = 1$ min.

To appreciate the importance of the exponent n, Eq. (21.25) can be rewritten as

$$T = \left(\frac{C}{V}\right)^{1/n}, \tag{21.26}$$

where it can be seen that for a constant value of C, the smaller the value of n, the lower is the tool life.

Cutting speed is the most important variable associated with tool life, followed by depth of cut and feed, f. For turning, Eq. (21.25) can be modified as

$$VT^n d^x f^y = C, \tag{21.27}$$

Table 21.3: Ranges of n Values for the Taylor Equation [Eq. (21.25)] for Various Tool Materials.

High-speed steels	0.08–0.2
Cast alloys	0.1–0.15
Carbides	0.2–0.5
Coated carbides	0.4–0.6
Ceramics	0.5–0.7

where d is the depth of cut and f is the feed in mm/rev, as shown in Fig. 21.2. The exponents x and y must be determined experimentally for each cutting condition. Taking $n = 0.15$, $x = 0.15$, and $y = 0.6$ as typical values encountered in machining practice, it can be seen that cutting speed, feed rate, and depth of cut are of decreasing importance. Equation (21.27) can be rewritten as

$$T = C^{1/n}V^{-1/n}d^{-x/n}f^{-y/n}, \tag{21.28}$$

or, using typical values for the exponents, as

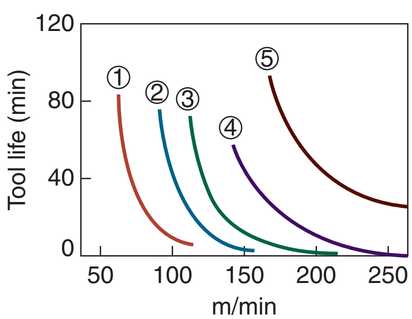
$$T \approx C^{7}V^{-7}d^{-1}f^{-4}. \tag{21.29}$$

For a constant tool life, the following observations can be made from Eq. (21.29):

- If the feed or the depth of cut is increased, the cutting speed must be decreased, and vice versa.
- Depending on the magnitude of the exponents, a reduction in speed can result in an increase in the volume of the material removed, because of the increased feed or depth of cut.

Tool-life Curves. *Tool-life curves* are plots of experimental data, obtained from cutting tests for various materials and under different cutting conditions, such as cutting speed, feed, depth of cut, tool material and geometry, and cutting fluids. Note in Fig. 21.16, for example, that (a) tool life decreases rapidly as the cutting speed increases, (b) the condition of the workpiece material has a strong influence on tool life, and (c) there is a large difference in tool life for different microstructures of the workpiece material.

Heat treatment of the workpiece is important, due largely to increasing workpiece hardness; for example, ferrite has a hardness of about 100 HB, pearlite 200 HB, and martensite 300 to 500 HB. Impurities



	Hardness (HB)	Ferrite	Pearlite
1 As cast	265	20%	80%
2 As cast	215	40	60
3 As cast	207	60	40
4 Annealed	183	97	3
5 Annealed	170	100	—

Figure 21.16: Effect of workpiece hardness and microstructure on tool life in turning ductile cast iron. Note the rapid decrease in tool life (approaching zero) as the cutting speed increases. Tool materials have been developed that resist high temperatures, such as carbides, ceramics, and cubic boron nitride, as described in Chapter 22.

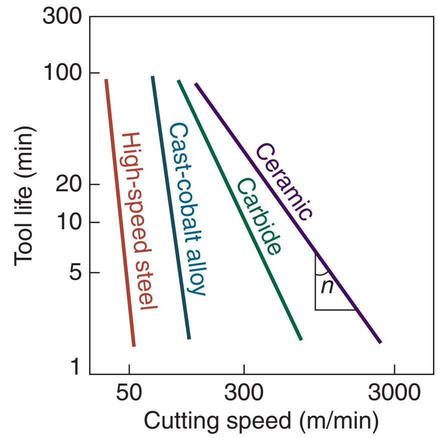


Figure 21.17: Tool-life curves for a variety of cutting-tool materials. The negative reciprocal of the slope of these curves is the exponent n in the Taylor tool-life equation [Eq. (21.25)], and C is the cutting speed at $T = 1$ min, ranging from about 60 to 3000 m/min in this figure.

and hard constituents in the material or on the workpiece surface, such as rust, scale, and slag, also are important factors, because their abrasive action (see Section 33.5) reduces tool life.

The exponent n can be determined from tool-life curves (Fig. 21.17). Note that the smaller the value of n, the faster the tool life decreases with increasing cutting speed. Although tool-life curves are somewhat linear over a limited range of cutting speeds, they rarely are linear over a wide range. Moreover, n can indeed become *negative* at low cutting speeds, meaning that tool-life curves actually can reach a maximum and then curve downward. Caution should therefore be exercised in using tool-life equations beyond the range of cutting speeds to which they are applicable.

Because temperature has a major influence on the physical and mechanical properties of materials (see Chapters 2 and 3), it is to be expected that temperature also strongly influences wear. Thus, as temperature increases, wear increases.

Example 21.2 Increasing Tool Life by Reducing the Cutting Speed

Given: Assume that for a given tool and workpiece combination, $n = 0.5$ and $C = 400$.

Find: Calculate the percentage increase in tool life when the cutting speed is reduced by 50%, using the Taylor equation [Eq. (21.25)] for tool life.

Solution: Since $n = 0.5$, the Taylor equation can be rewritten as $VT^{0.5} = 400$. Denote V_1 as the initial speed and V_2 as the reduced speed; thus, $V_2 = 0.5V_1$. Because C is a constant at 400,

$$0.5V_1\sqrt{T_2} = V_1\sqrt{T_1}.$$

Simplifying this equation,

$$\frac{T_2}{T_1} = \frac{1}{0.25} = 4.$$

Thus the change in tool life is

$$\frac{T_2 - T_1}{T_1} = \left(\frac{T_2}{T_1}\right) - 1 = 4 - 1 = 3,$$

or that tool life is increased by 300%. Note that a *reduction* in cutting speed has resulted in a major *increase* in tool life. Note also that, for this problem, the magnitude of C is not relevant.

Table 21.4: Allowable Average Wear Land (see VB in Fig. 21.15a) for Cutting Tools in Various Machining Operations.

	Allowable wear land (mm)	
Operation	**High-speed steel tools**	**Carbide tools**
Turning	1.5	0.4
Face milling	1.5	0.4
End milling	0.3	0.3
Drilling	0.4	0.4
Reaming	0.15	0.15

Note: Allowable wear for ceramic tools is about 50% higher. Allowable notch wear (see Section 21.5.3), $VB_{\max}$, is about twice that for VB.

Allowable Wear Land. Cutting tools need to be resharpened or replaced when (a) the surface finish of the machined workpiece begins to deteriorate, (b) cutting forces increase significantly, or (c) the temperature rises significantly. The *allowable wear land*, indicated as VB in Fig. 21.15a, is given in Table 21.4 for various machining conditions. For improved dimensional accuracy and surface finish, the allowable wear land may be smaller than the values given in the table. The *recommended cutting speed* for a high-speed steel tool (see Section 22.2) is generally the one that yields a tool life of 60 to 120 min, and for a carbide tool (Section 22.4), it is 30 to 60 min.

Optimum Cutting Speed. Recall that as cutting speed increases, tool life decreases rapidly. On the other hand, if the cutting speed is low, tool life is longer, but the rate at which material is removed is also low. Thus, there is an *optimum cutting speed*, based on economic or production considerations, where the tool life is long and production speeds are reasonably high. This topic is described in greater detail in Section 25.8.

Example 21.3 Effect of Cutting Speed on Material Removal

The effect of cutting speed on the volume of metal removed between tool changes or resharpenings can be appreciated by analyzing Fig. 21.16. Assume that a material is being machined, in the as-cast condition, with a hardness of 265 HB. Note that when the cutting speed is 60 m/min, tool life is about 40 min. Therefore, the tool travels a distance of 60 m/min $\times$ 40 min = 2400 m before it has to be replaced. However, when the cutting speed is increased to 120 m/min, tool life is reduced to about 5 min, and thus the tool travels 120 m/min $\times$ 5 min = 600 m before it has to be replaced.

Since the volume of material removed is directly proportional to the distance the tool has traveled, it can be seen that by *decreasing* the cutting speed, *more* material is removed between tool changes. Note, however, that the lower the cutting speed, the longer is the time required to machine a part, which has a significant economic impact on the operation (see Section 25.8).

21.5.2 Crater Wear

Crater wear occurs on the rake face of the tool, as shown in Fig. 21.15a and c, and Fig. 21.18, which also illustrates various types of tool wear and failures. It can be seen that crater wear alters the tool–chip contact geometry. The most significant factors that influence crater wear are (a) the temperature at the tool–chip interface and (b) the chemical affinity of the tool and workpiece materials. Additionally, the factors influencing flank wear also may affect crater wear.

Crater wear is generally attributed to a **diffusion** mechanism: the movement of atoms across the tool–chip interface. Because diffusion rate increases with increasing temperature, crater wear also increases as temperature increases. Note in Fig. 21.19, for example, how rapidly crater wear increases with temperature within a narrow range. Applying protective *coatings* to tools is an effective means of slowing

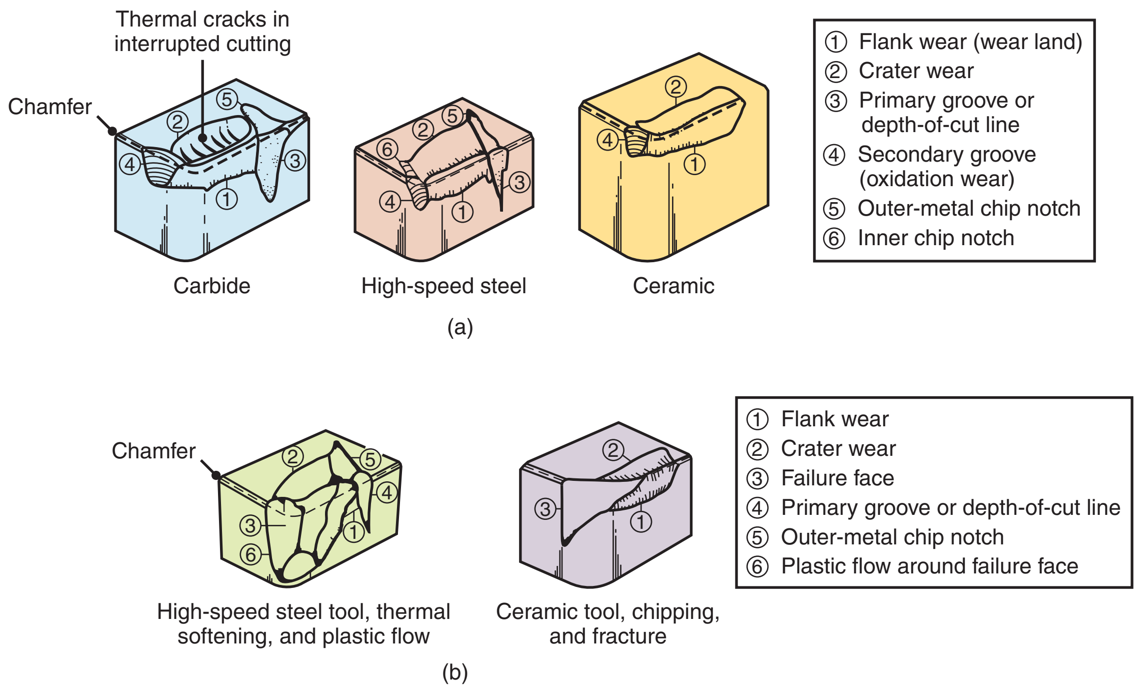


Figure 21.18: (a) Schematic illustrations of types of wear observed on various cutting tools. (b) Schematic illustrations of catastrophic tool failures. A wide range of parameters influence these wear and failure patterns. *Source:* Courtesy of V.C. Venkatesh.

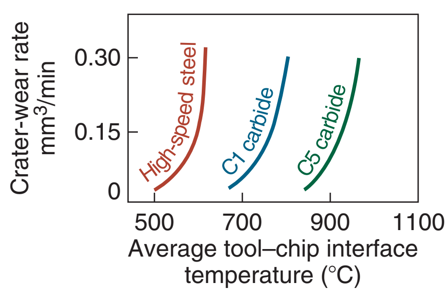


Figure 21.19: Relationship between crater-wear rate and average tool–chip interface temperature: (1) high-speed steel, (2) C1 carbide, and (3) C5 carbide (see Table 22.5). Note how rapidly crater-wear rate increases with an incremental increase in temperature. *Source:* After B.T. Chao and K.J. Trigger.

the diffusion process, and thus reducing crater wear. Typical tool coatings are titanium nitride, titanium carbide, diamondlike carbon, titanium carbonitride, and aluminum oxide, and are described in greater detail in Section 22.6.

In comparing Figs. 21.12 and 21.15a, it can be seen that the location of the *maximum depth* of crater wear, KT, coincides with the location of the *maximum temperature* at the tool–chip interface. An actual cross section of this interface, for steel machined at high speeds, is shown in Fig. 21.20. Note that the wear pattern on the tool face coincides with its discoloration pattern, an indication of the presence of high temperatures.

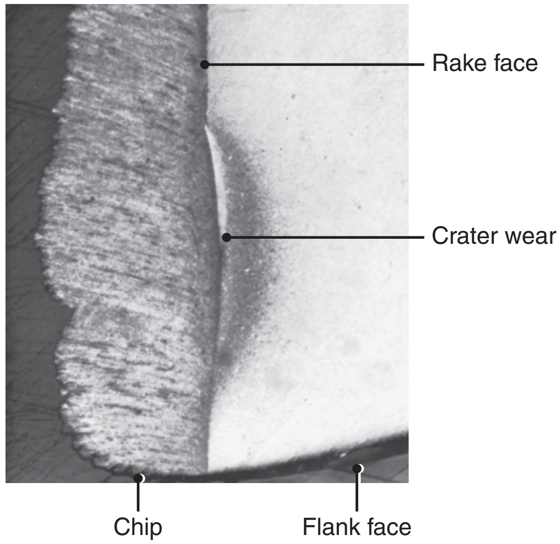


Figure 21.20: Interface of a cutting tool (right) and chip (left) in machining plain-carbon steel. The discoloration of the tool indicates the presence of high temperatures. Compare this figure with the temperature profiles shown in Fig. 21.12. *Source:* After P.K. Wright.

21.5.3 Other Types of Wear, Chipping, and Fracture

Nose wear (Fig. 21.15a) is the rounding of a sharp tool due to mechanical and thermal effects. It dulls the tool, affects the type of chip formation, and causes rubbing of the tool over the workpiece, raising temperature and inducing residual stresses on the machined surface. A related phenomenon is **edge rounding**, as shown in Fig. 21.15a.

An increase in temperature is particularly important for high-speed steel tools, as can be appreciated from Fig. 22.1. Tools may also undergo **plastic deformation**, because of temperature rises in the cutting zone where temperatures can easily reach 1000°C in machining steels, and can even be higher depending on the strength of the material machined.

Notches or **grooves** that develop on cutting tools, as shown in Figs. 21.15a and 21.18, have been attributed to the fact that the region where they occur is the boundary where the chip is no longer in contact with the tool. Known as the **depth-of-cut line** (DOC), see Fig. 21.15a, this boundary oscillates, because of inherent variations in the cutting operation. If sufficiently deep, the groove can lead to gross chipping of the tool tip because of (a) its now reduced cross section and (b) the notch sensitivity of the tool material.

Scale and *oxide layers* on a workpiece surface also contribute to **notch wear**, because these layers are hard and abrasive. Thus, light cuts should be avoided on such workpieces. In Fig. 21.3 for example, the depth of cut, t_o, should be greater than the thickness of the scale on the workpiece.

In addition to being subjected to wear, cutting tools may also undergo **chipping**, where a small fragment from the cutting edge of the tool breaks away. This phenomenon, which typically occurs in brittle tool materials such as ceramics, is similar to chipping of the tip of a pencil if it is too sharp. The chipped fragments from the cutting tool may be very small (called **microchipping** or **macrochipping**, depending on its size), or they may be relatively large, in which case they are variously called **gross chipping**, **gross fracture**, and **catastrophic failure** (Fig. 21.18).

Chipping also may occur in a region of the tool where there is a preexisting small crack or a defect during its production. Unlike wear, which is a gradual process, chipping is a sudden loss of the tool material, thus changing the tool's shape. As can be expected, chipping has a major detrimental effect on surface finish, surface integrity, and the dimensional accuracy of the workpiece being machined.

Two main causes of chipping are:

- **Mechanical shock**, such as impact due to interrupted cutting, as in turning a splined shaft on a lathe.
- **Thermal fatigue**, due to cyclic temperature variations within the tool in interrupted cutting.

Thermal cracks usually perpendicular to the cutting edge of the tool, as shown on the rake face of the carbide tool in Figs. 21.15d and 21.18a. Major variations in the composition or structure of the workpiece material also may cause chipping, due to differences in their thermal properties.

Chipping can be reduced by selecting tool materials with high impact and thermal-shock resistance, as described in Chapter 22. High positive-rake angles can contribute to chipping because of the small included angle of the tool tip, as can be visualized from Fig. 21.3. Also, it is possible for the crater-wear region to progress slowly toward the tool tip, thus weakening the tip because of reduced volume of material.

21.5.4 Tool-condition Monitoring

With rapid advances in computer-controlled machine tools and automated manufacturing, the reliable and repeatable performance of cutting tools is a major consideration. As described in Chapters 23 through 25, modern machine tools operate with little direct supervision by an operator. Moreover, they are typically enclosed, making it virtually impossible to closely monitor the machining operation and the condition of the cutting tool. It is thus essential to indirectly and continuously monitor the condition of the cutting tool. In machine tools, tool-condition monitoring systems are now integrated into computer numerical control and programmable logic controllers.

Techniques for tool-condition monitoring typically fall into two general categories: direct and indirect. The **direct method** for observing the condition of a cutting tool involves *optical* measurements of wear, such as periodic observation of changes in the tool profile. This is a common technique, and is done using a *toolmakers' microscope*. However, this method requires that the cutting operation be stopped for tool observation. Another direct method involves programming the tool to contact a sensor (*touch probe*) after each machining cycle; this approach allows the measurement of wear and/or the detection of broken tools.

Indirect methods involve correlating the tool condition with parameters such as cutting forces, power, temperature rise, workpiece surface finish, vibration, and chatter. A common technique is **acoustic emission** (AE), which utilizes a *piezoelectric transducer* mounted on a toolholder. The transducer picks up acoustic emissions (typically above 100 kHz) which result from the stress waves generated during cutting. By analyzing the signals, tool wear and chipping can be monitored. This technique is effective particularly in precision-machining operations, where cutting forces are low (because of the small amounts of material removed). Another effective use of AE is in detecting the fracture of small carbide tools at high cutting speeds.

A similar indirect technique consists of various sensors that are installed in the original machine tool, or are retrofitted on existing machines. The system continually monitors torque and forces during machining. The signals are analyzed and interpreted to differentiate between events: tool breakage, tool wear, a missing tool, overloading of the machine tool, or colliding with machine components. This system also can compensate automatically for tool wear, and thus improve the dimensional accuracy of the part being machined.

The design of these systems must be such that they are (a) nonintrusive to the machining operation, (b) accurate and repeatable in signal detection, (c) resistant to abuse, (d) robust for the shop-floor environment (see Sections 36.5.1 and 40.7), and (e) cost effective. **Sensors**, including the use of *infrared* and *fiber-optic techniques* for temperature measurement during machining, are important components of the system.

21.6 Surface Finish and Integrity

Surface finish influences not only the dimensional accuracy of machined parts but also their properties and performance in service. The term *surface finish* describes the *geometric features* of a surface (see Chapter 33); *surface integrity* pertains to *properties*, such as fatigue life and corrosion resistance, that are strongly influenced by the nature of the surface produced.

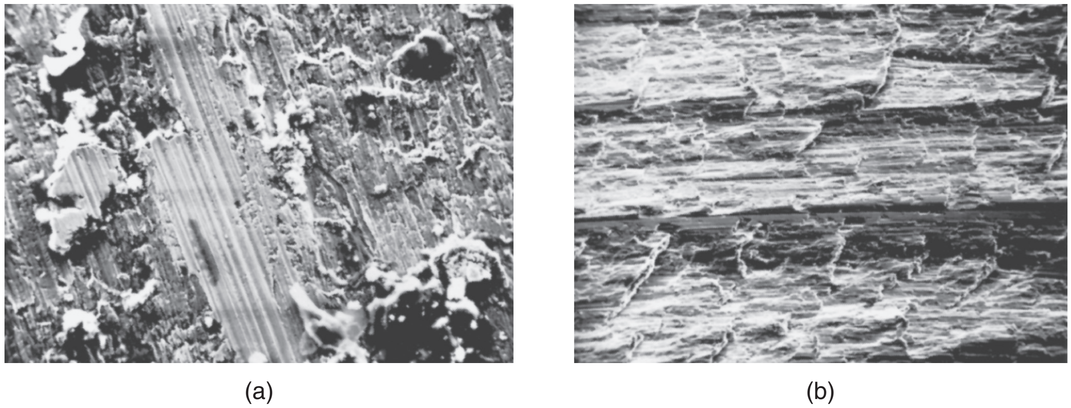


Figure 21.21: Machined surfaces produced on steel (highly magnified), as observed with a scanning-electron microscope: (a) turned surface and (b) surface produced by shaping. *Source:* After J T. Black and S. Ramalingam.

With its significant effect on changing the tool-tip profile, the *built-up edge* (see Fig. 21.6) has the greatest influence on surface finish. Fig. 21.12 shows the surfaces produced in two different cutting operations. Note the considerable damage to the surfaces from BUE and scuffing marks; they deviate from the straight grooves (tool marks) resulting from normal machining, as seen in Fig. 21.2. Ceramic and diamond cutting tools generally produce a better surface finish than other tools, largely because of their much lower tendency to form a BUE.

A *dull* tool has a large radius along its cutting edges, as in the tip of a dull pencil or the edge of a knife. Figure 21.22 illustrates the relationship between the radius of the cutting edge and the depth of cut in orthogonal cutting. Note that at small depths of cut, the rake angle effectively can become negative, and the tool simply may ride over the workpiece surface, instead of cutting it and producing chips. This is a phenomenon similar to trying to scrape a thin layer from the surface of a stick of butter with a dull knife.

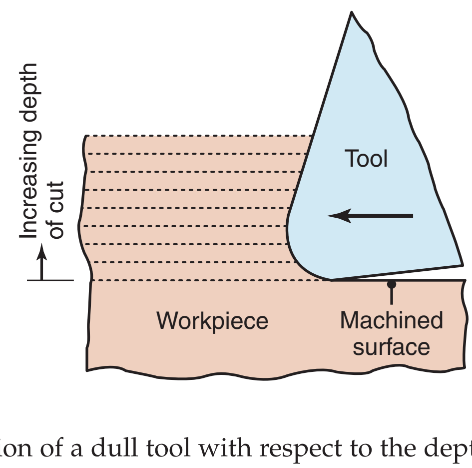


Figure 21.22: Schematic illustration of a dull tool with respect to the depth of cut in orthogonal machining (exaggerated). Note that the tool has a positive rake angle, but as the depth of cut decreases, the rake angle effectively can become negative. The tool then simply rides over the workpiece (without cutting) and burnishes its surface; this action raises the workpiece temperature and causes surface residual stresses.

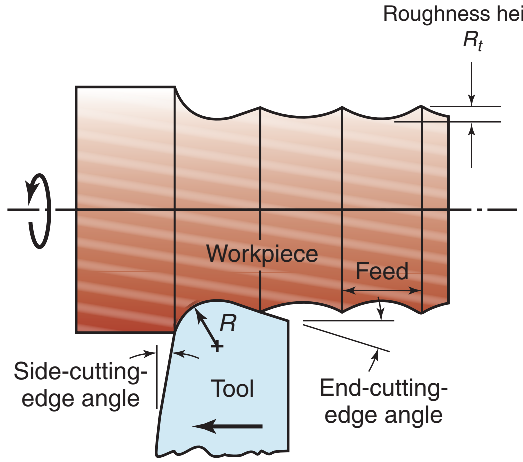


Figure 21.23: Schematic illustration of feed marks on a surface being turned (exaggerated).

If the tip radius of the tool (not to be confused with radius R in Fig. 21.15a) is large in relation to the depth of cut, the tool simply will rub over the machined surface. Rubbing will then generate heat and induce residual surface stresses, which in turn may cause surface damage, such as tearing and cracking. Consequently, the depth of cut should be greater than the radius on the cutting edge.

In a turning operation, as in other cutting processes described in the rest of Part IV of this text, the tool leaves a spiral profile (called **feed marks**) on the machined surface as it moves across the workpiece (Figs. 21.2 and 21.23). It will be noted that the higher the feed, f, and the smaller the tool-nose radius, R, the more prominent the feed marks will be. It can be shown that the surface roughness, for such a case, is given by

$$R_t = \frac{f^2}{8R}, \tag{21.30}$$

where R_t is the *roughness height*, as described in Section 33.3. Although not significant in rough machining operations, feed marks are important in finish machining.

Vibration and chatter are described in detail in Section 25.4. For now, it should be recognized that if the tool vibrates or chatters during machining, it will adversely affect the workpiece surface finish produced, for the reason that a vibrating tool periodically changes the dimensions of the cut and the surface properties. Chatter can cause chipping and premature failure of the more brittle cutting tools, such as ceramics and diamond.

Factors influencing *surface integrity* are as follow:

- Temperatures generated during cutting and possible metallurgical transformations
- Surface residual stresses
- Plastic deformation and strain hardening of the machined surfaces, tearing, and cracking.

Each of these factors can have various adverse effects on the machined part.

The difference between **finish machining** and **rough machining** should be emphasized. In finish machining, it is important to consider the *surface finish* to be produced, whereas in rough machining the main purpose is to remove a large amount of material at a high rate, regardless of the surface finish produced, since it will be greatly improved during finish machining. It is important to note that there be no *subsurface damage* resulting from rough machining that cannot later be removed during finish machining (see Fig. 21.21).

21.7 Machinability

The *machinability* of a material is usually defined in terms of four factors:

1. Surface finish and surface integrity of the machined part
2. Tool life
3. Force and power requirements
4. The level of difficulty in chip control after it is generated.

Thus, for instance, good machinability indicates good surface finish and surface integrity, long tool life, and low force and power requirements. As for chip control, chips that are long, thin, stringy, and curled can severely interfere with the machining operation by becoming entangled in the cutting area of the machine tool (see Fig. 21.7).

Because of the complex nature of machining operations, it is difficult to establish relationships that quantitatively define the machinability of a particular material. In practice, tool life and surface roughness generally are considered to be the most important factors in machinability. In subsequent chapters, several tables are presented in which, for various groups of materials, *specific recommendations* are given regarding such parameters as cutting speed, feed, depth of cut, cutting tool materials, tool shape, and type of cutting fluids.

21.7.1 Machinability of Ferrous Metals

This section describes the machinability of steels, alloy steels, stainless steels, and cast irons.

Steels. Carbon steels have a wide range of machinability, depending on their ductility and hardness. If a certain carbon steel is too ductile, chip formation can include built-up edge, leading to poor surface finish. If the steel is too hard, it can cause abrasive wear of the tool, because of the presence of carbides in the steel. The machinability of most steels is improved by cold working, which hardens the material and reduces the tendency for built-up edge formation.

An important group of steels is **free-machining steels**, containing sulfur and phosphorus. *Sulfur* forms manganese-sulfide inclusions (*second-phase particles*, Section 4.2.3), which act as stress raisers in the primary shear zone. As a result, the chips produced break up easily and are small, thus improving machinability. The size, shape, distribution, and concentration of these inclusions significantly influence machinability. Elements such as *tellurium* and *selenium*, both of which are chemically similar to sulfur, act as *inclusion modifiers* in resulfurized steels.

Phosphorus in steels has two major effects: (a) it strengthens the ferrite, causing increased hardness and resulting in better chip formation and surface finish, and (b) it increases hardness and thus causes the formation of short chips instead of continuous stringy ones, thereby improving machinability.

In **leaded steels**, a high percentage of lead solidifies at the tips of manganese-sulfide inclusions. In nonresulfurized grades of steel, lead takes the form of dispersed fine particles. Lead is insoluble in iron, copper, and aluminum and their alloys, and because of its low shear strength, it acts as a *solid* lubricant (see Section 33.7) and is smeared over the tool–chip interface during machining.

When the temperature rise is sufficiently high, such as at high cutting speeds and feeds, the lead melts directly in front of the tool, acting as a liquid lubricant. Lead also lowers the shear stress of the materials in the primary shear zone, thus reducing cutting forces and power consumption. Lead can be used with every grade of steel and its presence is identified by the letter L between the second and third numerals in steel identification (e.g., 10L45). In stainless steels, use of the letter L means low carbon, which improves their corrosion resistance.

Because lead is a well-known *toxin* and a pollutant, there are continuing serious environmental concerns about its use in steels. There is a continuing trend toward eliminating the use of lead in steels (**lead-free steels**). *Bismuth* and *tin* are substitutes for lead in steels, but they are shown to be not as effective.

Calcium-deoxidized steels contain oxide flakes of calcium silicates (CaSO) that reduce the strength of the secondary shear zone, and decrease tool–chip interface friction and wear. Because temperature increase is reduced correspondingly, these steels produce less crater wear, especially at high cutting speeds.

Alloy steels can have a wide variety of compositions and hardness, thus their machinability cannot be generalized. An important trend in machining these steels is *hard turning*, as described in detail in Section 25.6. Alloy steels at hardness levels of 45 to 65 HRC can be machined with polycrystalline cubic-boron-nitride (cBN) cutting tools (see Section 22.7), producing good surface finish, integrity, and dimensional accuracy.

Effects of Various Elements in Steels. The presence of *aluminum* and *silicon* is always harmful, because these elements combine with oxygen, forming aluminum oxide and silicates, which are hard and abrasive. As a result, tool wear increases and machinability is reduced.

Carbon and *manganese* have various effects on the machinability of steels, depending on their composition. Plain low-carbon steels (less than 0.15% C) can produce poor surface finish by forming a built-up edge. Cast steels can be abrasive, although their machinability is similar to that of wrought steels. Tool and die steels are very difficult to machine, and usually require annealing prior to machining or using other machining techniques (see Chapter 27).

Other alloying elements, such as *nickel*, *chromium*, *molybdenum*, and *vanadium*, that otherwise improve the properties of steels also generally reduce machinability. The effect of *boron* is negligible. Gaseous elements such as *hydrogen* and *nitrogen* can have particularly detrimental effects on the properties of steel. Oxygen has been shown to have a strong effect on the aspect ratio of the manganese-sulfide inclusions (see also Section 2.10.1): The higher the oxygen content, the lower the aspect ratio, and the higher the machinability.

It is important to also consider the possible detrimental effects of the alloying elements on the properties and strength of machined parts in service. At elevated temperatures, for example, lead causes *embrittlement* of steels (liquid-metal embrittlement and hot shortness; see Section 1.5.2), although it has no effect on mechanical properties at room temperature.

Sulfur can reduce the hot workability of steels (see Section 14.5) severely, because of the formation of iron sulfide, unless sufficient manganese is present to prevent such formation. At room temperature, the mechanical properties of **resulfurized steels** depend on the orientation of the deformed manganese-sulfide inclusions. **Rephosphorized steels** are significantly less ductile, and are produced solely for the purpose of improving their machinability.

Stainless Steels. Austenitic (300 series) steels generally are difficult to machine. Chatter can be a problem, thus necessitating machine tools with high stiffness. Ferritic stainless steels (300 series) have good machinability. Martensitic (400 series) steels are abrasive, tend to form a built-up edge, and require tool materials with high hot hardness and crater-wear resistance. *Precipitation-hardening stainless steels* are strong and abrasive, thus requiring hard and abrasion-resistant tool materials.

Cast Irons. *Gray irons* generally are machinable, although they can be abrasive, depending on composition, especially pearlite. Free carbides in castings reduce their machinability, and can also cause tool chipping or fracture. *Nodular* and *malleable irons* are machinable, using hard tool materials.

21.7.2 Machinability of Nonferrous Metals

In alphabetic order:

- **Aluminum** is generally very easy to machine, although the softer grades tend to form a built-up edge, resulting in poor surface finish. High speeds, high rake angles, and high relief angles are recommended. Wrought aluminum alloys with high silicon content and cast aluminum alloys are generally abrasive, hence they require harder tool materials. Dimensional tolerance control may be a problem in machining aluminum, because it has a high thermal expansion coefficient and a relatively low elastic modulus.

- **Beryllium** is generally machinable; however, because the fine particles produced during machining are toxic, this requires a controlled environment.
- **Cobalt-based alloys** are abrasive and highly work hardening; they require sharp, abrasion-resistant tool materials and low feeds and speeds.
- **Copper**, in the wrought condition, can be difficult to machine, because of built-up edge formation. Cast copper alloys are easy to machine; brasses are easy to machine, especially with the addition of lead (*leaded free-machining brass*); note, however, the toxicity of lead and associated environmental concerns. New brasses and bronzes have been developed that are lead free; examples are bismuth-tin bronze, with 1 to 6% Bi and popular for bearing races; aluminum bronzes; and Envirobrass, used for drinking water pipes. Bronzes are more difficult to machine than brass.
- **Magnesium** is very easy to machine, with good surface finish and long tool life; however, care should be exercised because of its high rate of oxidation (*pyrophoric*) and hence the danger of fire.
- **Molybdenum** is ductile and work hardening; sharp tools are essential to prevent produce poor surface finish.
- **Nickel-based alloys and superalloys** are work hardening, abrasive, and strong at high temperatures. Their machinability depends on their condition and improves with annealing.
- **Tantalum** is very work hardening, ductile, and soft; it produces a poor surface finish, and tool wear is high.
- **Titanium** and its alloys have very poor thermal conductivity (the lowest of all metals, see Table 3.2), thus causing a significant temperature rise and built-up edge. They are highly reactive and can be difficult to machine.
- **Tungsten** is brittle, strong, and very abrasive; thus its machinability is low, although it improves greatly at elevated temperatures.
- **Zirconium** has good machinability, but it requires a coolant-type cutting fluid (Section 22.12) because of the danger of explosion and fire.

21.7.3 Machinability of Miscellaneous Materials

Thermoplastics generally have low thermal conductivity and low elastic modulus, and they are thermally softening. Consequently, machining them requires sharp tools with positive rake angles (to reduce cutting forces), large relief angles, small depths of cut and feed, relatively high speeds, and proper workholding devices to support of the workpiece. To keep the chips from becoming gummy and sticking to cutting tools, external cooling of the cutting zone may be necessary. Cooling can be achieved with a jet of air, a vapor mist, or using water-soluble oils.

Thermosetting plastics are brittle and sensitive to thermal gradients during cutting; machining conditions generally are similar to those of thermoplastics.

Polymer-matrix composites are very abrasive, because of the fibers present, hence they are difficult to machine. Fiber tearing, pullout, and edge delamination are significant problems, and can lead to severe reduction in the load-carrying capacity of machined components. Machining of these components requires careful handling and removal of debris, in order to avoid contact with and inhaling of the fibers.

Metal-matrix and ceramic-matrix composites can be difficult to machine, depending on the properties of the matrix material and the type of reinforcing fibers.

Graphite is abrasive; it requires sharp, hard, and abrasion-resistant tools.

Ceramics have a steadily improved machinability, particularly with the development of *machinable* ceramics and nanoceramics (Section 8.2), and with the selection of appropriate processing parameters, such as ductile-regime cutting (described in Section 25.7).

Wood is a complex material, with properties varying with its grain direction. Consequently, the type of chips and the surfaces produced also vary significantly, depending on the type of wood and its condition. Woodworking, which dates back to 3000 B.C., still remains largely an art. The basic requirements are generally sharp tools and high cutting speeds.

21.7.4 Thermally Assisted Machining

Metals and alloys that are difficult to machine at room temperature can be machined more easily at elevated temperatures. In thermally assisted machining, also called *hot machining*, a source of heat (such as a torch, induction coil, electric current, laser-beam, electron-beam, or plasma arc) is focused onto an area just ahead of the cutting tool. First investigated in the early 1940s, this operation typically is carried out above the *homologous temperature* of $T/T_m = 0.5$ (see Section 1.8, and Tables 1.2 and 3.1). Steels, for example, are hot machined above the temperature range of 650° to 750°C.

Although difficult and complicated to perform in production plants, the general advantages of hot machining are: (a) reduced cutting forces, (b) increased tool life, (c) higher material-removal rates, and (d) reduced tendency for vibration and chatter.

Summary

- Machining processes are often necessary to impart the desired dimensional accuracy, geometric features, and surface-finish characteristics to components, particularly those with complex shapes that cannot be produced economically using other shaping techniques. On the other hand, machining generally takes more time, wastes material, in the form of chips, and generally doesn't affect the bulk properties of the workpiece; however, it may have adverse effects on surfaces produced.

- Commonly observed chip types in machining are continuous, built-up edge, discontinuous, and serrated. Important process variables in machining are tool geometry and tool material; cutting conditions, such as speed, feed, and depth of cut; use of cutting fluids; and characteristics of the workpiece material and the machine tool. Parameters influenced by these variables are forces and power consumption, tool wear, surface finish and surface integrity, temperature rise, and dimensional accuracy of the workpiece.

- Temperature rise in machining is an important phenomenon, since it can have adverse effects on tool life, as well as on the properties, dimensional accuracy, and surface integrity of the machined part.

- Two principal types of tool wear are flank wear and crater wear. Tool wear depends on workpiece and tool material characteristics; cutting speed, feed, depth of cut, and cutting fluids; as well as the characteristics of the machine tool. Tool failure also may occur by notching, chipping, and gross fracture.

- The surface finish of machined components can adversely affect product integrity. Important variables are the geometry and condition of the cutting tool, the type of chip produced, and process variables.

- Machinability is generally defined in terms of surface finish, tool life, force and power requirements, and chip control. The machinability of specific materials depends on their composition, general properties, and microstructure.

Key Terms

Acoustic emission
Allowable wear land
Built-up edge
Chip
Chip breaker
Chip curl
Chipping of tool
Clearance angle
Continuous chip
Crater wear
Cutting force
Cutting ratio
Depth-of-cut line
Diffusion
Discontinuous chip
Feed marks
Flank wear
Friction angle
Hot machining
Inclination angle
Machinability
Machine tool
Machining
Notch wear
Oblique cutting
Orthogonal cutting
Primary shear zone
Rake angle
Relief angle
Rephosphorized steel
Resulfurized steel
Secondary shear zone
Serrated chip
Shaving
Shear angle
Shear plane
Skiving
Specific energy
Surface finish
Surface integrity
Taylor equation
Thrust force
Tool-condition monitoring
Tool life
Turning
Wear land

Bibliography

Ahmad, J., **Machining of Polymer Composites**, Springer, 2008.
Altintas, Y., **Machining Automation: Metal Cutting Mechanics, Machine Tool Vibrations, and CNC Design**, 2nd. ed., Cambridge, 2012.
Boothroyd, G., and Knight, W.A., **Fundamentals of Metal Machining and Machine Tools**, 3rd ed., CRC Press, 2006.
Cormier, D., **McGraw-Hill Machining and Metalworking Handbook**, McGraw-Hill, 2006.
Davim, J.P. (ed.), **Machinability of Advanced Materials**, Wiley, 2014.
—, **Machining of Complex Sculpted Surfaces**, Springer, 2012.
—, **Machining of Composite Materials**, Wiley, 2009.
—, **Machining of Hard Materials**. Springer, 2011.
—, **Machining: Fundamentals and Recent Advances**, Springer, 2010.
—, **Surface Integrity in Machining**, Springer, 2010.
Dudzinski, D., Molinari, A., and Schulz, H., (eds.), **Metal Cutting and High-Speed Machining**, Springer, 2002.

Erdel, B., **High-Speed Machining**, Society of Manufacturing Engineers, 2003.
Jackson, M.J., and Morell, J., **Machining with Nanomaterials**, Springer, 2009.
Liang, S. and Shih, A.J., **Analysis of Machining and Machine Tools**, Springer, 2016.
Marinescu, I.D., **Handbook of Advanced Ceramics Machining**, CRC Press, 2006.
Shaw, M.C., **Metal Cutting Principles**, 2nd ed., Oxford, 2005.
Stephenson, D.A., and Agapiou, J.S., **Metal Cutting: Theory and Practice**, 2nd ed., CRC Press, 2005.
Trent, E.M., and Wright, P.K., **Metal Cutting**, 4th ed., Butterworth-Heinemann, 2000.
Tschatsch, H., **Applied Machining Technology**, Springer, 2009.
Youssef, H.A. and El-Hofy, H., **Machining Technology, Machine Tools and Operations**, CRC Press, 2008.

Review Questions

21.1. Explain why continuous chips are not necessarily desirable.

21.2. Name the factors that contribute to the formation of discontinuous chips.

21.3. What is the cutting ratio? Is it always greater than 1? Explain.

21.4. Explain the difference between positive and negative rake angles. What is the importance of the rake angle?

21.5. Explain how a dull tool can lead to negative rake angles.

21.6. Comment on the role and importance of the relief angle.

21.7. Explain the difference between discontinuous chips and segmented chips.

21.8. Why should we be interested in the magnitude of the thrust force in cutting?

21.9. What are the differences between orthogonal and oblique cutting?

21.10. What is a BUE? Why does it form?

21.11. Is there any advantage to having a built-up edge on a tool? Explain.

21.12. What is the function of chip breakers? How do they function? Do you need a chip breaker to eliminate continuous chips in oblique cutting? Explain.

21.13. Identify the forces involved in a cutting operation. Which of these forces contributes to the power required?

21.14. Explain the characteristics of different types of tool wear.

21.15. List the factors that contribute to poor surface finish in cutting.

21.16. Explain what is meant by the term machinability and what it involves. Why does titanium have poor machinability?

21.17. What is shaving?

Qualitative Problems

21.18. List reasons that machining operations may be required, and provide an example for each reason.

21.19. Are the locations of maximum temperature and crater wear related? If so, explain why.

21.20. Is material ductility important for machinability? Explain.

21.21. Explain why studying the types of chips produced is important in understanding cutting operations.

21.22. Why do you think the maximum temperature in orthogonal cutting is located at about the middle of the tool–chip interface? (*Hint:* Note that the two sources of heat are (a) shearing in the primary shear plane and (b) friction at the tool–chip interface.)

21.23. Tool life can be almost infinite at low cutting speeds. Would you then recommend that all machining be done at low speeds? Explain.

21.24. Explain the consequences of allowing temperatures to rise to high levels in cutting.

21.25. The cutting force increases with the depth of cut and decreasing rake angle. Explain why.

21.26. Why is it not always advisable to increase the cutting speed in order to increase the production rate?

21.27. What are the consequences if a cutting tool chips?

21.28. What are the effects of performing a cutting operation with a dull tool? A very sharp tool?

21.29. To what factors do you attribute the difference in the specific energies in machining the materials shown in Table 21.2? Why is there a range of energies for each group of materials?

21.30. Explain why it is possible to remove more material between tool resharpenings by lowering the cutting speed.

21.31. Noting that the dimension d in Fig. 21.4a is very small, explain why the shear strain rate in metal cutting is so high.

21.32. Explain the significance of Eq. (21.9).

21.33. Comment on your observations regarding Figs. 21.12 and 21.13.

21.34. Describe the consequences of exceeding the allowable wear land (Table 21.4) for various cutting-tool materials.

21.35. Comment on your observations regarding the hardness variations shown in Fig. 21.6a.

21.36. Why does the temperature in cutting depend on the cutting speed, feed, and depth of cut? Explain in terms of the relevant process variables.

21.37. You will note that the values of a and b in Eq. (21.24) are higher for high-speed steels than for carbides. Why is this so?

21.38. As shown in Fig. 21.14, the percentage of the total cutting energy carried away by the chip increases with increasing cutting speed. Why?

21.39. Describe the effects that a dull tool can have on cutting operations.

21.40. Explain whether it is desirable to have a high or low (a) n value and (b) C value in the Taylor tool-life equation.

21.41. The Taylor tool-life equation is directly applicable to flank wear. Explain whether or not it can be used to model tool life if other forms of wear are dominant.

21.42. The tool-life curve for ceramic tools in Fig. 21.17 is to the right of those for other tool materials. Why?

21.43. Why are tool temperatures low at low cutting speeds and high at high cutting speeds?

21.44. Can high-speed machining be performed without the use of a cutting fluid?

21.45. Given your understanding of the basic metal-cutting process, what are the important physical and chemical properties of a cutting tool?

21.46. Explain why the power requirements in cutting depend on the cutting force but not the thrust force.

21.47. State whether or not the following statements are true, explaining your reasons: (a) For the same shear angle, there are two rake angles that give the same cutting ratio. (b) For the same depth of cut and rake angle, the type of cutting fluid used has no influence on chip thickness. (c) If the cutting speed, shear angle, and rake angle are known, the chip velocity can be calculated. (d) The chip becomes thinner as the rake angle increases. (e) The function of a chip breaker is to decrease the curvature of the chip.

Quantitative Problems

21.48. Let $n = 0.5$ and $C = 500$ in the Taylor equation for tool wear. What is the percent increase in tool life if the cutting speed is reduced by (a) 40% and (b) 80%?

21.49. Assume that, in orthogonal cutting, the rake angle is 25° and the coefficient of friction is 0.25. Using Eq. (21.4), determine the percentage increase in chip thickness when the friction is doubled.

21.50. Derive Eq. (21.14).

21.51. Taking carbide as an example and using Eq. (21.24), determine how much the feed should be reduced in order to keep the mean temperature constant when the cutting speed is doubled.

21.52. Using trigonometric relationships, derive an expression for the ratio of shear energy to frictional energy in orthogonal cutting, in terms of angles α, β, and ϕ only.

21.53. An orthogonal cutting operation is being carried out under the following conditions: $t_o = 0.2$ mm, $t_c = 0.3$ mm, width of cut $= 6$ mm $V = 5$ m/s rake angle $= 15°$, $F_c = 650$ N, and $F_t = 250$ N. Calculate the percentage of the total energy that is dissipated in the shear plane.

21.54. Explain how you would go about estimating the C and n values for the four tool materials shown in Fig. 21.17.

21.55. Derive Eqs. (21.1) and (21.3).

21.56. Assume that, in orthogonal cutting, the rake angle, α, is 20° and the friction angle, β, is 35° at the chip–tool interface. Determine the percentage change in chip thickness when the friction angle is 45°. [*Note:* do not use Eq. (21.4) or Eq. (21.5)].

21.57. Show that, for the same shear angle, there are two rake angles that give the same cutting ratio.

21.58. With appropriate diagrams, show how the use of a cutting fluid can change the magnitude of the thrust force, F_t, in Fig. 21.11. Consider both heat transfer and lubrication effects.

21.59. In a cutting operation using a $-5°$ rake angle, the measured forces were $F_c = 1330$ N and $F_t = 740$ N. When a cutting fluid was used, these forces were $F_c = 1200$ N and $F_t = 710$ N. What is the change in the friction angle resulting from the use of a cutting fluid?

21.60. For a turning operation using a ceramic cutting tool, if the speed is increased by 50%, by what factor must the feed rate be modified to obtain a constant tool life? Use $n = 0.5$ and $y = 0.6$.

21.61. In Example 21.3, if the cutting speed V is doubled, will the answer be different? Explain.

21.62. Using Eq. (21.30), select an appropriate feed for $R = 1.5$ mm and a desired roughness of 0.6 μm. How would you adjust this feed to allow for nose wear of the tool during extended cuts? Explain your reasoning.

21.63. With a carbide tool, the temperature in a cutting operation is measured as 650 K when the speed is 90 m/min and the feed is 0.05 mm/rev. What is the approximate temperature if the speed is doubled? What speed is required to lower the maximum cutting temperature to 480 K?

21.64. The following flank wear data were collected in a series of machining tests using C6 carbide tools on 1045 steel (HB=192). The feed rate was 0.38 mm/rev, and the width of cut was 0.75 mm. (a) Plot flank wear as a function of cutting time. Using a 0.38 mm wear land as the criterion of tool failure, determine the lives for the two cutting speeds. (b) Plot your results on log–log plot and determine the values of n and C in the Taylor tool life equation. (Assume a straight line relationship.) (c) Using these results, calculate the tool life for a cutting speed of 90 m/min.

Cutting speed V (m/min)	Cutting time (min)	Flank wear (mm)
120	0.5	0.035
	2.0	0.0575
	4.0	0.075
	8.0	0.1375
	16.0	0.205
	24.0	0.28
	54.0	0.375
180	0.5	0.045
	2.0	0.0875
	4.0	0.15
	8.0	0.25
	13.0	0.3625
	14	0.4
240	0.5	0.125
	2.0	0.25
	4.0	0.35
	5.0	0.4
300	0.5	0.25
	1.0	0.325
	1.8	0.375
	2.0	0.4

21.65. The following data are available from orthogonal cutting experiments. In both cases depth of cut (feed) $t_o = 0.13$ mm, width of cut $b = 2.5$ mm, rake angle $\alpha = -5°$, and cutting speed $V = 2$ m/s.

	Workpiece material	
	Aluminum	Steel
Chip thickness, t_c (mm)	0.23	0.58
Cutting force, F_c (N)	430	890
Thrust force, F_t (N)	280	800

Determine the shear angle ϕ, friction coefficient μ, shear stress τ and shear strain γ on the shear plane, chip velocity V_c and shear velocity V_s, as well as energies u_f, u_s and u_t.

21.66. Estimate the cutting temperatures for the conditions of Problem 21.65 if the following properties apply:

	Workpiece material	
	Aluminum	Steel
Cutting energy, u (N-mm/mm^3)	1320	2740
Thermal diffusivity, K (mm^2/s)	97	14
Volumetric specific heat, ρc (N/mm^{2}°C)	2.6	3.3

21.67. Assume that you are an instructor covering the topics described in this chapter and you are giving a quiz on the numerical aspects to test the understanding of the students. Prepare two quantitative problems and supply the answers.

Synthesis, Design, and Projects

21.68. Tool life is increased greatly when an effective means of cooling and lubrication is implemented. Design methods of delivering this fluid to the cutting zone and discuss the advantages and limitations of your design.

21.69. Design an experimental setup whereby orthogonal cutting can be simulated in a turning operation on a lathe.

21.70. Describe your thoughts on whether chips produced during machining can be used to make useful products. Give some examples of possible products, and comment on their characteristics and differences if the same products were made by other manufacturing processes. Which types of chips would be desirable for this purpose?

21.71. Recall that cutting tools can be designed so that the tool–chip contact length is reduced by recessing the rake face of the tool some distance away from its tip. Explain the possible advantages of such a tool.

21.72. Recall that the chip-formation mechanism also can be observed by scraping the surface of a stick of butter with a sharp knife. Using butter at different temperatures, including frozen butter, conduct such an experiment. Keep the depth of cut constant and hold the knife at different angles (to simulate the tool rake angle), including oblique scraping. Describe your observations regarding the type of chips produced. Also, comment on the force that your hand feels while scraping and whether you observe any chatter when the butter is very cold.

21.73. Experiments have shown that it is possible to produce thin, wide chips, such as 0.08 mm thick and 10 mm wide, which would be similar to the dimensions of a rolled sheet. Materials have been aluminum, magnesium, and stainless steel. A typical setup would be similar to orthogonal cutting, by machining the periphery of a solid round bar with a straight tool moving radially inward. Describe your thoughts regarding producing thin metal sheets by this method, taking into account the metal's surface characteristics and properties.

21.74. Describe your thoughts regarding the recycling of chips produced during machining in a plant. Consider chips produced by dry cutting versus those produced by machining with a cutting fluid.

21.75. List products that can be directly produced from metal chips or shavings.

21.76. Obtain a wood planer and some wood specimens. Show that the chips produced depend on the direction of cut with respect to the wood grain. Explain why.

21.77. It has been noted that the chips from certain carbon steels are noticeably magnetic, even if the original workpiece is not. Research the reasons for this effect and write a one-page paper explaining the important mechanisms.

21.78. As we have seen, chips carry away the majority of the heat generated during machining. If chips did not have this capacity, what suggestions would you make in order to be able to carry out machining processes without excessive heat? Explain.

21.79. A common practice is to set the cutting speed so that the tool life is the same as a work shift, commonly eight hours. This allows the standard practice of changing a cutting tool at the start of each shift. List the advantages and disadvantages of following this practice.

Chapter 22

Cutting-tool Materials and Cutting Fluids

- Continuing the coverage of the fundamentals of machining in the preceding chapter, this chapter describes two essential elements in machining operations: cutting-tool materials and cutting fluids.
- The chapter opens with a discussion of the types and characteristics of cutting-tool materials, including high-speed steels, carbides, ceramics, cubic boron nitride, diamond, and coated tools.
- The types of cutting fluids in common use are then described, including their functions and how they affect machining.
- Trends in near-dry and dry machining, and in methods for cutting fluid application, are also described, and their significance with respect to environmentally friendly machining operations is explained.

22.1 Introduction

The selection of a cutting-tool material for a specific application is among the most important factors to consider in machining operations. This chapter describes the properties and performance characteristics of all major types of tool materials as a guide to tool selection. General guidelines and recommendations have been established by industry based on experience. More specific information on recommendations are presented beginning with Chapter 23.

As described in the preceding chapter, a cutting tool is subjected to (a) high temperatures, (b) high forces and contact stresses, and (c) rubbing along the tool–chip interface and along the machined surface. Consequently, cutting-tool material must possess the following characteristics:

- **Hot hardness**, so that the hardness, strength, and wear resistance of the tool can be maintained at the temperatures encountered in machining. Hot hardness ensures that the tool does not undergo any deformation, thus retaining its shape and sharpness. As shown in Fig. 22.1, tool-material hardness is a function of temperature. Note, for example, how rapidly carbon tool steels (Section 5.7) lose their hardness and how well ceramics (Chapter 8) maintain their hardness at high temperatures.

- **Toughness** and **impact strength** (Section 2.9), so that forces on the tool encountered repeatedly in interrupted cutting operations (such as milling, Section 24.2) or forces due to vibration and chatter during machining do not chip or fracture the tool.

- **Thermal shock resistance**, to withstand the rapid temperature cycling (Section 3.6), as encountered in interrupted cutting.

- **Wear resistance** (Section 33.5), so that an acceptable tool life is maintained before tool replacement is necessary.

- **Chemical stability** and **inertness**, with respect to the workpiece material, to avoid or minimize any adverse reactions, adhesion, and tool–chip diffusion that would contribute to tool wear.

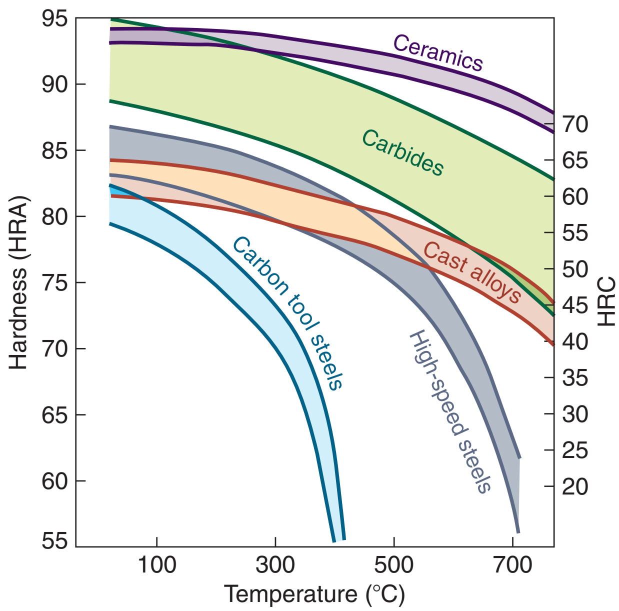


Figure 22.1: The hardness of various cutting-tool materials as a function of temperature. The wide range in each group of materials is due to the variety of tool compositions and treatments available for that particular group.

Table 22.1: General Characteristics of Tool Materials.

Property	High-speed steels	Cast-cobalt alloys	Carbides WC	Carbides TiC	Ceramics	Cubic boron nitride	Single-crystal diamond*
Hardness	83–86 HRA	82–84 HRA 46–62 HRC	90–95 HRA 1800–2400 HK	91–93 HRA 1800–3200 HK	91–95 HRA 2000–3000 HK	4000–5000 HK	7000–8000 HK
Compressive strength, MPa	4100–4500	1500–2300	4100–5850	3100–3850	2750–4500	6900	6900
Transverse rupture strength, MPa	2400–4800	1380–2050	1050–2600	1380–1900	345–950	700	1350
Impact strength, J	1.35–8	0.34–1.25	0.34–1.35	0.79–1.24	< 0.1	< 0.5	< 0.2
Modulus of elasticity, GPa	200	–	520–690	310–450	310–410	850	820–1050
Density, kg/m^3	8600	8000–8700	10,000–15,000	5500–5800	4000–4500	3500	3500
Volume of hard phase, %	7–15	10–20	70–90	–	100	95	95
Melting or decomposition temperature, °C	1300	–	1400	1400	2000	1300	700
Thermal conductivity, W/m K	30–50	–	42–125	17	29	13	500–2000
Coefficient of thermal expansion, $\times 10^{-6}/°C$	12	–	4–6.5	7.5–9	6–8.5	4.8	1.5–4.8

*The values for polycrystalline diamond are generally lower, except for impact strength, which is higher.

To respond to these demanding requirements, a variety of cutting-tool materials, with a wide range of mechanical, physical, and chemical properties, have been developed over the years (Table 22.1). The properties listed in the first column of the table are useful in determining desirable tool-material characteristics for a particular application. For example,

- Hardness and strength are important with respect to the mechanical properties of the workpiece material being machined.
- Impact strength facilitates making interrupted cuts, as in milling.
- Melting temperature of the tool material is important as compared to the temperatures developed in machining.
- Thermal conductivity and coefficient of thermal expansion indicate the resistance of the tool to thermal fatigue and shock.

A particular tool material may not have *all* of the desired properties for a particular machining operation, a situation that can readily be noted from Table 22.2, by observing the opposite directions of the long horizontal trendlines. Note, for example, that (a) high-speed steels are tough but have limited hot hardness, and (b) ceramics have high resistance to temperature and wear but they are brittle and thus can easily chip.

The operating characteristics of tool materials are shown in Table 22.3, listed in the order in which they were developed and implemented in industry. Many of these materials also are used for dies and molds in casting, forming, and shaping metallic and nonmetallic materials.

Table 22.2: General Characteristics of Cutting-tool Materials. These Materials Have a Wide Range of Compositions and Properties; Overlapping Characteristics Exist in Many Categories of Tool Materials.

	High-speed steels	Cast-cobalt alloys	Uncoated carbides	Coated carbides	Ceramics	Polycrystalline cubic boron nitride	Diamond
Hot hardness	⟶						
Toughness	⟵						
Impact strength	⟵						
Wear resistance	⟶						
Chipping resistance	⟵						
Cutting speed	⟵						
Thermal-shock resistance	⟵						
Tool material cost	⟶						
Depth of cut	Light	Light	Light	Light	Light	Light	Very light
Depth of cut	to	to	to	to	to	to	for single-crystal
Depth of cut	heavy	heavy	heavy	heavy	heavy	heavy	diamond
Processing method	Wrought,	Cast	Cold	CVD	Cold pressing	High-pressure,	High-pressure,
Processing method	cast,	and	pressing	or	and sintering	high-	high-
Processing method	HIP*	HIP	and	PVD**	or HIP	temperature	temperature
Processing method	sintering	sintering	sintering		sintering	sintering	sintering

Source: After R. Komanduri.

* Hot-isostatic pressing.

** Chemical-vapor deposition, physical-vapor deposition.

Table 22.3: General Operating Characteristics of Cutting-tool Materials.

Tool materials	General characteristics	Modes of tool wear or failure	Limitations
High-speed steels	High toughness, resistance to fracture, wide range of roughing and finishing cuts, good for interrupted cuts	Flank wear, crater wear	Low hot hardness, limited hardenability, and limited wear resistance
Uncoated carbides	High hardness over a wide range of temperatures, toughness, wear resistance, versatile, wide range of applications	Flank wear, crater wear	Cannot be used at low speeds because of cold welding of chips and microchipping
Coated carbides	Improved wear resistance over uncoated carbides, better frictional and thermal properties	Flank wear, crater wear	Cannot be used at low speeds because of cold welding of chips and microchipping
Ceramics	High hardness at elevated temperatures, high abrasive wear resistance	Depth-of-cut line notching, microchipping, gross fracture	Low strength and low thermomechanical fatigue strength
Polycrystalline cubic boron nitride (cBN)	High hot hardness, toughness, cutting-edge strength	Depth-of-cut line notching, chipping, oxidation, graphitization	Low strength, and lower chemical stability than ceramics at higher temperature
Diamond	High hardness and toughness, abrasive wear resistance	Chipping, oxidation, graphitization	Low strength, and low chemical stability at higher temperatures

Carbon steels are the oldest tool materials (Fig. 22.6), and have been used widely for drills, taps, broaches, and reamers. Low-alloy and medium-alloy steels were later developed for similar applications but with longer tool life. Although inexpensive and easily shaped and sharpened, these steels do not have sufficient hot hardness and wear resistance for machining at high speeds, where the temperature rises rapidly. Their use is limited to very low speed cutting operations, particularly in woodworking.

The following topics are described in this chapter:

- Characteristics, applications, and limitations of cutting-tool materials, and their costs.
- Applicable range of processing variables for optimal performance.
- Types and characteristics of cutting fluids and their specific applications in machining.

22.2 High-speed Steels

High-speed steel (HSS) tools are so named because they were developed to machine at speeds higher than was previously possible. Introduced in the early 1900s, these steels are the most highly alloyed of the tool steels (Section 5.7). They can be hardened to various depths, have good wear resistance, and are relatively inexpensive. Because of their toughness, hence high resistance to fracture, they are suitable especially for (a) high positive rake-angle tools (those with small included angles), (b) interrupted cuts, (c) machine tools with low stiffness thus subject to vibration and chatter, and (d) tools with complex geometries, such as drills, reamers, taps, and gear cutters. Their most important limitation is due to their lower hot hardness, whereby applicable cutting speeds are low as compared with those of carbide tools (see Fig. 22.1).

There are two basic types of high-speed steels: *molybdenum* (M-series) and *tungsten* (T-series). The M-series contains up to about 10% Mo, with Cr, V, W, and Co as alloying elements. The T-series contains 12 to 18% W, with Cr, V, and Co as alloying elements. Carbides in these steels constitute about 10 to 20% by volume. The M-series generally has higher abrasion resistance than the T-series, undergoes less distortion during heat treating (Section 4.7), and is less expensive. Consequently, 95% of all high-speed steel tools are made of the M-series steels. These steels and their characteristics are listed in Table 5.8.

High-speed steel tools are available in wrought (forged or rolled), cast, and powder-metallurgy (sintered) conditions; they are also available *coated*, for improved performance (Section 22.5). High-speed steel tools also may be subjected to *surface treatments*, such as case hardening for improved hardness and wear resistance (Section 4.10) or steam treatment at elevated temperatures to develop a hard, black oxide layer (*bluing*) for improved performance, including a lower tendency for built-up edge formation.

The major alloying elements in HSS are chromium, vanadium, tungsten, cobalt, and molybdenum. Their role in cutting tools may be summarized as follows (see also Table 5.2):

- *Chromium* improves toughness, wear resistance, and high-temperature strength.
- *Vanadium* improves toughness, abrasion resistance, and hot hardness.
- *Tungsten* and *cobalt* have similar effects, namely, improved strength and hot hardness.
- *Molybdenum* improves wear resistance, toughness, and high-temperature strength and hardness.

22.3 Cast-cobalt Alloys

Introduced in 1915, *cast-cobalt alloys* have the composition ranges of 38–53% Co, 30–33% Cr, and 10–20% W. Because of their high hardness, typically 58 to 64 HRC, they have good wear resistance and can maintain their hardness at elevated temperatures. They are not as tough as high-speed steels and are sensitive to impact forces; consequently, they are less suitable than high-speed steels for interrupted cutting operations. Commonly known as *Stellite* tools, they are cast and ground into relatively simple shapes.

22.4 Carbides

The two groups of tool materials just described possess the required toughness, impact strength, and thermal shock resistance, but they also have important limitations, particularly with respect to strength and hot hardness. Consequently, they cannot be used as effectively where high cutting speeds, hence high temperatures, are involved; such speeds often are necessary to improve productivity.

To meet the challenge for increasingly higher cutting speeds, *carbides*, also known as *cemented* or *sintered carbides*, were introduced in the 1930s. Because of their high hardness over a wide range of temperatures (Fig. 22.1), high elastic modulus, high thermal conductivity, and low thermal expansion, carbides are among the most important, versatile, and cost-effective tool and die materials for a wide range of applications. The two major groups of carbides are *tungsten carbide* and *titanium carbide*. In order to differentiate them from the coated tools, described in Section 22.5, plain-carbide tools are referred to as **uncoated carbides**.

22.4.1 Tungsten Carbide

Tungsten carbide (WC) typically consists of tungsten-carbide particles bonded together in a cobalt matrix. They are made using powder-metallurgy techniques (Chapter 17), hence the term *sintered carbides* or *cemented carbides*. Tungsten-carbide particles are first combined typically with cobalt, resulting in a *composite material* with a cobalt matrix surrounding the carbide particles. These particles, which are 1 to 5 μm in size, are then pressed and sintered into *inserts* (Section 22.4.3 and Fig. 22.2). Tungsten carbides frequently are also compounded with *titanium carbide* and *niobium carbide* to impart special properties to the material.

The amount of cobalt present, ranging typically from 6 to 16%, significantly affects the properties of tungsten-carbide tools. As the cobalt content increases, the strength, hardness, and wear resistance decrease,

Figure 22.2: Typical cutting tool inserts with various shapes and chip-breaker features: Round inserts also are available, as can be seen in Figs. 22.3c and 22.4. The holes in the inserts are standardized for interchangeability in toolholders. *Source:* Courtesy of Kennametal, Inc.

while toughness increases because of the higher toughness of cobalt. Tungsten-carbide tools are generally used for cutting steels, cast irons, and abrasive nonferrous materials.

Micrograin Carbides. Cutting tools also are made of submicron and ultra-fine-grained (*micrograin*) carbides, including tungsten carbide, titanium carbide, and tantalum carbide. Grain size is typically in the range from 0.2 to 0.8 μm. Compared with the traditional carbides, these tools are stronger, harder, and more wear resistant, thus improving productivity. In one application, microdrills, with diameters on the order of 100 μm, have been made from micrograin carbides, and used in the fabrication of microelectronic circuit boards (Section 28.13).

Functionally Graded Carbides. In these tools, the composition of the carbide in the insert has a *gradient* through its near-surface depth, instead of being uniform as it is in common carbide inserts. The gradient has a smooth distribution of compositions and phases, with functions similar to those described as desirable properties of coatings on cutting tools. Graded mechanical properties eliminate stress concentrations and increase tool life; however, they are more expensive and cannot be justified for all applications.

22.4.2 Titanium Carbide

Titanium carbide (TiC) has a nickel–molybdenum matrix; it has higher wear resistance than tungsten carbide but is not as tough. Titanium carbide is suitable for machining hard materials, mainly steels and cast irons, and for machining at speeds higher than those for tungsten carbide.

22.4.3 Inserts

Tool changing operations can be time consuming, reducing productivity. The need for a more effective method led to the development of *inserts*, which are individual cutting tools with several cutting edges (Fig. 22.2). A square insert has eight cutting edges, and a triangular insert has six. Inserts are typically clamped on the *toolholder*, using a variety of locking mechanisms (Fig. 22.3). When one edge of the insert is worn, it is **indexed** (rotated in its holder) to make another edge available. A wide variety of other toolholders is available for specific applications, including those with quick insertion and removal features.

The strength of the insert's cutting edge depends on its shape; the smaller the included angle (see top of Fig. 22.4), the lower is the strength of the edge. In order to further improve edge strength and prevent chipping, insert edges are usually honed, chamfered, or made with a negative land (Fig. 22.5).

Chip-breaker features (Fig. 21.7 and Section 21.2.1) on inserts are for the purposes of (a) controlling chip flow during machining, (b) eliminating long continuous chips, (c) reducing heat generated, and (d) reducing the tendency for vibration and chatter. Carbide inserts are available with a wide variety of complex

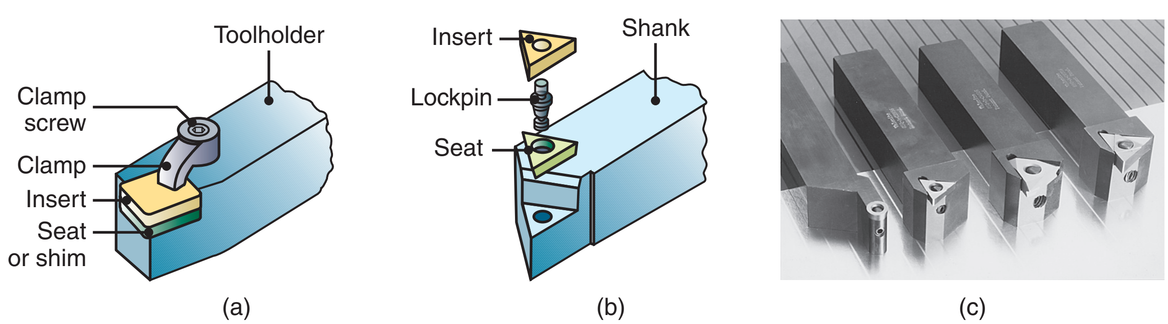


Figure 22.3: Methods of mounting inserts on toolholders: (a) clamping and (b) wing lockpins. (c) Examples of inserts mounted with threadless lockpins, which are secured with side screws. *Source:* Courtesy of Sandvik.

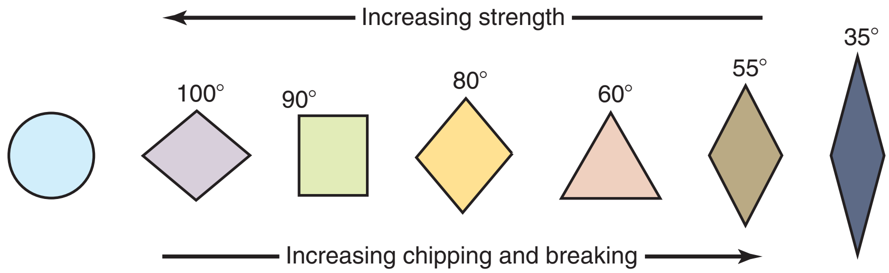


Figure 22.4: Relative edge strength and tendency for chipping of inserts with various shapes. Strength refers to the cutting edge indicated by the included angles. *Source:* Courtesy of Kennametal Inc.

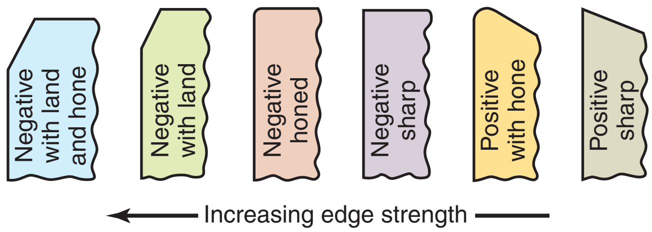


Figure 22.5: Edge preparation for inserts to improve edge strength. *Source:* Courtesy of Kennametal Inc.

chip-breaker features (Fig. 22.2). The selection of a particular chip-breaker feature depends on the feed and depth of cut of the operation, the workpiece material, the type of chip produced during machining, and whether it is a roughing or a finishing machining operation.

Stiffness of the machine tool (Section 25.3) is of major importance in using carbide tools; light feeds, low speeds, and chatter are detrimental because they tend to damage the tool's cutting edge. Light feeds, for example, concentrate the forces and temperature closer to the edge of the tool, increasing its tendency to chip off.

22.4.4 Classification of Carbides

Carbide tool grades are classified using the letters P, M, K, N, S, and H (Tables 22.4 and 22.5) for a range of applications, including the traditional C grades used in the United States.

Table 22.4: ISO Classification of Carbide Cutting Tools According to Use.

			Designation in increasing order of wear resistance and decreasing order of toughness in each category (in increments of 5)	
Symbol	Workpiece Material	Color code	Uncoated	Coated
P	Ferrous metals with long chips	Blue	P01, P05–P20	P20-P50
M	Stainless steels with long or short chips	Yellow	M10–M20	M20-M40
K	Cast iron with short chips	Red	K05–K20	K05-K30
N	Non-Ferrous metals	Green	N10-20	N05-N30
S	High-temperature alloys	Orange	S10-20	S20-S30
H	Hardened Materials	Gray	—	H10

Table 22.5: Classification of Tungsten Carbides According to Selected Machining Applications.

ISO standard	ANSI classification number (grade)	Materials to be machined	Machining operation	Type of carbide	Characteristics of Cut	Characteristics of Carbide
K30–K40	C1	Cast iron, nonferrous metals, and nonmetallic materials requiring abrasion resistance	Roughing	Wear-resistant grades; generally straight WC–Co with varying grain sizes	Increasing cutting speed ↓↑ Increasing feed rate	Increasing hardness and wear resistance ↓↑ Increasing strength and binder content
K20	C2		General purpose			
K10	C3		Light finishing			
K01	C4		Precision finishing			
P30–P50	C5	Steels requiring crater and deformation resistance	Roughing	Crater-resistant grades; various WC–Co compositions with TiC and/or TaC alloys	Increasing cutting speed ↓↑ Increasing feed rate	Increasing hardness and wear resistance ↓↑ Increasing strength and binder content
P20	C6		General purpose			
P10	C7		Light finishing			
P01	C8		Precision finishing			

Note: The ISO and ANSI comparisons are approximate.

22.5 Coated Tools

The difficulty of machining newly developed materials efficiently and the need for improving their performance has led to important developments in *coated tools*. Compared to the tool materials themselves, coatings have such properties as:

- Higher chemical inertness and resistance to wear and cracking
- Higher hot hardness and impact resistance
- Lower friction
- Acting as a diffusion barrier between the tool and the chip.

Coated tools can last 10 times more than those of uncoated tools, thus allowing for high cutting speeds and reducing both the time required for machining and production costs. As can be seen from Fig. 22.6, machining time has been reduced steadily by a factor of more than 100 since 1900, an improvement that has had a major impact on the economics of machining operations in manufacturing. This progress includes the continued improvements in the design and construction of modern machine tools and their computer controls (Chapter 25 and Part IX).

22.5.1 Coating Materials and Coating Methods

Common coating materials are *titanium nitride* (TiN), *titanium carbide* (TiC), *titanium carbonitride* (TiCN), and *aluminum oxide* (Al_2O_3). Typically in the thickness range from 2 to 15 μm, coatings are applied by two techniques, described in greater detail in Section 34.6:

1. **Chemical-vapor deposition** (CVD), including **plasma-assisted chemical-vapor deposition**
2. **Physical-vapor deposition** (PVD).

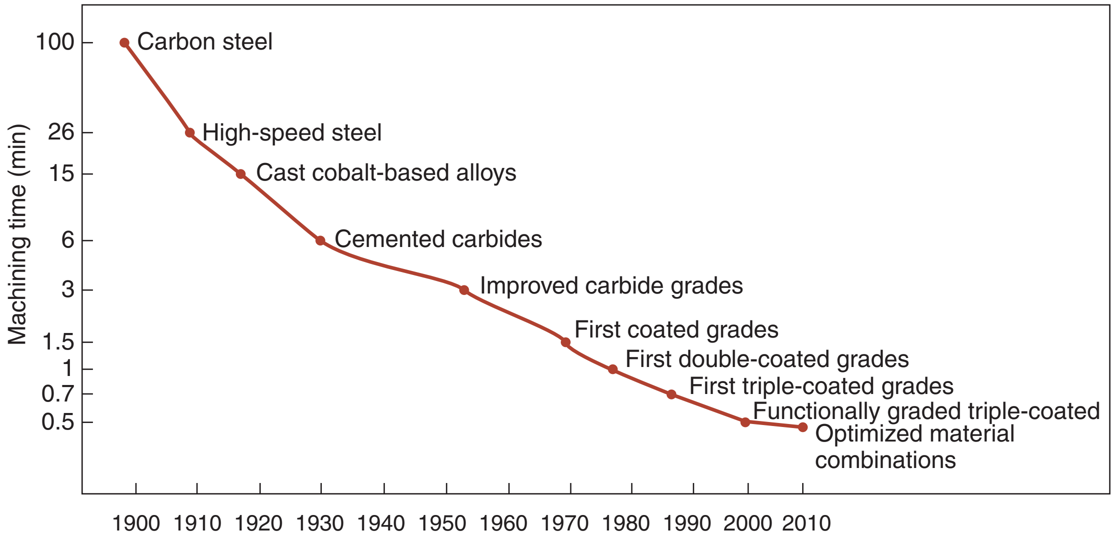


Figure 22.6: Relative time required to machine with various cutting-tool materials, indicating the year the tool materials were first introduced. Note that machining time has been reduced by two orders of magnitude within a hundred years. *Source:* Courtesy of Sandvik.

The CVD process is the most common method for carbide tools, with multiphase and ceramic coatings. However, the PVD-coated carbides with TiN coatings have higher cutting-edge strength, lower friction, and a lower tendency to form a built-up edge, and the coatings are smoother and more uniform in thickness, which is generally in the range from 2 to 4 μm. Another technology, used particularly for multiphase coatings, is *medium-temperature chemical-vapor deposition* (MTCVD), developed to machine ductile (nodular) iron and stainless steels and to provide higher resistance to crack propagation than CVD coatings provide.

Coatings should have the following general characteristics:

- **High hardness** at elevated temperatures, to resist wear.
- **Chemical stability** and **inertness** to the workpiece material, in order to reduce wear.
- **Low thermal conductivity**, to prevent temperature rise in the substrate.
- **Compatibility** and **good bonding**, to prevent flaking or spalling from the substrate, which may be carbide or high-speed steel.
- **Little or no porosity**, to maintain integrity and strength.

Coating effectiveness is enhanced by the hardness, toughness, and thermal conductivity of the substrate. Honing (Section 26.7) of the cutting edges is an important procedure for maintaining coating strength, as otherwise the coating may peel off or chip at sharp edges and corners.

Titanium-nitride. Titanium-nitride coatings have low friction, high hardness, good high temperature resistance, and good adhesion to the substrate. They greatly improve the life of high-speed steel and carbide tools, carbide tools, drill bits, and cutters. Titanium-nitride-coated tools (gold in color), perform well at higher cutting speeds and feeds, and flank wear is significantly lower than that of uncoated tools (Fig. 22.7). Flank surfaces can be reground after use, so long as regrinding that face does not remove the coating on the rake face of the tool. Coated tools do not perform as well at low cutting speeds, because the coating can be worn off by chip adhesion, thus the use of appropriate cutting fluids, to minimize adhesion, is important.

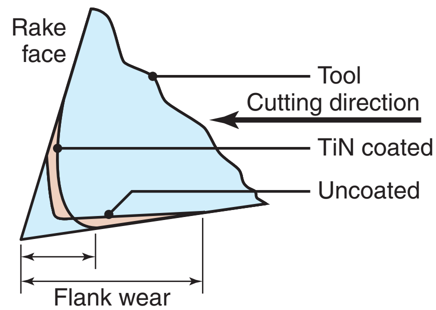


Figure 22.7: Schematic illustration of typical wear patterns on uncoated high-speed steel tools and titanium-nitride-coated tools. Note that flank wear is significantly lower for the coated tool.

Titanium-carbide Coatings. Titanium-carbide coatings on tungsten-carbide inserts have high flank-wear resistance in machining abrasive materials.

Ceramic Coatings. Because of their chemical inertness, low thermal conductivity, resistance to high temperature, and resistance to flank and crater wear, ceramics are good coating materials for cutting tools. The most commonly used ceramic coating is *aluminum oxide* (Al_2O_3). However, because they are very stable (e.g., not chemically reactive), oxide coatings generally do not bond well to the substrate.

Multiphase Coatings. The desirable properties of coatings can be combined and optimized using *multiphase coatings*. Carbide tools are available with two or three layers of such coatings, and are particularly effective in machining cast irons and steels. For example, TiC can be deposited first over the substrate, followed by Al_2O_3, and then TiN; the first layer must bond well with the substrate, the outer layer should resist wear and have low thermal conductivity, and the intermediate layer should bond well and be compatible with both layers.

Typical applications of multiple-coated tools are:

- High-speed, continuous cutting: TiC/Al_2O_3
- Heavy-duty, continuous cutting: TiC/Al_2O_3/TiN
- Light, interrupted cutting: TiC/TiC + TiN/TiN.

Coatings can also be deposited in *alternating multiphase layers*; their thickness is on the order of 2 to 10 μm, this is thinner than regular multiphase coatings (Fig. 22.8). The reason for using thinner coatings is that coating hardness increases with decreasing grain size, a phenomenon similar to the increase in the strength of metals with decreasing grain size (see Section 1.5.1); thus, thinner layers are harder than thicker layers.

A typical multiphase-coated carbide tool may consist of the following layers, starting from the top:

1. TiN: low friction
2. Al_2O_3: high thermal stability
3. TiCN: fiber reinforced, with a good balance of resistance to flank wear and crater wear, effective particularly for interrupted cutting
4. A thin carbide substrate: high fracture toughness
5. A thick carbide substrate: hard and resistant to plastic deformation at high temperatures.

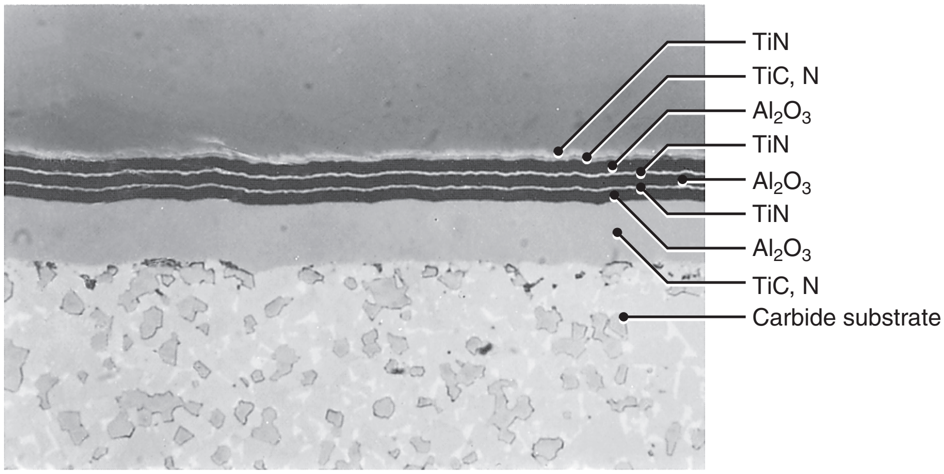


Figure 22.8: Multiphase coatings on a tungsten-carbide substrate. Three alternating layers of aluminum oxide are separated by very thin layers of titanium nitride. Inserts with as many as 13 layers of coatings have been made. Coating thicknesses are typically in the range from 2 to 10 μm. *Source:* Courtesy of Kennametal Inc.

Diamond Coatings. The properties and applications of diamond, diamond coatings, and *diamondlike carbon* are described in Sections 8.7 and 34.13, and the use of these materials as cutting tools is given in Section 22.9. *Polycrystalline diamond* (PCD) is used widely as a coating material for tools, particularly on tungsten-carbide and silicon-nitride inserts. Diamond-coated tools are particularly effective in machining (a) nonferrous metals, (b) abrasive materials, such as aluminum alloys containing silicon, (c) fiber-reinforced and metal-matrix composite materials, and (d) graphite. As many as tenfold improvements in tool life have been obtained over the lives of other coated tools.

Diamond-coated inserts have thin films deposited on substrates through PVD or CVD techniques (Section 34.6). Thick diamond films are produced by growing a large sheet of pure diamond, which is then laser cut to shape and brazed to a carbide insert. *Multilayer nanocrystal diamond coatings* have interlocking layers of diamond that give strength to the coating. As with all coatings, it is essential that the diamond film adheres well to the substrate and to minimize the difference in thermal expansion between the diamond and substrate materials selected (see Section 3.6).

22.5.2 Miscellaneous Coating Materials

The hardness of some of the following coatings approaches that of cubic boron nitride (Fig. 2.15).

1. **Titanium carbonitride** (TiCN) and *titanium-aluminum nitride* (TiAlN) are effective in machining stainless steels. TiCN (deposited by physical-vapor deposition) is harder and tougher than TiN, and can be used on carbides and high-speed steel tools. TiAlN is effective in machining aerospace alloys.

2. Chromium-based coatings, such as *chromium carbide* (CrC), have been found to be effective in machining softer metals that have a tendency to adhere to the cutting tool, such as aluminum, copper, and titanium. Other coating materials include **zirconium nitride** (ZrN) and hafnium nitride (HfN).

3. **Nanolayer coatings**, such as carbide, boride, nitride, oxide, or some combination of these materials (see also Section 8.8).

4. **Composite coatings**, using a variety of materials.

22.5.3 Ion Implantation

In this process, ions are introduced into the surface of the tool, improving its surface properties without affecting the tool's dimensions (Section 34.7). *Nitrogen-ion* implanted carbide tools have been used successfully on alloy steels and stainless steels. *Xenon-ion* implantation of tools is also under development.

22.6 Alumina-based Ceramics

Ceramic tool materials, introduced in the early 1950s, consist primarily of fine-grained, high-purity **aluminum oxide** (Section 8.2). They are cold pressed into insert shapes under high pressure, then sintered at high temperature. The end product is referred to as *white* (*cold-pressed*) *ceramics*. Additions of titanium carbide and zirconium oxide help improve properties, such as toughness and thermal-shock resistance.

Alumina-based ceramic tools have very high abrasion resistance and hot hardness (Fig. 22.9). Chemically, they are more stable than high-speed steels and carbides; they have less tendency to adhere to metals during machining, and thus a correspondingly lower tendency to form a built-up edge. Consequently, in machining cast irons and steels, good surface finish is obtained using ceramic tools. On the other hand, recall that ceramics generally lack toughness.

Ceramic inserts are available in shapes similar to those for carbide inserts (Section 22.4.3). They are effective in high-speed, uninterrupted cutting operations. To reduce thermal shock, cutting should be performed either dry or with a copious amount of cutting fluid, applied in a steady stream (Section 22.12). Improper or intermittent application of the fluid can cause thermal shock, possibly leading to fracture of the tool.

The shape of ceramic inserts and their setup are important. Negative rake angles (i.e., large included angles) generally are to be preferred to avoid chipping, due to the poor tensile strength of ceramics. Tool failure can be reduced by increasing the stiffness and damping capacity of machine tools, mountings, and workholding devices, thus reducing vibration and chatter.

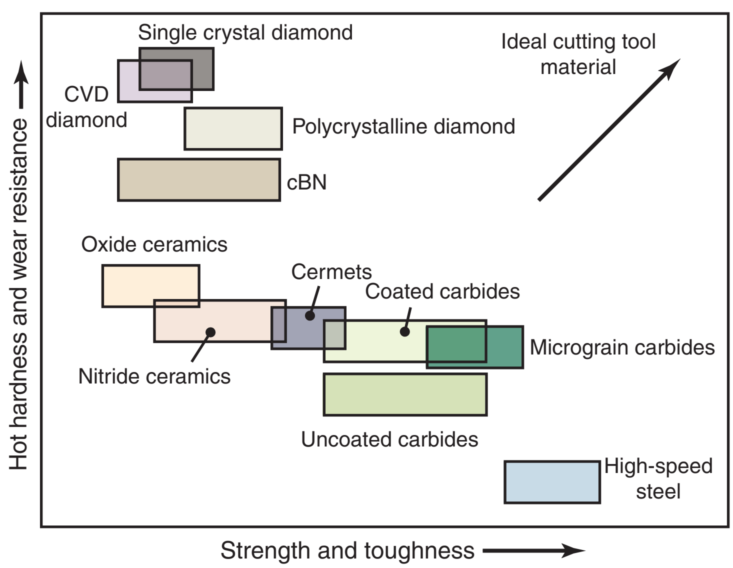


Figure 22.9: Ranges of mechanical properties for various groups of tool materials (see also Tables 22.1 through 22.5).

Cermets. *Cermets* (from the words *cer*amic and *met*al), first used in the early 1950s, consist of ceramic particles in a metallic matrix. They are referred to as *black* or *hot-pressed ceramics* (carboxides). A typical cermet consists of 70% aluminum oxide and 30% titanium carbide; other cermets contain molybdenum carbide, niobium carbide, and tantalum carbide. Although they have chemical stability and resistance to built-up edge formation, the brittleness and high cost of cermets have been a limitation to their wider use. Their performance is somewhere between that of ceramics and carbides, and has been particularly suitable for light roughing cuts and high-speed finishing cuts. Chip-breaker features are important for cermet inserts.

22.7 Cubic Boron Nitride

Next to diamond, *cubic boron nitride* (cBN) is the hardest material available. Introduced in 1962 under the trade name *Borazon*, cubic boron nitride is made by bonding a 0.5 to 1 mm layer of **polycrystalline cubic boron nitride** to a carbide substrate, by sintering under high pressure and high temperature. While the carbide provides shock resistance, the cBN layer provides very high wear resistance and cutting-edge strength (Fig. 22.10).

The thermochemical stability of cBN is a significant advantage; it can be used safely up to 1200°C. At elevated temperatures, cBN maintains high chemically inertness to iron and nickel, hence there is no wear due to diffusion. Its resistance to oxidation is high, making it particularly suitable for machining hardened ferrous and high-temperature alloys (see *hard machining*, Section 25.6) and for high-speed machining operations (Section 25.5). cBN also is used as an abrasive; however, because these tools are brittle, the stiffness of the machine tool and the fixturing is important in order to avoid vibration and chatter. Furthermore, in order to avoid chipping and cracking due to thermal shock, machining generally should be performed dry, particularly in interrupted cutting operations.

22.8 Silicon-Nitride-based Ceramics

Developed in the 1970s, *silicon-nitride* (SiN)-*based ceramic* materials consist of silicon nitride, with various additions of aluminum oxide, yttrium oxide, and titanium carbide. These tools have high toughness, hot hardness, and good thermal-shock resistance. An example of a SiN-based material is **sialon**, named after the elements *si*licon, *al*uminum, *o*xygen, and *n*itrogen. Sialon has higher thermal-shock resistance than silicon nitride; it is recommended for machining cast irons and nickel-based superalloys at intermediate cutting speeds. Because of their chemical affinity to iron at elevated temperatures, however, SiN-based tools are not suitable for machining steels.

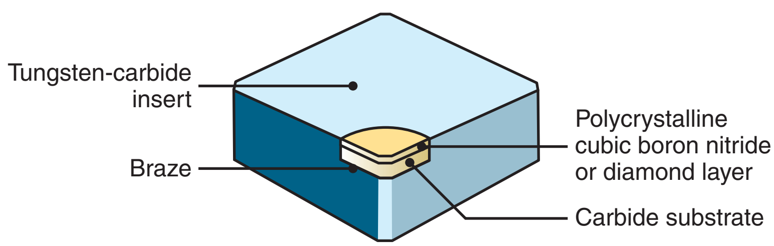
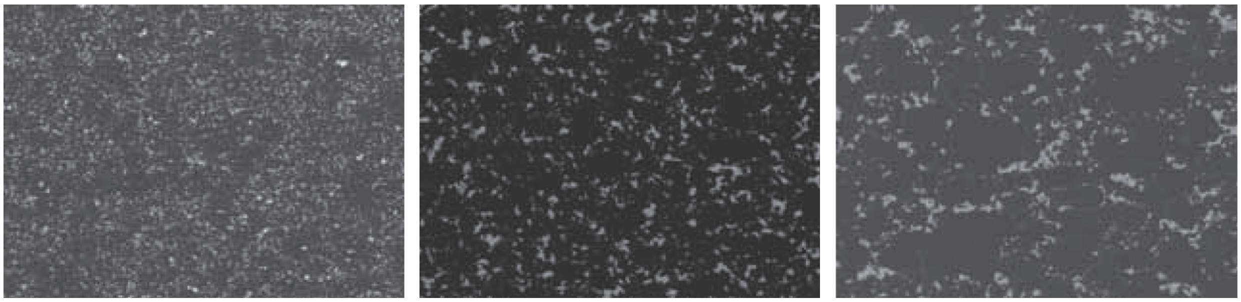


Figure 22.10: An insert of a polycrystalline cubic boron nitride or a diamond layer on tungsten carbide.



(a) (b) (c)

Figure 22.11: Microphotographs of diamond compacts: (a) fine-grained diamond, with mean grain size around 2 μm; (b) medium grain, with mean grain size around 10 μm; (c) coarse grain, with grain size around 25 μm. *Source:* Courtesy of Kennametal, Inc.

22.9 Diamond

Described in Section 8.7, diamond, as a cutting tool, has highly desirable properties such as low friction, high wear resistance, and the ability to maintain a sharp cutting edge. Diamond is used when good surface finish and dimensional accuracy are required, particularly when machining soft nonferrous alloys and abrasive nonmetallic and metallic materials, especially some aluminum–silicon alloys. **Synthetic** or **industrial diamond** is widely used because natural diamond has flaws, and thus its performance can be unpredictable.

Although *single-crystal diamond* of various carats (1 carat=200 mg) can be used for special applications, they have been replaced largely by **polycrystalline diamond** (PCD) tools, called **compacts** (also used as dies for fine wire drawing; see Section 15.9). These diamond tools consist of very small synthetic crystals (Fig. 22.11), fused under a high-pressure, high-temperature process, to a thickness of about 0.5 to 1 mm, and bonded to a carbide substrate; this product is similar to cBN tools (Fig. 22.10). A unique feature is that the random orientation of the diamond crystals of the structure prevents the propagation of cracks, thus significantly improving its toughness (see also Section 2.10.2). Fine grains are used when a high cutting-edge quality and higher strength are required; coarse grains are preferred for increased abrasion resistance.

Because diamond is brittle, tool shape and its sharpness are important. Low rake angles are generally used to provide a strong cutting edge, because of the larger included angles. Proper mounting and crystal orientation are important for optimum tool life. Wear may occur through microchipping (caused by thermal stresses and oxidation) and through transformation to carbon (caused by the heat generated during machining). **Diamond tools** can be used satisfactorily at almost any speed, but are most suitable for light, uninterrupted finishing cuts. To minimize tool fracture, the single-crystal diamond must be resharpened as soon as it becomes dull. Diamond is not recommended for machining plain-carbon steels or for titanium, nickel, and cobalt-based alloys, because of its strong chemical affinity to carbon at elevated temperatures (resulting in diffusion)).

22.10 Whisker-reinforced Materials and Nanomaterials

To further improve the performance and wear resistance of cutting tools, continued progress is being made in developing new tool materials, with enhanced properties, such as:

- High fracture toughness
- Resistance to thermal shock

- Cutting-edge strength
- Creep resistance
- Hot hardness.

Whiskers are important as reinforcing fibers in composite tool materials. Examples of *whisker-reinforced cutting tools* include (a) silicon-nitride-based tools reinforced with silicon-carbide whiskers and (b) aluminum-oxide-based tools reinforced with 25% to 40% silicon-carbide whiskers, sometimes with the addition of *zirconium oxide* (ZrO_2). Silicon-carbide whiskers are typically 5 to 100 μm long and 0.1 to 1 μm in diameter. The high reactivity of silicon carbide with ferrous metals, however, makes SiC-reinforced tools unsuitable for machining irons and steels.

Nanomaterials are also becoming important in advanced cutting-tool materials (see Section 8.8). Suitable nanomaterials are carbides and ceramics. Often, they are applied as a thin coating, usually in an attempt to obtain a reasonable tool life without requiring a coolant (see *dry machining*, Section 22.12.1).

22.11 Tool Costs and Reconditioning of Tools

Tool costs vary widely, depending on the tool material, size, shape, chip-breaker features, and quality. The approximate cost for a typical 12.5-mm *insert* is (a) $10 to $15 for uncoated carbides, (b) $10 to $25 for coated carbides, (c) $30 to $50 for ceramics, (d) $10 to $90 for diamond-coated carbides, (e) $75 to $225 for cubic boron nitride, and (f) $150 to $200 for a diamond-tipped insert.

After reviewing the costs involved in machining and considering all of the aspects involved in the total operation, it can be seen that the cost of an individual insert is relatively insignificant. Tooling costs in machining have been estimated to be on the order of 2% to 4% of the manufacturing costs. This small amount is due to the fact that a single insert typically can perform a large amount of material removal before it is indexed to use all its cutting edges, and eventually recycled. Note from Section 21.5 that the expected tool life can be in the range of 30–60 minutes. Considering that a square insert, for example, has eight cutting edges, a tool can last a long time before it is removed from the machine tool and replaced.

Cutting tools can be **reconditioned** by resharpening them, using tool and cutter grinders with special fixtures (Section 26.4). This operation may be carried out by hand or on computer-controlled grinders. Advanced methods of shaping cutting tools also are available, as described in Chapter 27. Reconditioning of coated tools is done by *recoating* them, usually in special facilities available for these purposes. It is important to ensure that reconditioned tools have the same geometric features as the original. Often, a decision has to be made whether further **reconditioning of tools** is economically viable. *Recycling* of tools is always a significant consideration, especially if they contain expensive and strategically important materials, such as tungsten and cobalt.

22.12 Cutting Fluids

Cutting fluids are used in machining operations for the following purposes:

- Reduce friction and wear, thus improving the tool life and surface finish of the workpiece.
- Cool the cutting zone, thus improving tool life and reducing the temperature and thermal distortion of the workpiece.
- Reduce cutting forces and energy consumption.
- Flush away the chips from the cutting zone, preventing the chips from interfering with the cutting operation, particularly in drilling and tapping.
- Protect the machined surface from environmental corrosion.

Depending on the type of machining operation, the cutting fluid required may be a **coolant**, a **lubricant**, or both. The effectiveness of fluids depends on several factors, such as the type of machining operation, tool and workpiece materials, cutting speed, and the method of application. Water is an excellent coolant, and can effectively reduce the high temperatures developed in the cutting zone. However, it is not an effective lubricant and it does not reduce friction, and can cause corrosion of workpieces and machine-tool components.

The necessity for a cutting fluid depends on the *severity* of the particular machining operation, defined as (a) the level of temperatures and forces encountered and the ability of the tool materials to withstand them, (b) the tendency for built-up edge formation, (c) the ease with which chips produced can be removed from the cutting zone, and (d) how effectively the fluids can be supplied to the proper region at the tool–chip interface. The relative severities of specific machining processes, in increasing order of severity, are: sawing, turning, milling, drilling, gear cutting, thread cutting, tapping, and internal broaching.

There are operations, however, in which the cooling action of cutting fluids can be detrimental. Cutting fluids may cause the chip to become *more curly* (see Fig. 21.9c), and thus concentrate the heat closer to the tool tip, reducing tool life. In interrupted cutting operations, such as milling with multiple-tooth cutters, cooling of the cutting zone leads to thermal cycling of the cutter teeth, which can cause *thermal cracks* by the mechanisms of thermal fatigue or thermal shock.

Cutting-fluid Action. The basic mechanisms of lubrication in metalworking operations are described in Section 33.6. Studies have shown that the cutting fluid gains access to the tool–chip interface by seeping from the *sides* of the chip (perpendicular to the page in Figs. 21.11 and 21.12), through the *capillary action* of the interlocking network of surface asperities in the interface.

Because of the small size of this capillary network, the cutting fluid should have a *small molecular size* and possess *wetting* (*surface tension*) characteristics. Grease, for example, cannot be an effective lubricant in machining, whereas low-molecular-weight oils suspended in water, known as *emulsions*, are very effective. Note also that in discontinuous machining operations, cutting fluids have more access to tool–chip-workpiece interfaces, although the tools become more susceptible to thermal shock.

Example 22.1 Effects of Cutting Fluids on Machining

Given: A machining operation is being carried out with a cutting fluid that is an effective lubricant.

Find: Describe the changes in the cutting operation mechanics if the fluid supply is interrupted.

Solution: Since the cutting fluid is a good lubricant, the following chain of events will take place after the fluid is shut off:

1. Friction at the tool–chip interface will increase.
2. The shear angle will decrease, in accordance with Eq. (21.3).
3. The shear strain will increase, as seen from Eq. (21.2).
4. The chip will become thicker.
5. A built-up edge is likely to form.

As a result of these changes, the following events will occur:

1. The shear energy in the primary zone will increase.
2. The frictional energy in the secondary zone will increase.
3. The total energy will increase.

4. The temperature in the cutting zone will rise, causing greater tool wear.
5. Surface finish of the workpiece will begin to deteriorate, and dimensional accuracy may be difficult to maintain, because of the increased temperature and thermal expansion of the workpiece during machining.

Types of Cutting Fluids. Four general types of cutting fluids are commonly used in machining operations:

1. **Oils** include mineral, animal, vegetable, compounded, and, more recently, synthetic oils. They typically are used for low-speed operations where temperature rise is not significant.
2. **Emulsions**, also called *soluble oils*, are a mixture of oil, water, and additives. They generally are used for high-speed machining operations where the temperature rise is significant. The presence of water makes emulsions highly effective coolants, and the presence of oil reduces or eliminates the tendency of water to cause oxidation of workpiece surfaces.
3. **Semisynthetics** are chemical emulsions containing some mineral oil diluted in water, and additives that reduce the size of the oil particles, thus making them more effective.
4. **Synthetics** are chemicals with additives, diluted in water; they contain no oil.

Because of the complex interactions among the cutting fluid, workpiece materials, temperature, and processing variables, the selection and application of fluids cannot be generalized. Recommendations for cutting fluids for various specific machining operations are given in Chapters 23 and 24.

Methods of Cutting-fluid Application. There are four basic methods of cutting-fluid applications in machining:

1. **Flooding.** This is the most common method, as shown in Fig. 22.12 and indicating good and poor flooding practices. Fluid flow rates typically range from 10 L/min for single-point tools to 225 L/min per cutter for multiple-tooth cutters, as in milling. In some operations, such as drilling and milling, fluid pressures in the range from 700 to 14,000 kPa are used to flush away the chips produced to prevent their interfering with the operation.
2. **Mist.** This type of cooling supplies fluid to inaccessible areas, in a manner similar to using an aerosol can, and provides better visibility of the workpiece being machined. This method is particularly effective with water-based fluids and at air pressures ranging from 70 to 600 kPa. However, it has limited cooling capacity, and requires venting to prevent the inhalation of airborne fluid particles by the operator and other personnel nearby.
3. **High-pressure systems.** Heat generation in machining can be a significant factor. Particularly effective is the use of high-pressure *refrigerated coolant systems* to increase the rate of heat removal. High pressures are also used to deliver the cutting fluid via specially designed nozzles; they aim a powerful jet of fluid to the cutting zone, particularly into the *clearance* or *relief face* of the tool (see Fig. 21.3). The pressures are usually in the range from 5.5 to 35 MPa, and also act as a chip breaker in situations where the chips produced would otherwise be long and continuous, interfering with the cutting operation. Proper cycling and continuous filtering of the fluid is essential to maintain workpiece surface quality.

   A design which achieves good performance with lower pressure required is shown in Fig. 22.13. The method has been found to be especially effective in machining titanium and other difficult-to-machine materials, with tool life increases over 300%. Instead of applying coolant to the workpiece surface or chip at a distance remote from the cutting zone, the coolant is applied on the *side* of the insert, whereby the temperature rise in the tool and chip can be reduced significantly (Fig. 22.13b).

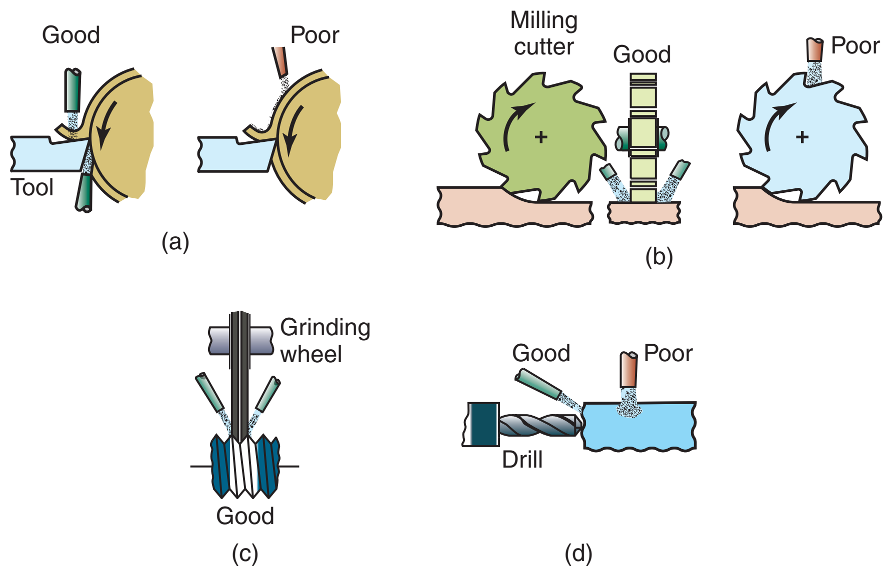


Figure 22.12: Schematic illustration of the proper methods of applying cutting fluids (flooding) in various machining operations: (a) turning, (b) milling, (c) thread grinding, and (d) drilling.

4. **Through the cutting-tool system.** For a more effective application, narrow passages can be produced in cutting tools and in toolholders, through which cutting fluids can be supplied under high pressure. Two applications of this method are (a) gun drilling (see Fig. 23.22; note the long, small hole through the body of the drill itself) and (b) boring bars (Fig. 23.18a; note the long hole through the shank (toolholder)), to which an insert is clamped. Similar designs have been developed, including those that deliver the cutting fluid through the spindle of the machine tool.

Effects of Cutting Fluids. The selection of a cutting fluid should also include considerations such as its effects on:

- Workpiece material
- Machine tool components
- Health considerations
- The environment.

In selecting an appropriate cutting fluid, one should consider the following factors:

1. Fluids containing sulfur should not be used with nickel-based alloys
2. Fluids containing chlorine should not be used with titanium, because of increased corrosion
3. Machined parts should be cleaned, when necessary, to remove any fluid residue (Section 34.16)
4. Cutting fluids may adversely affect the machine tool components; their compatibility with various metallic and nonmetallic materials in the machine also must be considered.

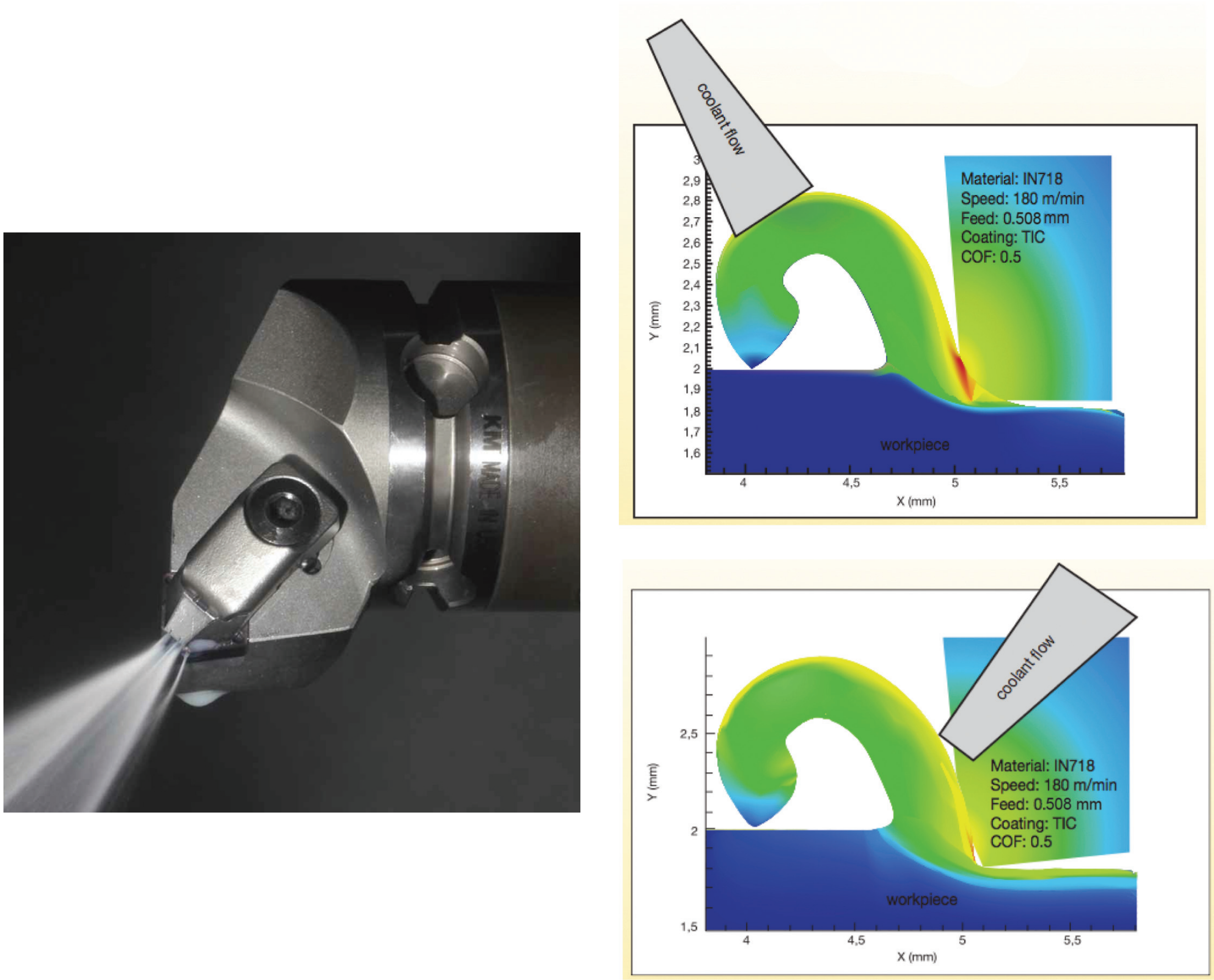


Figure 22.13: (a) A turning insert with coolant applied through the tool; (b) comparison of temperature distributions for conventional and through-the-tool application. *Source:* Courtesy of Kennametal, Inc.

The *health effects* of machine operators in contact with fluids also should be of primary concern. Mist, fumes, smoke, and odors from cutting fluids can cause severe skin reactions and respiratory problems, especially in using fluids with such chemical constituents as sulfur, chlorine, phosphorus, hydrocarbons, biocides, and various additives. The use of dry or near-dry machining techniques (see the following section), as well as in the design of machine tools with enclosed working areas (see Fig. 25.2) should be considered.

Cutting fluids may undergo chemical changes as they are used and recycled over time. These changes may be due to environmental effects or to contamination from various sources, including metal chips, fine particles, and *tramp oil* (oils that are from leaks in hydraulic systems, the sliding members of machine tools, and from lubricating systems for the machines). Several techniques, such as settling, skimming, centrifuging, and filtering, are used for clarifying used cutting fluids.

22.12.1 Near-dry and Dry Machining

For economic and environmental reasons, there has been a continuing worldwide trend to minimize or eliminate the use of metalworking fluids since the mid-1990s. This trend has led to the practice of *near-dry machining* (NDM), with significant benefits such as:

- Alleviating the environmental impact of using cutting fluids, improving air quality in manufacturing plants, and reducing health hazards.

- Reducing the cost of machining operations, including the cost of maintenance, recycling, and disposal of cutting fluids.

The significance of this approach becomes apparent when one notes that, in the United States alone, millions of gallons of metalworking fluids are consumed each year. Furthermore, it has been estimated that metalworking fluids constitute about 7% to 17% of the total machining costs.

The principle behind near-dry cutting is the application of a fine mist of an air–fluid mixture containing a very small amount of cutting fluid, which may be reformulated to contain vegetable oil. The mixture is delivered to the cutting zone through the spindle of the machine tool, typically through a 1-mm-diameter nozzle and under a pressure of 600 kPa. It is used at rates on the order of 1 to 100 cc/hour, which is estimated to be, at most, one ten-thousandth of that used in flood cooling. Consequently, the process is also known as *minimum-quantity lubrication* (MQL).

Cryogenic Machining. More recent developments in machining include the use of cryogenic gases, such as *nitrogen* or *carbon dioxide*, as a coolant. With small-diameter nozzles and at a temperature of –200°C, liquid nitrogen can be injected into the cutting zone. Because of the reduced temperature, tool hardness is maintained and hence tool life is improved, thus allowing for higher cutting speeds. The chips are also less ductile, thus machinability is increased. There is no adverse environmental impact, and the nitrogen simply evaporates.

Dry machining also is a viable alternative. With major advances in cutting tools, dry machining has been shown to be effective in various machining operations, especially turning, milling, and gear cutting, on steels, steel alloys, and cast irons, although generally not for aluminum alloys.

One of the functions of a metal-cutting fluid is to *flush* chips from the cutting zone. Although this function appears to be a challenge with dry machining, tool designs have been developed that allow the application of *pressurized air*, often through the tool shank (see Fig. 22.3b). Some gases, such as carbon dioxide, can also have a boundary lubrication benefit.

Summary

- Cutting tool materials have a wide range of mechanical and physical properties, such as hot hardness, toughness, chemical stability and inertness, and resistance to chipping and wear. A variety of cutting-tool materials are now available, the most commonly used being high-speed steels, carbides, ceramics, cubic boron nitride, and diamond.

- Several tool coatings have been developed, resulting in major improvements in tool life, surface finish, and the economics of machining operations. Common coating materials are titanium nitride, titanium carbide, titanium carbonitride, and aluminum oxide. The trend is toward multiphase coatings for even better performance.

- The selection of appropriate tool materials depends not only on the material to be machined, but also on processing parameters and the characteristics of the machine tool.

- Cutting fluids are important in machining operations, as they reduce friction, wear, cutting forces, and power requirements. Generally, slower cutting operations and those with high tool pressures require a fluid with good lubricating characteristics. In high-speed operations, where the temperature rise can be significant, fluids with good cooling capacity and some lubricity are required. The selection of cutting fluids must take into account their possible adverse effects on the machined parts, on machine tools and their components, on personnel, and on the environment.

Key Terms

Alumina-based ceramics
aluminum oxide
Carbides
Cast-cobalt alloys
Ceramic
Cermets
Chemical stability
Chip-breaker
Coated tools
Coolants
Cryogenic machining
Cubic boron nitride
Cutting fluids
Diamond coatings
Diamond tools
Dry machining
Flooding
High-speed steels
Inserts
Lubricants
Micrograin carbides
Mist
Multiphase coatings
Near-dry machining
Polycrystalline cubic boron nitride
Polycrystalline diamond
Reconditioning of tools
Sialon
Silicon-nitride-based ceramics
Stellite
Titanium carbide
Titanium nitride
Tool costs
Toughness
Tungsten carbide
Uncoated carbides
Wear resistance
Whiskers

Bibliography

ASM Specialty Handbook: **Tool Materials**, ASM International, 1995.
Astakhov, V.P., **Tribology of Metal Cutting**, Elsevier, 2007.
Astakhov, V.P., and Joksch, S., **Metalworking Fluids for Cutting and Grinding: Fundamentals and Recent Advances**, Woodhead, 2012.
Bryson, W.E., **Heat Treatment, Selection, and Application of Tool Steels**, 2nd ed., Hanser, 2005.
Byers, J.P. (ed.), **Metalworking Fluids**, 2nd ed., CRC Press, 2006.
Jackson, M.J., and Hitchiner, M.P., **High Performance Grinding and Advanced Cutting Tools**, Springer, 2013.
Davim, J.P., (ed.), **Sustainable Machining**, Springer, 2017.

Mesquita, R.A., **Tool Steels: Properties and Performance**, CRC Press, 2016.
Nachtman, E.S., and Kalpakjian, S., **Lubricants and Lubrication in Metalworking Operations**, Marcel Dekker, 1985.
Roberts, G.A., Krauss, G., and Kennedy, R., **Tool Steels**, 5th ed., ASM International, 1997.
Shaw, M.C., **Metal Cutting Principles**, 2nd ed., Oxford, 2005.
Smith, G.T., **Cutting Tool Technology: Industrial Handbook**, Springer, 2008.
Trent, E.M., and Wright, P.K., **Metal Cutting**, 4th ed., Butterworth-Heinemann, 2000.

Review Questions

22.1. What are the major properties required of cutting-tool materials? Why?

22.2. What is the composition of a typical carbide tool?

22.3. Why were cutting-tool inserts developed?

22.4. Why are some tools coated? What are the common coating materials?

22.5. Explain the applications and limitations of ceramic tools.

22.6. List the major functions of cutting fluids.

22.7. Why is toughness important for cutting tool materials?

22.8. Is the elastic modulus important for cutting tool materials? Explain.

22.9. Explain how cutting fluids penetrate the tool–chip interface.

22.10. List the methods by which cutting fluids are typically applied in machining operations.

22.11. Describe the advantages and limitations of (a) single-crystal and (b) polycrystalline diamond tools.

22.12. What is a cermet? What are its advantages?

22.13. Explain the difference between M-series and T-series high-speed steels.

22.14. Why is cBN generally preferred over diamond for machining steels?

22.15. What are the advantages of dry machining?

Qualitative Problems

22.16. Explain why so many different types of cutting-tool materials have been developed over the years. Why are they still being developed further?

22.17. Which tool-material properties are suitable for interrupted cutting operations? Why?

22.18. Describe the reasons for and advantages of coating cutting tools with multiple layers of different materials.

22.19. Make a list of the alloying elements used in high-speed steels. Explain what their functions are and why they are so effective in cutting tools.

22.20. As stated in Section 22.1, tool materials can have conflicting properties when used for machining operations. Describe your observations regarding this matter.

22.21. Explain the economic impact of the trend shown in Fig. 22.6.

22.22. Why does temperature have such an important effect on tool life?

22.23. Ceramic and cermet cutting tools have certain advantages over carbide tools. Why, then, are they not completely replacing carbide tools?

22.24. What precautions would you take in machining with brittle tool materials, especially ceramics? Explain.

22.25. Can cutting fluids have any adverse effects in machining? If so, what are they?

22.26. Describe the trends you observe in Table 22.2.

22.27. Why are chemical stability and inertness important in cutting tools?

22.28. Titanium-nitride coatings on tools reduce the coefficient of friction at the tool–chip interface. What is the significance of this property?

22.29. Describe the necessary conditions for optimal utilization of the capabilities of diamond and cubic-boron-nitride cutting tools.

22.30. Negative rake angles generally are preferred for ceramic, diamond, and cubic-boron-nitride tools. Why?

22.31. Do you think that there is a relationship between the cost of a cutting tool and its hot hardness? Explain.

22.32. Make a survey of the technical literature, and give some typical values of cutting speeds for high-speed steel tools and for a variety of workpiece materials.

22.33. In Table 22.1, the last two properties listed can be important to the life of a cutting tool. Why?

22.34. It has been stated that titanium-nitride coatings allow cutting speeds and feeds to be higher than those for uncoated tools. Survey the technical literature and prepare a table showing the percentage increase of speeds and feeds that would be made possible by coating the tools.

22.35. Note in Fig. 22.1 that all tool materials—especially carbides—have a wide range of hardnesses for a particular temperature. Describe each of the factors that are responsible for this wide range.

22.36. Referring to Table 22.1, state which tool materials would be suitable for interrupted cutting operations. Explain.

22.37. Which of the properties listed in Table 22.1 is, in your opinion, the least important in cutting tools? Explain.

22.38. If a drill bit is intended only for woodworking applications, what material is it most likely to be made from? (*Hint:* Temperatures rarely rise to 400°C in woodworking.) Explain.

22.39. What are the consequences of a coating on a tool having a different coefficient of thermal expansion than the substrate material?

22.40. Discuss the relative advantages and limitations of near-dry machining. Consider all relevant technical and economic aspects.

22.41. Emulsion cutting fluids typically consist of 95% water and 5% soluble oil and chemical additives. Why is the ratio so unbalanced? Is the oil needed at all?

22.42. List and explain the considerations involved in determining whether a cutting tool should be reconditioned, recycled, or discarded after use.

22.43. In order of importance, list the important properties of cutting tool materials.

Quantitative Problems

22.44. Review the contents of Table 22.1. Plot several curves to show relationships, if any, among parameters such as hardness, transverse rupture strength, and impact strength. Comment on your observations.

22.45. Obtain data on the thermal properties of various commonly used cutting fluids. Identify those which are basically effective coolants (such as water-based fluids) and those which are basically effective lubricants (such as oils).

22.46. The first column in Table 22.2 shows 10 properties that are important to cutting tools. For each of the tool materials listed in the table, add numerical data for each of these properties. Describe your observations, including any data that overlap.

Synthesis, Design, and Projects

22.47. Describe in detail your thoughts regarding the technical and economic factors involved in tool-material selection.

22.48. One of the principal concerns with coolants is degradation due to biological attack by bacteria. To prolong the life of a coolant, chemical biocides often are added, but these biocides greatly complicate the disposal of the coolant. Conduct a literature search concerning the latest developments in the use of environmentally benign biocides in cutting fluids.

22.49. How would you go about measuring the effectiveness of cutting fluids? Describe your method and explain any difficulties that you might encounter.

22.50. Contact several different suppliers of cutting tools, or search their websites. Make a list of the costs of typical cutting tools as a function of various sizes, shapes, and features.

22.51. There are several types of cutting-tool materials available today for machining operations, yet much research and development is being carried out on all these materials. Discuss why you think such studies are being conducted.

22.52. Assume that you are in charge of a laboratory for developing new or improved cutting fluids. On the basis of the topics presented in this chapter and in Chapter 21, suggest a list of topics for your staff to investigate. Explain why you have chosen those topics.

22.53. Tool life could be greatly increased if an effective means of cooling and lubrication were developed. Design methods of delivering a cutting fluid to the cutting zone, and discuss the advantages and shortcomings of your design.

22.54. List the concerns you would have if you needed to economically machine carbon fiber-reinforced polymers or metal matrix composites with graphite fibers in an aluminum matrix.

Chapter 23

Machining Processes: Turning and Hole Making

Examples:

Case Studies:

- With the preceding two chapters as background, this chapter describes specific machining processes that are capable of generating round external or internal shapes.
- The most common machine tool used for such operations is the lathe, available in several types and degrees of automation.

- The wide variety of operations that can be performed on lathes are then described in detail, including turning, drilling, profiling, facing, grooving, thread cutting, and knurling.
- The chapter also describes operations such as boring, drilling, reaming, and tapping, and the characteristics of the machine tools associated with these processes.

Typical parts made: Machine components; engine blocks and heads; parts with complex shapes, close tolerances, and good surface finish; and externally and internally threaded parts.

Alternative processes: Precision casting, additive manufacturing, powder metallurgy, powder injection molding, abrasive machining, thread rolling, and rotary swaging.

23.1 Introduction

This chapter describes machining processes with the capability of producing parts with rotational symmetry. Typical products made are as small as miniature screws for the hinges of eyeglass frames, and as large as turbine shafts for hydroelectric power plants and rolls for rolling mills.

One of the most basic machining processes is **turning**, meaning that the part is rotated while it is being machined. The blank is generally a workpiece made by various processes, such as casting, forging, extrusion, drawing, or powder metallurgy, as described in Chapters 11-16. Turning operations, which typically are carried out on a **lathe** or by similar *machine tools*, are outlined in Fig. 23.1 and Table 23.1. These machines are highly versatile and capable of performing several machining operations that produce a wide variety of shapes, such as:

- **Turning:** to produce cylindrical, conical, curved, or grooved parts (Fig. 23.1a through d), such as shafts, spindles, and pins.
- **Facing:** to produce a flat surface at the end of the part and perpendicular to its axis (Fig. 23.1e); *face grooving* produces grooves for O-ring seats (Fig. 23.1f).
- **Machining with form tools:** (Fig. 23.1g) to produce various axisymmetric shapes for functional or for aesthetic purposes.
- **Boring:** to enlarge a hole or cylindrical cavity made by a previous process, or to produce circular internal grooves (Fig. 23.1h).
- **Drilling:** to produce a hole (Fig. 23.1i) which then may be followed by boring it, to improve its dimensional accuracy and surface finish.
- **Cutting off:** also called **parting**, to cut a piece from the end of a longer piece (Fig. 23.1j).
- **Threading:** to produce external or internal threads (Fig. 23.1k).
- **Knurling:** to produce a regularly shaped texture on cylindrical surfaces, as in making knobs and handles (Fig. 23.1l).

The machining operations summarized above are typically performed on a *lathe* (Fig. 23.2), available in a wide variety of designs, sizes, capacities, and computer-controlled features (Section 23.3 and Chapter 25). As shown in Figs. 21.2 and 23.3, turning is carried out at various (a) rotational speeds, N, of the workpiece clamped in a spindle, (b) depths of cut, d, and (c) feeds, f.

This chapter describes turning process parameters, cutting tools, process capabilities, and characteristics of the machine tools that are used to produce a variety of parts with *round* shapes. Design considerations to improve productivity for each group of processes also are described.

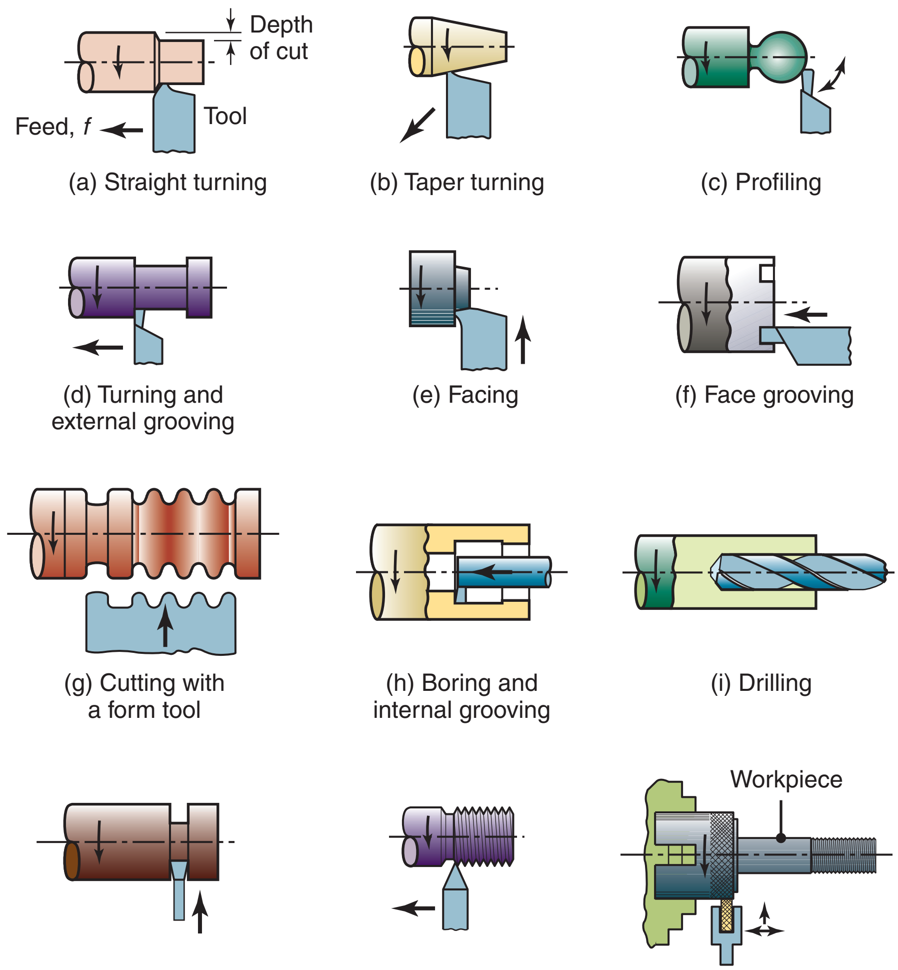


Figure 23.1: Miscellaneous cutting operations that can be performed on a lathe. Note that all parts are axisymmetric. The tools used, their shape, and the processing parameters are described in detail throughout this chapter.

23.2 The Turning Process

The majority of turning operations involve using simple single-point cutting tools. The geometry of a typical right-hand cutting tool is shown in Figs. 21.10 and 23.4. The tools are described by a standardized nomenclature; each type of workpiece material has its own optimum set of angles, developed largely through experience over many years (Table 23.2).

The major processing parameters that have a direct influence on machining processes, and the importance of controlling these parameters for optimized performance, have been described in Chapter 21. This section outlines the important turning-process parameters of tool geometry and material-removal rate, and gives data regarding recommended cutting practices, including tool materials, depth-of-cut, feed, cutting speed, and cutting fluids.

Tool Geometry. The various angles in a single-point cutting tool have specific functions in machining operations. These angles are measured in a coordinate system, consisting of the three major axes of the tool shank, as shown in Fig. 23.4.

Table 23.1: General Characteristics of Machining Processes and Typical Dimensional Tolerances.

Process	Characteristics	Typical dimensional tolerances, ±mm
Turning	Turning and facing operations on all types of materials, uses single-point or form tools; engine lathes require skilled labor; low production rate (but medium-to-high rate with turret lathes and automatic machines) requiring less skilled labor	Fine: 0.025–0.13 Rough: 0.13
Boring	Internal surfaces or profiles with characteristics similar to turning; stiffness of boring bar important to avoid chatter	0.025
Drilling	Round holes of various sizes and depths; high production rate; labor skill required depends on hole location and accuracy specified; requires boring and reaming for improved accuracy	0.075
Milling	Wide variety of shapes involving contours, flat surfaces, and slots; versatile; low-to-medium production rate; requires skilled labor	0.13–0.25
Planing	Large flat surfaces and straight contour profiles on long workpieces, low-quantity production, labor skill required depends on part shape	0.08–0.13
Shaping	Flat surfaces and straight contour profiles on relatively small workpieces; low-quantity production; labor skill required depends on part shape	0.05–0.13
Broaching	External and internal surfaces, slots, and contours; good surface finish; costly tooling; high production rate; labor skill required depends on part shape	0.025–0.15
Sawing	Straight and contour cuts on flat or structural shapes; not suitable for hard materials unless saw has carbide teeth or is coated with diamond; low production rate; generally low labor skill	0.8

- **Rake angle** is important in controlling both the direction of chip flow and the strength of the tool tip. Positive rake angles improve the cutting operation by reducing forces and temperatures, but they also result in a small included angle of the tool tip (as in Figs. 21.3 and 23.4), and are therefore prone to premature tool chipping and failure, depending on the toughness of the tool material.
- **Side rake angle** is more important than the **back rake angle**, which usually controls the direction of chip flow; these angles typically are in the range from –5° to 5°.

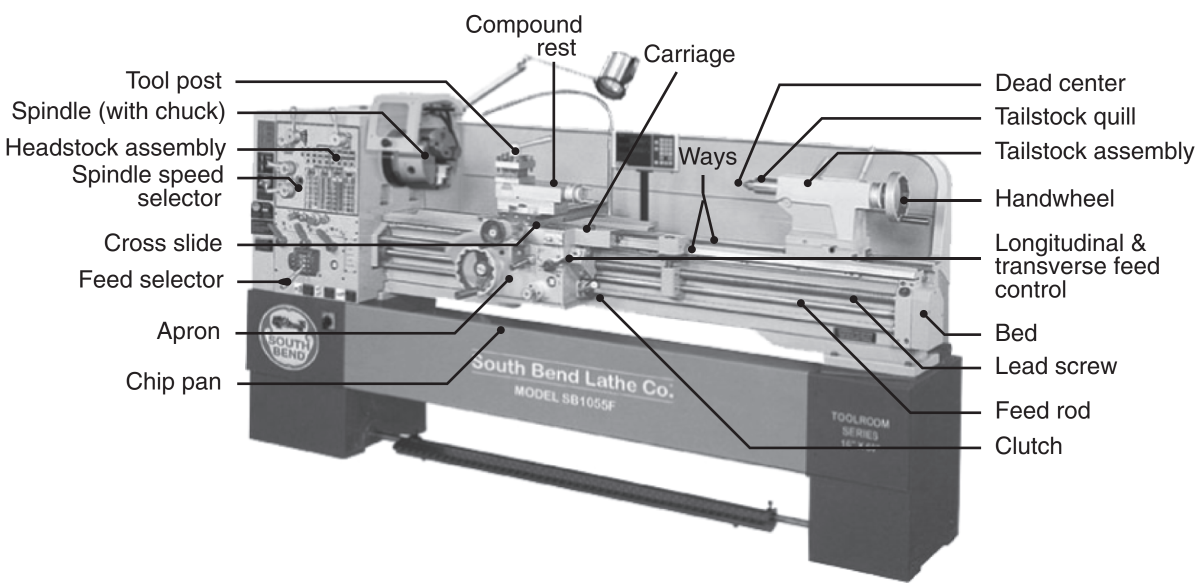


Figure 23.2: General view of a typical lathe, showing various components. *Source:* Courtesy of South Bend Lathe Co.

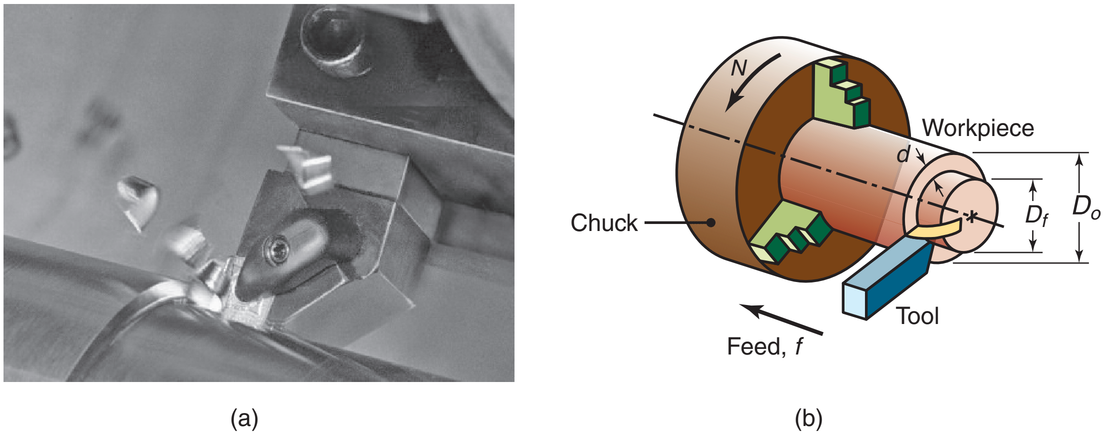


Figure 23.3: (a) Photograph of a turning operation, showing insert and discontinuous chips. The cutting tool is traveling from right to left in this photograph. (b) Schematic illustration of the basic turning operation, showing depth of cut, d; feed, f; and spindle rotational speed, N, in rev/min. The cutting speed is the surface speed of the workpiece at the tool tip. *Source:* (a) Courtesy of Kennametal Inc.

- **Cutting-edge angle** affects type of chip formation, tool strength, and cutting forces; typically, this angle is around 15°.
- **Relief angle** controls interference and rubbing at the tool–workpiece interface. If it is too large, the tool tip may chip off; if it is too small, flank wear may be excessive. This angle is typically 5°.
- **Nose radius** affects surface finish and tool-tip strength. The smaller the nose radius (meaning a sharp tool), the rougher the surface finish of the workpiece and the lower the strength of the tool, A large nose radius can, however, lead to tool *chatter* (see Section 25.4).

Material-removal Rate. The *material-removal rate* (MRR) in turning is the volume of material removed per unit time, and has the units of mm^3/min. Referring to Figs. 21.2 and 23.3, note that a ring-shaped layer of material is removed for each revolution of the workpiece; it has a cross-sectional area equal

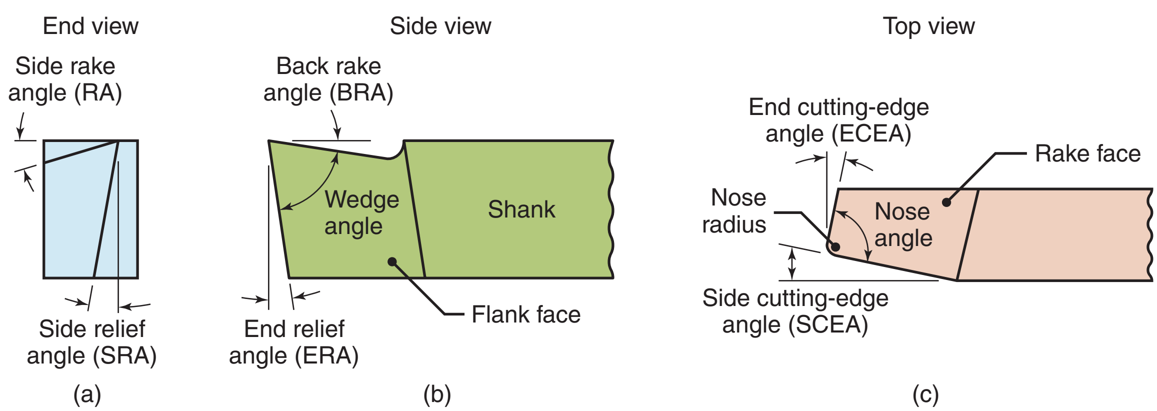


Figure 23.4: Designations for a right-hand tool. *Right-hand* means that the tool travels from right to left, as shown in Fig. 23.3b.

Table 23.2: General Recommendations for Tool Angles in Turning.

	High-speed steel					Carbide inserts				
Material	**Back rake**	**Side rake**	**End relief**	**Side relief**	**Side and end cutting edge**	**Back rake**	**Side rake**	**End relief**	**Side relief**	**Side and end cutting edge**
Aluminum and magnesium alloys	20	15	12	10	5	0	5	5	5	15
Copper alloys	5	10	8	8	5	0	5	5	5	15
Steels	10	12	5	5	15	-5	-5	5	5	15
Stainless steels	5	8–10	5	5	15	-5–0	-5–5	5	5	15
High-temperature alloys	0	10	5	5	15	5	0	5	5	45
Refractory alloys	0	20	5	5	5	0	0	5	5	15
Titanium alloys	0	5	5	5	15	-5	-5	5	5	5
Cast irons	5	10	5	5	15	-5	-5	5	5	15
Thermoplastics	0	0	20–30	15–20	10	0	0	20–30	15–20	10
Thermosets	0	0	20–30	15–20	10	0	15	5	5	15

to the product of the distance the tool travels in one revolution (the feed, f) and the depth of cut, d. The volume is the product of the cross-sectional area $(f)(d)$ and the average circumference of the ring that is machined, πD_{avg}, where

$$D_{\text{avg}} = \frac{D_o + D_f}{2}.$$

For light cuts on large-diameter workpieces, the average diameter may be replaced by D_o.

The rotational speed of the workpiece is N, and the material removal rate per revolution is $(\pi)(D_{\text{avg}})(d)(f)$. Since there are N revolutions per minute, the removal rate is

$$\text{MRR} = \pi D_{\text{avg}} d f N. \tag{23.1}$$

Note that Eq. (23.1) also can be written as

$$\text{MRR} = d f V, \tag{23.2}$$

where V is the cutting speed and MRR has the same unit of mm^3/min.

The cutting time, t, for a workpiece of length l can be calculated by noting that the tool travels at a feed rate of $fN = (\text{mm/rev})(\text{rev/min}) = \text{mm/min}$. Since the distance traveled is l mm, the cutting time is

$$t = \frac{l}{fN}. \tag{23.3}$$

The foregoing equations and the terminology used are summarized in Table 23.3. The cutting time in Eq. (23.3) does not include the time required for *tool approach* and *retraction*. Because the time spent in noncutting cycles of a machining operation is nonproductive, the time involved in approaching and retracting tools to and from the workpiece is an important consideration, affecting the overall economics of machining. Advanced machine tools are designed and built to minimize this time (see also Chapters 25, 37, and 38). One method of accomplishing this is to rapidly traverse the tools during noncutting cycles, followed by a slower movement as the tool engages the workpiece and starts cutting.

Forces in Turning. The three principal forces acting on a cutting tool in turning are shown in Fig. 23.5. These forces are important in the design of machine tools, as well as in the deflection of tools and workpieces, particularly in precision-machining operations (Section 25.7). It is essential that the machine tool and its components be able to withstand these forces without undergoing significant deflections, vibrations, and chatter.

The **cutting force**, F_c, acts downward on the tool, and thus tends to deflect it downward. This force supplies the energy required for the cutting operation. It can be calculated using the data given in Table 21.2, or from the energy per unit volume, described in Section 21.3. The product of the cutting force and its distance from the workpiece center is the *torque* on the spindle. The product of the torque and the spindle speed is the *power* required in the turning operation.

The **thrust force**, F_t, acts in the longitudinal direction; it is also called the **feed force**, because it is in the feed direction of the tool. This force tends to deflect the tool towards the right and away from the chuck in Fig. 23.5. The **radial force**, F_r, acts in the radial direction and tends to deflect the tool away from the workpiece. Because of the several factors involved in the cutting process, forces F_t and F_r are difficult to calculate directly, and are usually determined experimentally.

Roughing and Finishing Cuts. In machining, the usual procedure is to first take one or more *roughing cuts*, typically at high feed rates and large depths of cut. The material-removal rates are high, and there is little consideration for dimensional tolerance and surface roughness of the workpiece. These cuts are then followed by a *finishing cut*, typically done at a lower feed and smaller depth of cut, for a good surface finish.

Tool Materials, Feeds, and Cutting Speeds. The general characteristics of cutting-tool materials have been described in Chapter 22. A broad range of applicable cutting speeds and feeds for various tool materials

Table 23.3: Summary of Turning Parameters and Formulas.

N	=	Rotational speed of the workpiece, rpm
f	=	Feed, mm/rev
v	=	Feed rate, or linear speed of the tool along workpiece length, mm/min
	=	fN
V	=	Surface speed of workpiece, m/min
	=	$\pi D_o N$(for maximum speed)
	=	$\pi D_{\text{avg}} N$(for average speed)
l	=	Length of cut, mm
D_o	=	Original diameter of workpiece, mm
D_f	=	Final diameter of workpiece, mm
D_{avg}	=	Average diameter of workpiece, mm
	=	$(D_o + D_f)/2$
d	=	Depth of cut, mm
	=	$(D_o - D_f)/2$
t	=	Cutting time, s
	=	l/fN
MRR	=	mm^3/min
	=	$\pi D_{\text{avg}} d f N$
Torque	=	N-m
	=	$F_c D_{\text{avg}}/2$
Power	=	kW
	=	$(\text{Torque})\,(\omega)$, where $\omega = 2\pi N$ rad/min

Note: The units given are those that are commonly used; however, appropriate units must be used and checked in the formulas.

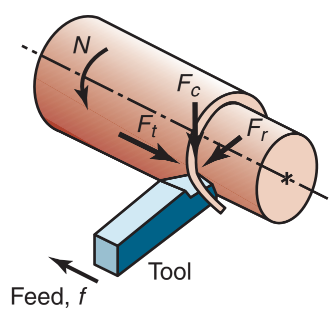


Figure 23.5: Forces acting on a cutting tool in turning. F_c is the cutting force, F_t is the thrust or feed force (in the direction of feed), and F_r is the radial force that tends to deflect the tool away from the workpiece being machined.

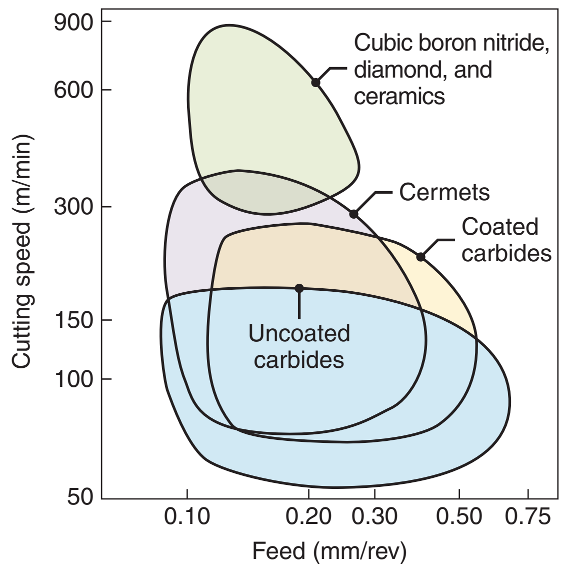


Figure 23.6: The range of applicable cutting speeds and feeds for a variety of tool materials.

is given in Fig. 23.6, as a general guideline in turning operations. Specific recommendations regarding turning-process parameters for various workpiece materials and cutting tools are given in Table 23.4. It is not uncommon to exceed these values in practice, especially with an effective coolant and a well-controlled process.

Case Study 23.1 Brake Disk Machining

An automotive brake manufacturer produces brake disks (see Fig. 23.7) by facing them on a lathe, using the processing parameters in Table 23.5. The disks are made from a cast blank, machined on a lathe; the mounting holes on the axle and the cooling holes in the disk are then produced on a CNC drill press. The material used is a gray cast iron (ASTM Class 25; see Table 12.4), using a silicon nitride insert. However, this material can have very poor machinability because of insufficient aging or variations in its composition. In addition, the cutting conditions have to be modified in order to increase production rate.

Aluminum oxide (Al_2O_3) and polycrystalline cubic boron nitride were investigated as alternative cutting tool materials. As can be seen in Table 23.4, cBN is the only tool material that would allow for an increased cutting speed, as compared to SiN for the gray cast iron workpiece. Based on the recommendations given in Table 23.4, the machining parameters shown in Table 23.5 were selected.

Using the cBN insert, it was found that the tool life could be dramatically increased to 4200 disks per tool edge, as compared to only 40 with the silicon nitride, so that the higher cost of cBN could be economically justified as well. Moreover, because of the longer life achieved, the tool change time was dramatically reduced, and the machine utilization was increased from 82% to 94%. Thus, a change to polycrystalline cBN led to a simultaneous improvement in economy and production rate.

Source: Courtesy of Kennametal, Inc.

Cutting Fluids. Many metallic and nonmetallic materials can be machined without a cutting fluid, but in most cases the application of a cutting fluid can significantly improve the operation. General recommendations for cutting fluids appropriate for various workpiece materials are given in Table 23.6. However, recall the major trend toward and the benefits of near-dry and dry machining (Section 22.12).

Table 23.4: General Recommendations for Turning Operations. These recommendations are for guidance only, and are often exceeded in practice.

		General-purpose starting conditions			Range for roughing and finishing		
Workpiece material	Cutting tool	Depth of cut, mm	Feed, mm/rev	Cutting speed,	Depth of cut,	Feed, mm/rev	Cutting speed, m/min
Low-C and free machining steels	Uncoated carbide	1.5–6.3	0.35	90	0.5–7.6	0.15–1.1	60–135
	Ceramic-coated carbide	″	″	245–275	″	″	180–495
	Triple-coated carbide	″	″	185–200	″	″	90–245
	TiN-coated carbide	″	″	105–150	″	″	60–230
	Al_2O_3 ceramic	″	0.25	395–440	″	″	365–550
	Cermet	″	0.30	215–290	″	″	180–455
Medium and high-C steels	Uncoated carbide	1.2–4.0	0.30	75	2.5–7.6	0.15–0.75	135–225
	Ceramic-coated carbide	″	″	185–230	″	″	120–410
	Triple-coated carbide	″	″	120–150	″	″	75–215
	TiN-coated carbide	″	″	90–200	″	″	45–215
	Al_2O_3 ceramic	″	0.25	335	″	″	245–455
	Cermet	″	0.25	170–245	″	″	105–305
Cast iron, gray	Uncoated carbide	1.25–6.3	0.32	90	0.4–12.7	0.1–0.75	75–185
	Ceramic-coated carbide	″	″	200	″	″	120–365
	TiN-coated carbide	″	″	90–135	″	″	60–215
	Al_2O_3 ceramic	″	0.25	455–490	″	″	365–855
	SiN ceramic	″	0.32	730	″	″	200–990
	Polycrystalline cBN	″	″	1000	″	″	200–1160

Table 23.4: General Recommendations for Turning Operations. These recommendations are for guidance only, and are often exceeded in practice. (*cont.*)

		General-purpose starting conditions			Range for roughing and finishing		
Workpiece material	**Cutting tool**	**Depth of cut, mm**	**Feed, mm/rev**	**Cutting speed,**	**Depth of cut,**	**Feed, mm/rev**	**Cutting speed, m/min**
Stainless steel, austenitic	Triple-coated carbide	1.5–4.4	0.35	150	0.5–12.7	0.08–0.75	75–230
	TiN-coated carbide	″	″	85–160	″	″	55–200
	Cermet	″	0.30	185–215	″	″	135–315
High-temperature alloys, nickel based	Uncoated carbide	2.5	0.15	25–45	0.25–6.3	0.1–0.3	15–30
	Ceramic-coated carbide	″	″	45	″	″	20–60
	TiN-coated carbide	″	″	30–55	″	″	20–85
	Al_2O_3 ceramic	″	″	260	″	″	185–395
	SiN ceramic	″	″	215	″	″	90–215
	Polycrystalline cBN	″	″	150	″	″	120–185
Titanium alloys	Uncoated carbide	1.0–3.8	0.15	35–60	0.25–6.3	0.1–0.4	10–75
	TiN-coated carbide	″	″	30–60	″	″	15–170
Aluminum alloys Free machining	Uncoated carbide	1.5–5.0	0.45	490	0.25–8.8	0.08–0.62	200–670
	TiN-coated carbide	″	″	550	″	″	60–915
	Cermet	″	″	490	″	″	215–795
	Polycrystalline diamond	0.1–4.0	0.1–0.4	760	″	″	1000–5000 (3200–16,250
High silicon	Polycrystalline diamond	″	″	530	″	″	365–915

Table 23.4: General Recommendations for Turning Operations. These recommendations are for guidance only, and are often exceeded in practice. (*cont.*)

		General-purpose starting conditions			Range for roughing and finishing		
Workpiece material	Cutting tool	Depth of cut, mm	Feed, mm/rev	Cutting speed,	Depth of cut,	Feed, mm/rev	Cutting speed, m/min
Copper alloys	Uncoated carbide	1.5–5.0	0.25	260	0.4–7.51	0.15–0.75	105–535
	Ceramic-coated carbide	″	″	365	″	″	215–670
	Triple-coated carbide	″	″	215	″	″	90–305
	TiN-coated carbide	″	″	90–275	″	″	45–455
	Cermet	″	″	245–425	″	″	200–610
	Polycrystalline diamond	″	″	520	0.05–2.0	0.03–0.3	400–1300
Tungsten alloys	Uncoated carbide	2.5	0.2	75	0.25–5.0	0.12–0.45	55–120
	TiN-coated carbide	″	″	85	″	″	60–150
Thermoplastics and thermosets	TiN-coated carbide	1.2	0.12	170	0.12–5.0	0.08–0.35	90–230
	Polycrystalline diamond	″	″	395	″	″	250–730
Composites, graphite reinforced	TiN-coated carbide	1.9	0.2	200	0.12–6.3	0.12–1.5	105–290
	Polycrystalline diamond	″	″	760	″	″	550–1310

Source: Based on data from Kennametal Inc.
Note: Cutting speeds for high-speed steel tools are about one-half those for uncoated carbides.

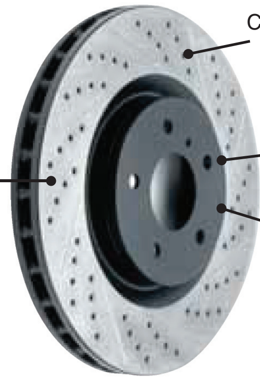


Figure 23.7: Brake disk, used in Case Study 23.1. *Source:* Courtesy of Kennametal, Inc.

Table 23.5: Process Parameter Comparison for SiN and cBN Tools in Facing of a Brake Disk.

	Cutting tool material	
Parameter	**SiN**	**cBN**
Depth of cut, mm	1.5	2.0
Feed, mm/rev	0.5	0.4
Cutting speed	700	1000
Tool life	40	4200
Machine utility	82%	94%

Table 23.6: General Recommendations for Cutting Fluids for Machining (see also Section 33.7).

Material	Type of fluid
Aluminum	D, MO, E, CSN
Beryllium	MC, E, CSN
Copper	D, E, CSN
Magnesium	D, MO
Nickel	MC, E, CSN
Refractory metals	MC, E, EP
Steels	
Carbon and low-alloy	D, MO, E, CSN, EP
Stainless	D, MO, E, CSN
Titanium	CSN, EP, MO
Zinc	C, MC, E, CSN
Zirconium	D, E, CSN

Note: CSN = chemical and synthetics; D = dry; E = emulsion; EP = extreme pressure; FO = fatty oil; and MO = mineral oil.

Example 23.1 Material-removal Rate and Cutting Force in Turning

Given: A 150-mm-long, 10-mm-diameter, 304 stainless-steel rod is being reduced in diameter to 8 mm by turning on a lathe. The spindle rotates at $N = 400$ rpm, and the tool is traveling at an axial speed of 200 mm/min.

Find: Calculate the cutting speed, material-removal rate, cutting time, power dissipated, and cutting force.

Solution: The cutting speed is the tangential speed of the workpiece. The maximum cutting speed is at the outer diameter, D_o, and is obtained from the equation

$$V = \pi D_o N.$$

Thus,

$$V = (\pi)(0.010)(400) = 12.57 \text{ m/min}.$$

The cutting speed at the machined diameter is

$$V = (\pi)(0.008)(400) = 10.05 \text{ m/min}.$$

From the information given, note that the depth of cut is

$$d = \frac{10 - 8}{2} = 1 \text{ mm} = 0.001 \text{ m}.$$

and the feed is

$$f = \frac{200}{400} = 0.5 \text{ mm/rev} = 0.0005 \text{ m/rev}.$$

According to Eq. (23.1), the material-removal rate is then

$$\text{MRR} = (\pi)(9)(1)(0.5)(400) = 5655 \text{ mm}^3/\text{min}.$$

The actual time to cut, according to Eq. (23.4), is

$$t = \frac{150}{(0.5)(400)} = 0.75 \text{ min}.$$

The power required can be calculated by referring to Table 21.2 and taking an average value for stainless steel as 4.1 W-s/mm^3. Therefore, the power dissipated is

$$\text{Power} = \frac{(4.1)(5655)}{60} = 386 \text{ W}.$$

The cutting force, F_c, is the tangential force exerted by the tool. Since power is the product of torque, T, and rotational speed in radians per unit time,

$$T = \frac{(386)}{(400)(2\pi/60)} = 9.2 \text{ Nm}.$$

Since $T = (F_c)(D_{\text{avg}}/2)$,

$$F_c = \frac{(9.2)(2)}{(0.009)} = 2.0 \text{ kN}.$$

23.3 Lathes and Lathe Operations

Lathes generally are considered to be the oldest machine tools. Although woodworking lathes originally were developed after 1000 B.C., *metalworking lathes*, with lead screws, were not built until the late 1700s. The most common lathe originally was called an *engine lathe*, because it was powered with overhead pulleys and belts from a nearby engine on the factory floor. Lathes became equipped with individual electric motors beginning in the late 19th century.

The maximum spindle speed of lathes is typically around 4000 rpm, but may be only about 200 rpm for large lathes. For special applications, speeds may range to 10,000 rpm or even higher for very *high-speed machining* (see Section 25.5). The cost of lathes ranges from about $2,000 for bench types to over $100,000 for larger units.

23.3.1 Lathe Components

Lathes are equipped with a variety of components and accessories, as shown in Fig. 23.2. Their basic features and functions are:

Bed. The bed supports all major components of the lathe; it has a large mass and is built rigidly, usually from gray or nodular cast iron. The top portion of the bed has two **ways**, with various cross sections that are hardened and machined for wear resistance and good dimensional accuracy during turning. In a *gap-bed lathe*, a section of the bed in front of the headstock can be removed to accommodate workpieces with larger diameters (see also Section 25.3 on advanced materials for machine-tool structures).

Carriage. The carriage, or *carriage assembly*, slides along the ways; it consists of an assembly of the *cross-slide*, *tool post*, and *apron*. The cutting tool is mounted on the *tool post*, usually with a *compound rest* that swivels for tool positioning and adjustments. The *cross-slide* moves radially in and out, controlling the radial position of the cutting tool in such operations as facing (see Fig. 23.1e). The apron is equipped with mechanisms for both manual and mechanized movement of the carriage and the cross-slide by means of the *lead screw*.

Headstock. The headstock is fixed to the left side of the bed and is equipped with motors, pulleys, and V-belts, supplying power to a *spindle* at various rotational speeds, which can be set through manually controlled selectors or by electrical controls. Most headstocks are equipped with a set of gears, and some have various drives to provide a *continuously variable* range of speed to the spindle. Headstocks have a *hollow spindle* to which workholding devices (such as *chucks* and *collets*; see Section 23.3.2) are mounted; long bars or tubing can thus be fed through them for various turning operations. The dimensional accuracy of the spindle is important for precision in turning, particularly in high-speed machining. Preloaded tapered or ball bearings are typically used to rigidly support the spindle.

Tailstock. The tailstock, which can slide along the ways and be clamped at any position, supports the right end of the workpiece. It is equipped with a *center*, which may be fixed (called *dead center*) or it may be free to rotate with the workpiece (*live center*). Drills and reamers (Sections 23.5 and 23.6) can be mounted on the tailstock *quill* (a hollow cylindrical piece with a tapered hole) to drill axial holes in the workpiece.

Feed Rod and Lead Screw. The feed rod is powered by a set of gears through the headstock. It rotates during the lathe operation, and provides movement to the carriage and the cross-slide by means of gears, a friction clutch, and a keyway along the length of the rod. Closing a *split nut* around the lead screw engages the rod with the carriage; the split nut is also used for cutting threads accurately.

Lathe Specifications. A lathe is generally specified by the following parameters:

- *Swing*, the maximum diameter of the workpiece that can be accommodated (Table 23.7); it may be as much as 2 m.
- Maximum distance between the headstock and tailstock centers.
- Length of the bed.

Table 23.7: Typical Capacities and Maximum Workpiece Dimensions for Machine Tools.

Machine tool	Maximum dimension (m)	Power (kW)	Maximum speed (rpm)
Lathes (swing/length)			
Bench	0.3/1	< 1	3000
Engine	3/5	70	12,000
Turret	0.5/1.5	60	6000
Automatic screw machines	0.1/0.3	20	10,000
Boring machines (work diameter/length)			
Vertical spindle	4/3	200	300
Horizontal spindle	1.5/2	70	2000
Drilling machines			
Bench and column (drill diameter)	0.1	10	12,000
Radial (column to spindle distance)	3	–	–
Numerical control (table travel)	4	–	–

Note: Larger capacities are available for special applications.

23.3.2 Workholding Devices and Accessories

Workholding devices are important, since they must hold the workpiece securely in place while machining. As shown in Fig. 23.3, one end of the workpiece is clamped to the lathe spindle either by a chuck, collet, face plate (see Fig. 23.8d), or a mandrel.

A **chuck** is usually equipped with three or four *jaws*. *Three-jaw* chucks generally have a geared-scroll design which makes the jaws self-centering. They are used for round workpieces, such as bar stock, pipes,

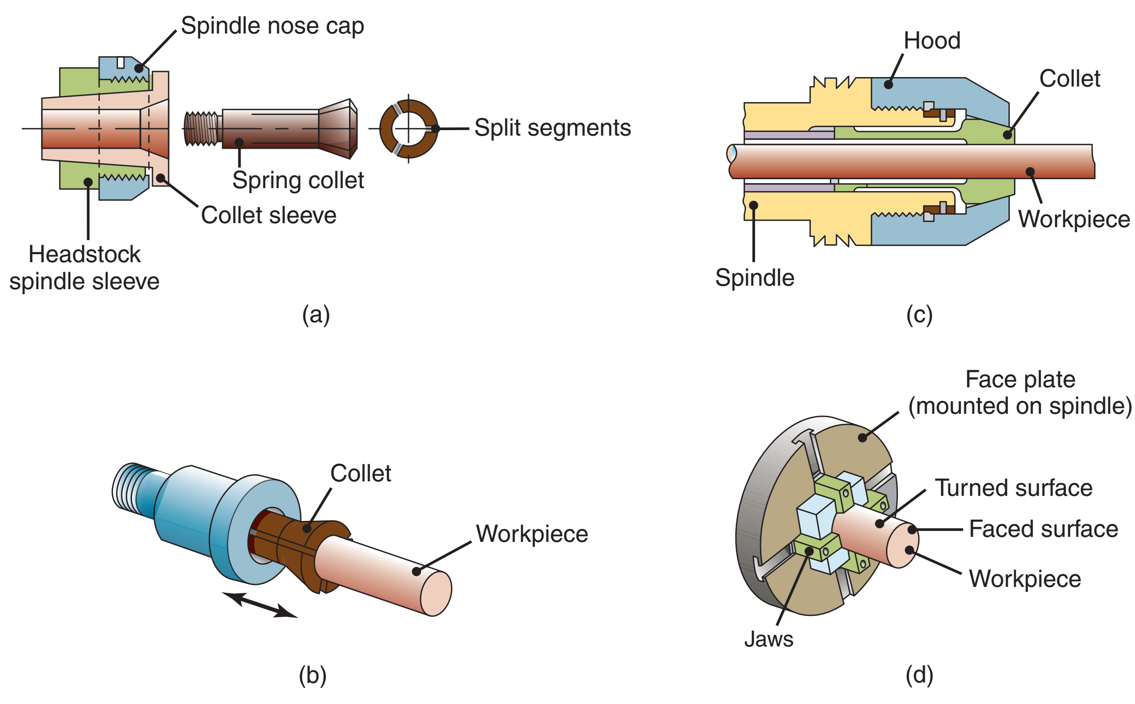


Figure 23.8: (a) and (b) Schematic illustrations of a draw-in type of collet. The workpiece is placed in the collet hole, and the conical surfaces of the collet are forced inward by pulling it with a draw bar into the sleeve. (c) A push-out type of collet. (d) Workholding of a workpiece on a face plate.

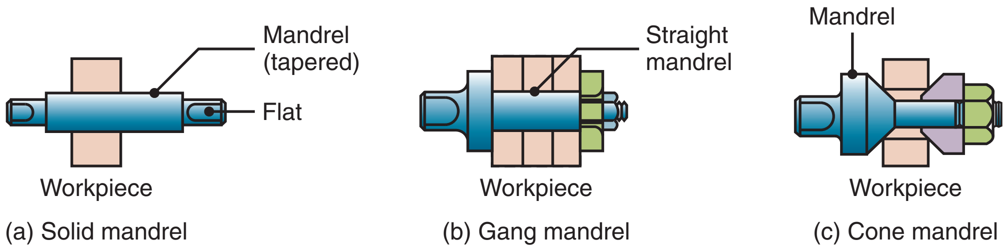


Figure 23.9: Various types of mandrels to hold workpieces for turning; they are usually mounted between centers on a lathe. Note that in (a) both the cylindrical and the end faces of the workpiece can be machined, whereas in (b) and (c) only the cylindrical surfaces can be machined.

and tubing, and typically can be centered to within 0.025 mm. *Four-jaw* chucks have jaws that can be moved and adjusted independently of each other; thus, they can be used for square, rectangular, or odd-shaped workpieces. The jaws in some types of chucks can be reversed to permit clamping of hollow workpieces, such as pipes and tubing, either on their outside or inside surfaces. Also available are jaws made of low-carbon steel (called *soft jaws*) that can be machined into desired shapes. Because of their low strength and hardness, soft jaws also conform to small irregularities on workpieces, thus ensuring better clamping. Chucks can be *power* or *manually actuated*, using a chuck wrench.

Power chucks, actuated either pneumatically or hydraulically, are used in automated equipment for high production rates, including loading of parts using industrial robots (Section 37.6). Also available are several types of power chucks, with lever- or wedge-type mechanisms for actuating the jaws. Chucks are available in various designs and sizes. Their selection depends on the type and speed of operation, workpiece size, production and dimensional accuracy requirements, and the jaw clamping forces required. By controlling the jaw forces, an operator can ensure that the part does not slip or distort in the chuck during machining. High spindle speeds can significantly reduce jaw forces due to *centrifugal forces*.

A **collet** is basically a longitudinally split, tapered bushing. The workpiece, generally with a maximum diameter of 25 mm, is placed inside the collet, and the collet is pulled (*draw-in collet*; Fig. 23.8a and b) or pushed (push-out collet; Fig. 23.8c) mechanically into the spindle. The tapered surfaces shrink the segments of the collet radially, tightening them onto the workpiece. Collets are used for round or other shapes. An advantage to using a collet, rather than a three- or four-jaw chuck, is that the collet grips nearly the entire circumference of the part, making it well suited particularly for parts with small cross sections.

Face plates are used for clamping irregularly shaped workpieces; they are round and have several slots and holes through which the workpiece is bolted or clamped (Fig. 23.8d). **Mandrels** (Fig. 23.9) are placed inside hollow or tubular workpieces, and are used to hold workpieces that require machining on both ends or on their cylindrical surfaces.

Accessories. Several devices are used as accessories and attachments for lathes. Among these are the following:

- Carriage and cross-slide stops, to stop the carriage at a predetermined distance along the bed.
- Devices for turning parts having a variety of tapers.
- Various attachments for milling, boring, drilling, thread cutting, gear-cutting, sawing, and grinding operations.

23.3.3 Lathe Operations

In a typical turning operation, the workpiece is clamped by any one of the work-holding devices described previously. Long and slender parts must be supported by a *steady* rest placed on the bed, or by a *follow* rest, to keep the part from deflecting excessively under the cutting forces. The rests usually are equipped with three adjustable fingers or rollers that support the workpiece while allowing it to rotate freely. Steady rests are clamped directly on the ways of the lathe (as in Fig. 23.2), whereas follow rests are clamped on the carriage and travel with it.

The cutting tool is attached to the tool post, which is driven by the lead screw. The cutting tool removes material by traveling along the bed. A *right-hand* tool travels toward the headstock, and a *left-hand* tool travels toward the tailstock. Facing operations are done by moving the tool radially inward with the cross-slide.

Form tools are used to machine various shapes on solid, round workpieces (Fig. 23.1g), by moving the tool radially inward while the part is rotating. Form cutting is not suitable for deep and narrow grooves or sharp corners, because of vibration and chatter. As a general rule, (a) the formed length of the part should not be greater than about 2.5 times the minimum diameter of the part and (b) cutting fluids should be used. The stiffness of the machine tools and workholding devices also are important considerations.

Boring involves machining inside hollow workpieces or enlarging a hole; it is similar to turning. Out-of-shape round holes also can be straightened by boring. Boring large workpieces is described in Section 23.4.

Drilling (Section 23.5) can be performed on a lathe by mounting the drill bit in a chuck or in the tailstock quill. The workpiece is clamped in a workholder on the headstock, and the drill bit is advanced by rotating the handwheel of the tailstock. The concentricity of the hole can be improved by subsequently boring the drilled hole. For better dimensional accuracy and surface finish, drilled holes may later be **reamed** (Section 23.6) on lathes, in a manner similar to drilling.

The cutting tools for *parting*, *grooving*, and *thread cutting* are specially shaped for their specific purpose or are available as inserts. *Knurling* is performed on a lathe, with hardened rolls (see Fig. 23.1l); the surface of the rolls is a replica of the profile to be generated. The rolls are pressed radially against the rotating workpiece while the tool moves axially along the part.

23.3.4 Types of Lathes

Bench Lathes. As the name suggests, these lathes are placed on a workbench or a table; they have low power and are usually operated by hand feed. *Toolroom bench lathes* have higher precision, enabling the machining of parts to close dimensional tolerances.

Special-purpose Lathes. These lathes are used for such applications as railroad wheels, gun barrels, and rolling-mill rolls.

Tracer Lathes. These lathes have special attachments for turning parts with various contours. Also called a *duplicating lathe* or *contouring lathe*, the cutting tool follows a path that duplicates the contour of a template, similar to a pencil following the shape of a stencil. These machines have largely been replaced by *numerical-control lathes* and *turning centers* (Section 25.2), although tracer attachments are available for engine lathes.

Automatic Lathes. Mechanisms have been developed that enable machining operations on a lathe to follow a certain prescribed sequence. In a *fully automatic lathe*, parts are fed and removed automatically; in *semiautomatic* machines, these functions are performed by the operator, although machining remains automatic. Automatic lathes, either with a horizontal or a vertical spindle, are suitable for medium- to high-volume production.

Lathes without tailstocks are called *chucking machines* or *chuckers*. They are used for machining individual pieces with regular or irregular shapes, and are either single- or multiple-spindle types. In another type of an automatic lathe, the bar stock is fed periodically into the lathe, and after a part is machined, it is cut off from the end of the bar stock.

Automatic Bar Machines. Also called **automatic screw machines**, these machines are designed for high-production-rate machining of screws and similar threaded parts; all operations are performed automatically, with tools attached to a special turret. After each part is machined to finished dimensions, the bar

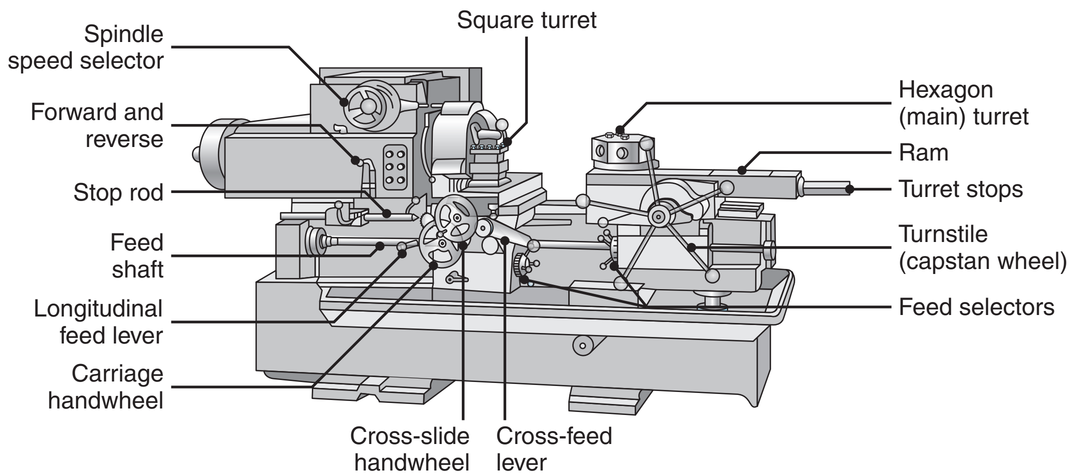


Figure 23.10: Schematic illustration of the components of a turret lathe; note the two turrets: square and hexagonal (main).

stock is fed forward automatically through the hole in the spindle, and then the part cut off. Automatic bar machines may be equipped with single or multiple spindles; capacities typically range from 3- to 150-mm diameter bar stock.

Single-spindle automatic bar machines are similar to turret lathes, and are equipped with various cam-operated mechanisms; they are capable of high-precision machining of small-diameter parts. There are two types of single-spindle machines. In *Swiss-type automatics*, the cylindrical surface of a solid-bar stock is machined, using a series of tools that move radially and in the same plane toward the workpiece. The bar stock is clamped close to the headstock spindle, to minimize deflections due to cutting forces.

Multiple-spindle automatic bar machines typically have from four to eight spindles, arranged in a circle on a large drum, with each spindle carrying an individual workpiece. The cutting tools are arranged in various positions in the machine, and move in both axial and radial directions. Each part is machined in stages as it moves from one station to the next. Because all operations are carried out simultaneously, the cycle time per part is reduced.

Turret Lathes. These machine tools are capable of performing multiple cutting operations, such as turning, boring, drilling, thread cutting, and facing (Fig. 23.10). Several cutting tools, usually as many as six, are mounted on the hexagonal *main turret*, which is rotated after each specific operation is completed. The lathe usually has a *square turret* on the cross-slide, and is equipped with as many as four cutting tools. The workpiece, generally long, round bar stock, is advanced a preset distance through the chuck. After the part is machined, it is cut off by a tool mounted on the square turret, which moves radially into the workpiece. The rod then is advanced the same preset distance, and the next part is machined.

Turret lathes are versatile; the operations may be carried out automatically or by hand, using the *capstan wheel*. Once set up, these machines do not require highly skilled operators. *Vertical turret lathes* are more suitable for short, heavy workpieces, with diameters as large as 1.2 m.

The turret lathe shown in Fig. 23.10 is known as a **ram-type** turret lathe, in which the ram slides in a separate base on the saddle. The short stroke of the turret slide limits this machine to relatively short workpieces and light cuts, in both small- and medium-quantity production. In another design, called the **saddle type**, the main turret is installed directly on the saddle, which slides along the bed; the length of the stroke is limited only by the length of the bed. This type of lathe is constructed more heavily, and is used for machining large workpieces.

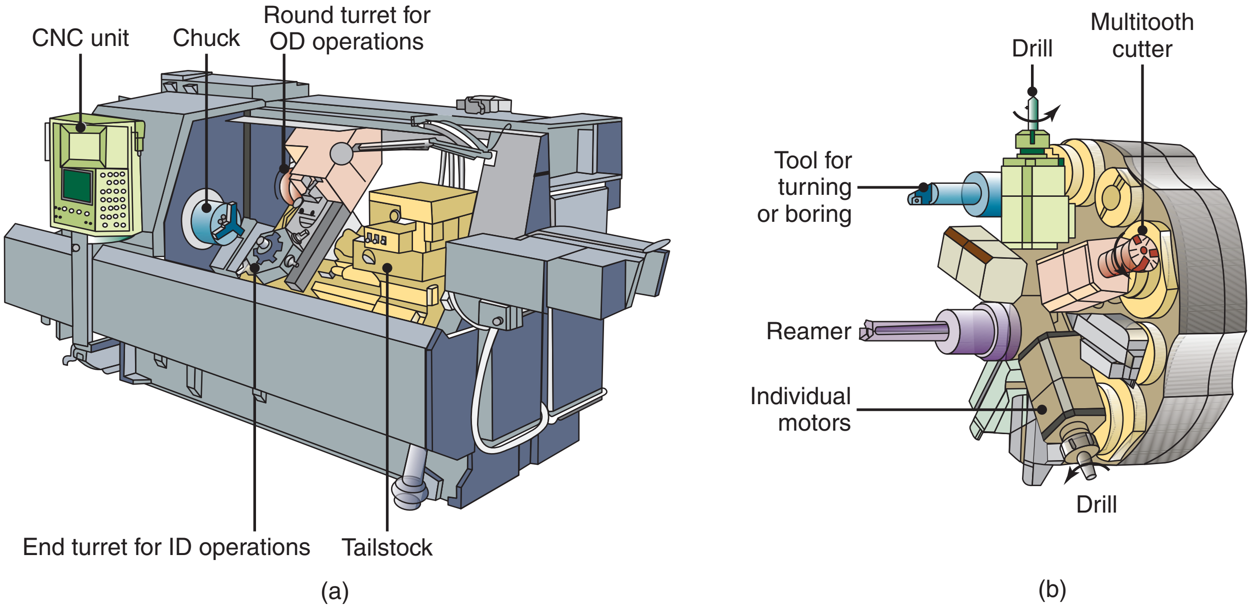


Figure 23.11: (a) A computer-numerical-control lathe with two turrets. (b) A typical turret equipped with 10 tools, some of which are powered.

Computer-controlled Lathes. The features of a computer-controlled lathe (CNC) are shown in Fig. 23.11a. These lathes generally have one or more turrets; each turret is equipped with a variety of tools and performs several machining operations on different surfaces of the workpiece (Fig. 23.11b), with diameters as much as 1 m.

Computer-controlled lathes are designed to operate faster and with higher power, and are equipped with *automatic tool changers* (ATCs). Their operations are reliably repetitive, maintain dimensional accuracy, require less skilled labor, and are suitable for low- to medium-volume production.

Example 23.2 Typical Parts Made on CNC Turning Machine Tools

The capabilities of CNC turning-machine tools are illustrated in the machined parts shown in Fig. 23.12, indicating the workpiece material, the number of cutting tools used, and the machining times. These parts also can be made on manual or turret lathes, although not as effectively or consistently.

Source: Courtesy of Monarch Machine Tool Company.

Example 23.3 Machining of Complex Shapes

Note in Example 23.2 that the parts are axisymmetric. The capabilities of CNC turning are further illustrated in Fig. 23.13, which shows three additional, more complex parts: a pump shaft, a crankshaft, and a tubular part with an internal rope thread. Descriptions of these parts are as follows; as in most operations, machining such parts consists of both roughing and finishing cuts:

1. *Pump shaft* (Fig. 23.13a). This part, as well as a wide variety of similar parts with external and internal features, including camshafts, was produced on a CNC lathe with two turrets. The lathe is similar in construction to the machine tool shown in Fig. 23.11a; each turret can hold as many

as eight tools. To produce this particular shape, the upper turret is programmed in such a manner that its radial movement is synchronized with the shaft rotation (Fig. 23.13b).

The spindle rotation is monitored directly, a processor performs a high-speed calculation, and the CNC then issues a command to the cam turret in terms of that angle. The machine has *absolute-position feedback*, using a high-accuracy scale system. The CNC compares the actual value with the one commanded, then performs an automatic compensation, using a built-in learning function. The turret has a lightweight design for smooth operation, which also reduces inertial forces.

The shaft can be made of aluminum or stainless steel. The machining parameters for aluminum are given in Table 23.8 (see Part (a) in the first column of the table). These parameters may be compared with the data given in Table 23.4, which has only a broad and approximate range as a guideline. The inserts were a K10 (C3) uncoated carbide, with a compacted polycrystalline diamond (see Fig. 22.10). The OD machining in the table shown refers to the two straight cylindrical ends of the part. The total machining time for an aluminum shaft was 24 min; for stainless steel, it was 55 min, because the cutting speed for stainless steel is considerably lower than that for aluminum.

2. *Crankshaft* (Fig. 23.13c). This part is made of ductile (nodular) cast iron; the machining parameters are shown in Part (b) of Table 23.8. The insert was K10 carbide. The machining time was 25 min, which is of the same order of magnitude as that for the pump shaft described above.

3. *Tubular part with internal rope threads* (Fig. 23.13d). This part, made of 304 stainless steel, was machined under the conditions given for Part (c) in Table 23.8. The starting blank was a straight tubular piece, similar to a bushing. The cutting tools were coated carbide and cermet. The boring bar was made of tungsten carbide, for increased stiffness, and, thus, improved dimensional accuracy and surface finish. For the threaded portion of the part, the dimensional accuracy was ± 0.05 mm, with a surface finish of $R_a = 2.5\ \mu$m.

   Machining time for this part was 1.5 min, much shorter than those for the previous two parts. The reason is that (a) this part is shorter, (b) less material is removed, (c) it does not have the eccentricity features of the first two parts, so the radial movement of the cutting tool is not a function of the angular position of the part, and (d) the cutting speed is higher.

Source: Based on technical literature supplied by Okuma Corp.

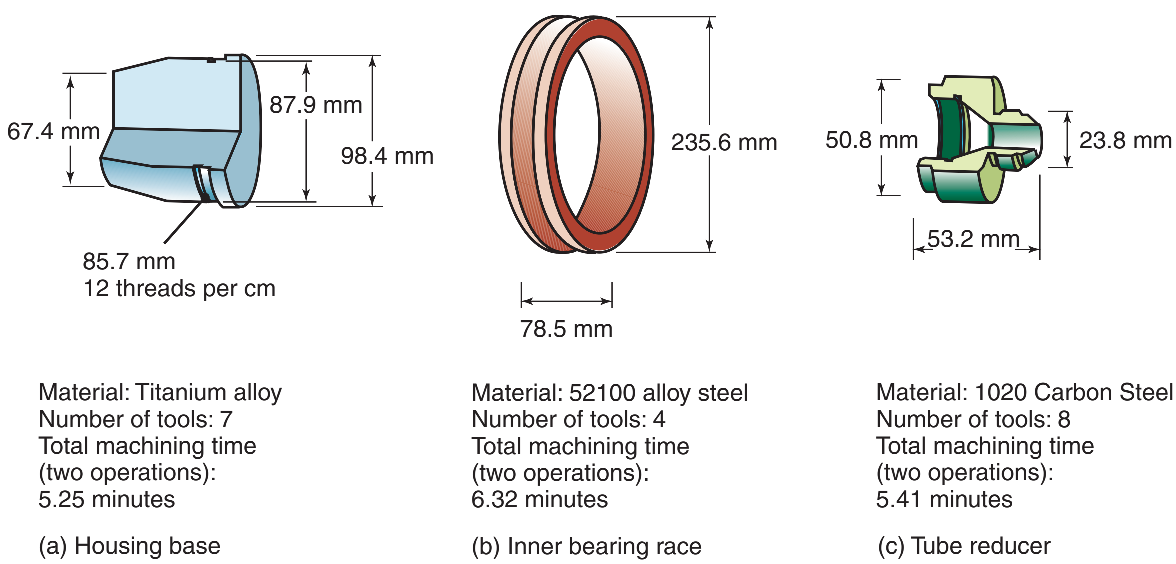


Figure 23.12: Typical parts made on CNC lathes.

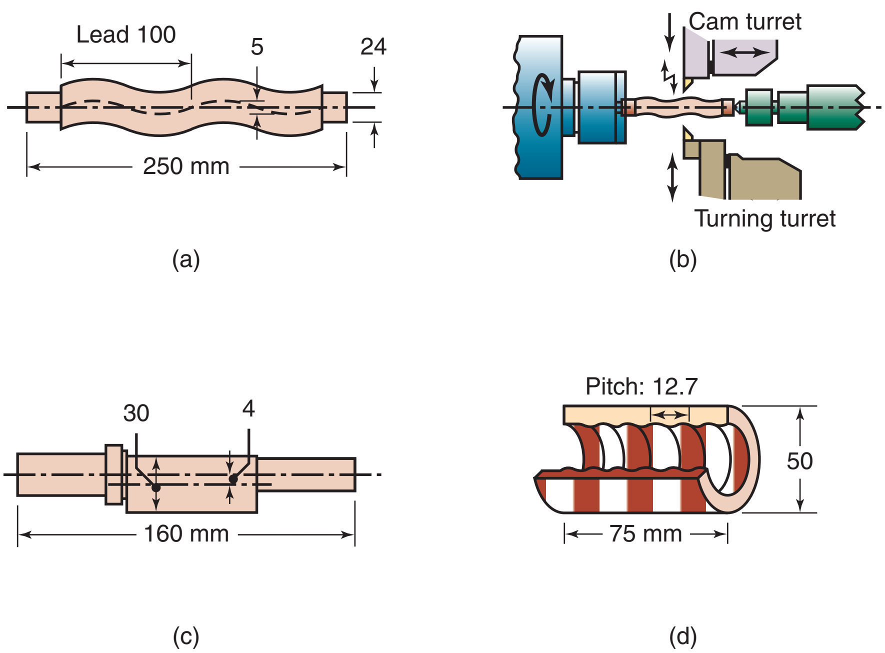


Figure 23.13: Examples of complex shapes that can be produced on a CNC lathe.

Table 23.8: Machining Summary for Example 23.3.

Operation	Speed (rpm)	Cutting speed	Depth of cut	Feed	Tool
Parts a and b:					
Outer diameter (OD)					
Roughing	1150	160 m/min	3 mm	0.3 mm/rev	K10 (C3)
Finishing	1750	250	0.2	0.15	K10 (C3)
Lead					
Roughing	300	45	3	0.15	K10
Finishing	300	45	0.1	0.15	Diamond compact
Part c: Eccentric shaft					
Roughing	200	5–11	1.5	0.2	K10 (C3)
Finishing	200	5–11	0.1	0.05	K10 (C3)
Part d: Internal thread					
Roughing	800	70	1.6	0.15	Coated carbide
Finishing	800	70	0.1	0.15	Cermet

23.3.5 Turning-process Capabilities

Relative *production rates* in turning, as well as in other machining operations described in the rest of this chapter and in Chapter 24, are shown in Table 23.8. These rates have an important bearing on productivity in machining operations. Note that there are major differences in production rates among the processes listed. The differences are due not only to the inherent characteristics of the processes and machine tools, but also to various factors, such as the setup times and the types and sizes of the workpieces to be machined.

The ratings given in Table 23.9 are relative, and there can be significant variations in special applications. For example, heat-treated, high-carbon cast-steel rolls for rolling mills can be machined on special lathes, using cermet tools and at material-removal rates as high as 6000 cm^3/min. Also called **high-removal-rate machining**, the process has at least two important requirements: (a) very high machine-tool rigidity, to avoid chatter and associated tool breakage, and (b) high power, of up to 450 kW.

The surface finish (Fig. 23.14) and dimensional accuracy (Fig. 23.15) obtained in turning and related operations depend on several factors: the characteristics and condition of the machine tool, stiffness, vibration and chatter, processing parameters, tool geometry, tool wear, cutting fluids, machinability of the workpiece material, and, when applicable, operator skill. A wide range of surface finishes can be obtained, as shown in Fig. 23.14 (see also Fig. 33.5).

23.3.6 Design Considerations and Guidelines for Turning Operations

Several considerations are important in designing parts to be machined economically by turning operations. Machining, in general, should be avoided whenever possible, because:

1. Machining can take considerable time, thus increasing production costs

2. Material and latent material energy are wasted (see Section 40.5), even though chips can be recycled

3. Economic considerations are paramount, and it may be more economical to produce a component through shaping operations.

Table 23.9: Typical Production Rates for Various Machining Operations.

Operation	Rate
Turning	
Engine lathe	Very low to low
Tracer lathe	Low to medium
Turret lathe	Low to medium
Computer-controlled lathe	Low to medium
Single-spindle chuckers	Medium to high
Multiple-spindle chuckers	High to very high
Boring	Very low
Drilling	Low to medium
Milling	Low to medium
Planing	Very low
Gear cutting	Low to medium
Broaching	Medium to high
Sawing	Very low to low

Note: Production rates indicated are relative: *Very low* is about 1 or more parts per hour, *medium* is approximately 100 parts per hour, and *very high* is 1000 or more parts per hour.

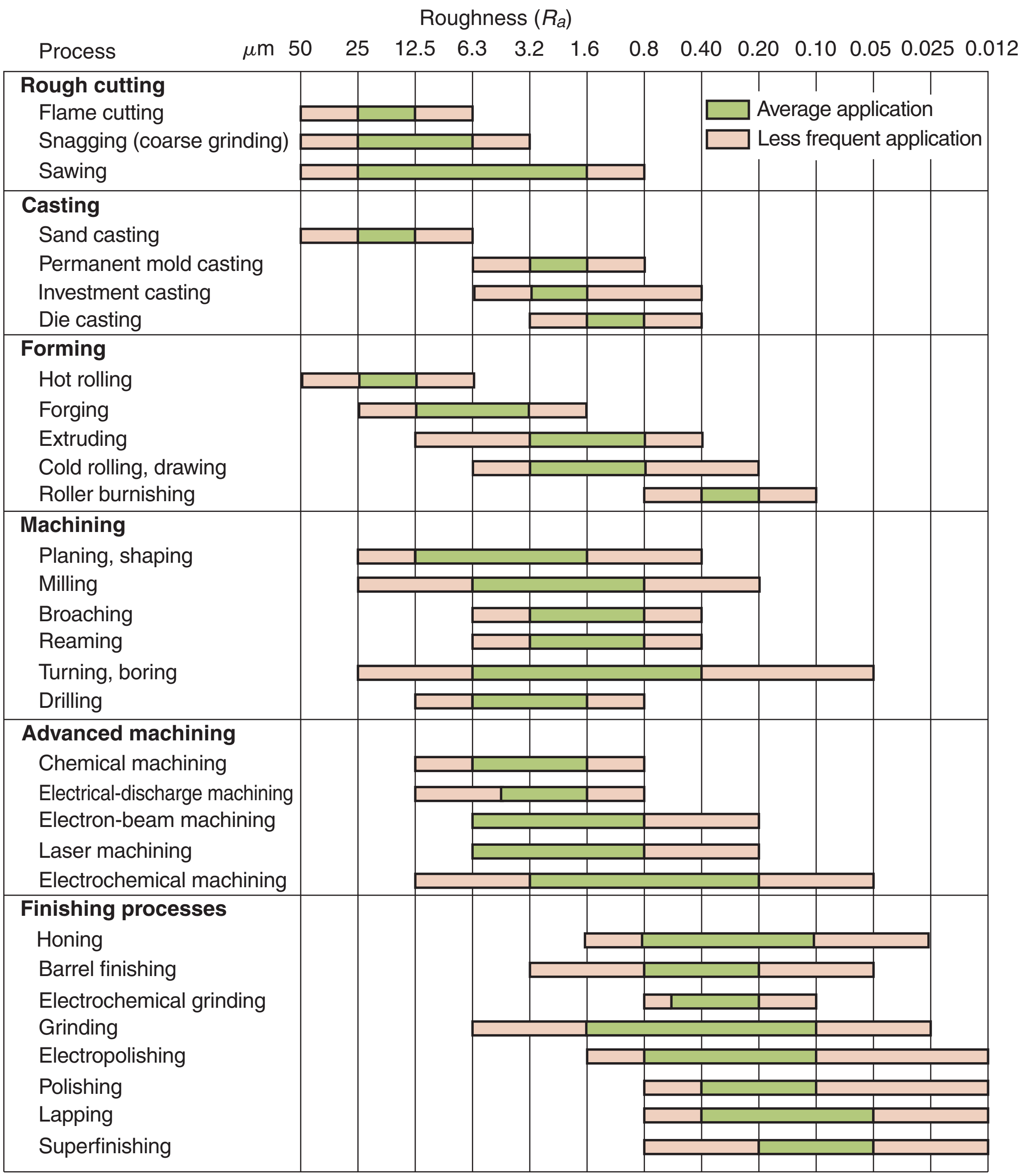


Figure 23.14: The range of surface roughnesses obtained in various processes; note the wide range within each group, especially in turning and boring.

The following general design guidelines should be considered for machining:

1. Parts should be designed so that they can be fixtured and clamped easily into workholding devices. Thin, slender workpieces are difficult to support properly and must be able to withstand clamping and cutting forces (see also *flexible fixturing*, Section 37.8).

2. The dimensional accuracy and surface finish specified should be as wide as permissible, but the part must still be able to function properly.

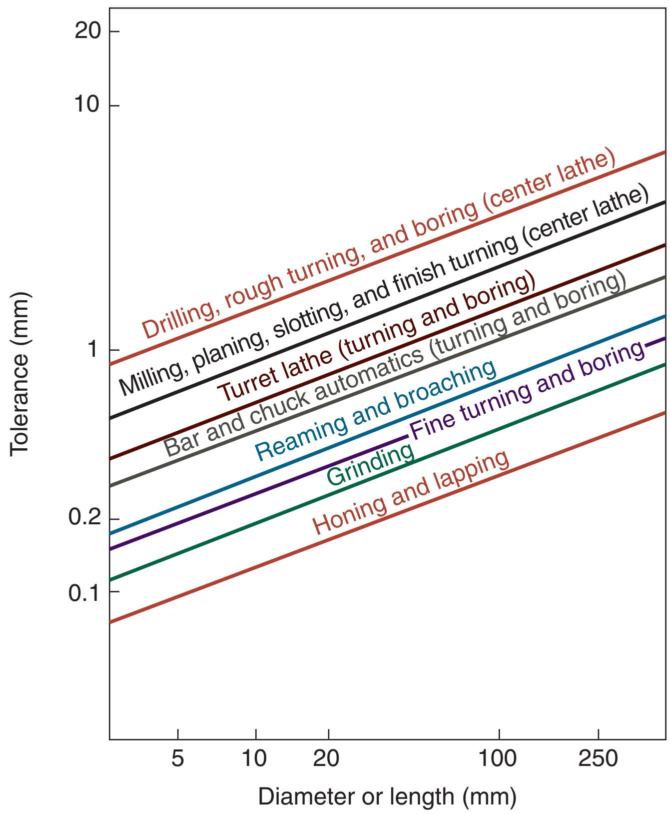


Figure 23.15: Range of dimensional tolerances in various machining processes as a function of workpiece size. Note that there is one order of magnitude difference between small and large workpieces.

3. Sharp corners, tapers, steps, and major dimensional variations in the part should be avoided.

4. Blanks to be machined should be as close to final part dimensions as possible, such as by near-net-shape forming, so as to reduce production cycle time.

5. Parts should be designed so that cutting tools can travel directly across the workpiece without any obstruction.

6. Design features should use commercially available, standard cutting tools, inserts, and toolholders whenever practicable.

7. Workpiece materials should preferably be selected for their machinability (Section 21.7).

Guidelines for Turning Operations. The following list outlines generally accepted guidelines for turning operations; see also Table 23.10 for probable causes of turning problems.

1. Minimize tool overhang

2. Support the workpiece rigidly

3. Use machine tools with high stiffness and high damping capacity

4. When tools begin to vibrate and chatter (Section 25.4), modify one or more of the process parameters, such as tool geometry, cutting speed, feed rate, depth of cut, and use of cutting fluid (see also adaptive control, Section 37.4, and stability lobes, Section 25.4).

Table 23.10: General Troubleshooting Guide for Turning Operations.

Problem	Probable causes
Tool breakage	Tool material lacks toughness, improper tool angles, machine tool lacks stiffness, worn bearings and machine components, machining parameters too high
Excessive tool wear	Machining parameters too high, improper tool material, ineffective cutting fluid, improper tool angles
Rough surface finish	Built-up edge on tool; feed too high; tool too sharp, chipped, or worn; vibration and chatter
Dimensional variability	Lack of stiffness of machine tool and work-holding devices, excessive temperature rise, tool wear
Tool chatter	Lack of stiffness of machine tool and work-holding devices, excessive tool overhang, machining parameters not set properly

23.3.7 Chip Collection Systems

The chips produced in machining operations must be collected and disposed of properly. The volume of chips produced can be very high, particularly in ultra-high-speed machining and high-removal-rate machining operations. For example, in a drilling operation on steel during which only 15 cm^3 of metal is removed, the loose bulk volume of the chips can, depending on chip type (see Section 21.2.1), be in the range of 600 to 12,000 cm^3. Likewise, the milling of 15 cm^3 of steel produces 450 to 700 cm^3 of chips, while cast iron produces 105 to 225 cm^3 of chips.

Also called **chip management**, the operation involves collecting chips from their source in an efficient manner and removing them from the work area. Long and stringy chips are more difficult to collect than short chips (produced by using tools with chipbreaker features; see Figs. 21.7 and 22.2). The type of chip produced must therefore be an integral aspect of the chip-collecting system.

Chips can be collected by any of the following methods:

- Using gravity, dropping them directly onto a steel conveyor belt
- Dragging the chips from a settling tank
- Using augers with feed screws
- Using magnetic conveyors, though for ferrous chips only
- Employing vacuum methods of chip removal and collection.

Modern machine tools are designed with automated chip-handling features.

There may be a considerable amount of *cutting fluid* residue on the chips produced. The cutting fluid and sludge can be separated from chips using *chip wringers* (centrifuges). Chip-processing systems usually require considerable floor space in a plant, and can cost from $60,000 for small shops to over $1 million for large facilities.

Collected chips may be recycled to reduce material and energy burden of machining (see Section 40.5). Prior to their removal from a manufacturing plant, the large volume of chips can be reduced to as little as one-fifth of their loose volume by *compacting* them into briquettes or by *shredding* them. Dry chips are more valuable for recycling, because of reduced environmental contamination. The method chosen for chip disposal depends on economics and on compliance with local, state, and federal regulations.

23.3.8 Thread Cutting

Screw threads are typically on the outside or inside of a cylindrical piece, and they may be a (*straight thread*) or a (*tapered thread*). Machine screws, bolts, and nuts have straight threads, as do threaded rods for such applications as the lead screw in lathes and a wide variety of machinery components (Fig. 23.2). Tapered threads are commonly used for water or gas pipes and plumbing supplies. Threads may be *right handed* or *left handed* and can have various profiles.

Although threads traditionally have been machined, they are now increasingly made by **thread rolling** (Section 13.5). Rolled fastener threads constitute the largest quantity of external threaded parts produced. Threads can also be integrally cast or molded, but there are limitations to their dimensional accuracy, surface finish, and minimum dimensions.

Threads can be machined either externally or internally, in a process called *thread cutting* or *threading*. External threads may also be cut with a *die* or by milling. Internal threads are typically produced by tapping, using a *tap* (Section 23.7). Threads may subsequently be ground, with high dimensional accuracy and surface finish, for such applications as power screw drives in machines.

Screw-thread Cutting on a Lathe. A typical thread-cutting operation on a lathe is shown in Fig. 23.16a. The cutting tool, the shape of which depends on the type of thread to be cut, is mounted on a holder and moved along the length of the workpiece by the lathe's lead screw. This movement is achieved by the engagement of a *split nut*, also called a *half nut*, inside the apron of the lathe (see Fig. 23.2).

The axial movement of the tool in relation to the workpiece rotation determines the *lead* of the screw thread (i.e., the axial distance moved in one complete revolution of the screw). For a fixed spindle speed, the slower the tool movement, the finer will be the thread. The cutting tool may be fed radially into the workpiece, thus cutting both sides of the thread at the same time, as in form cutting described earlier. However, this method usually produces a poor surface finish.

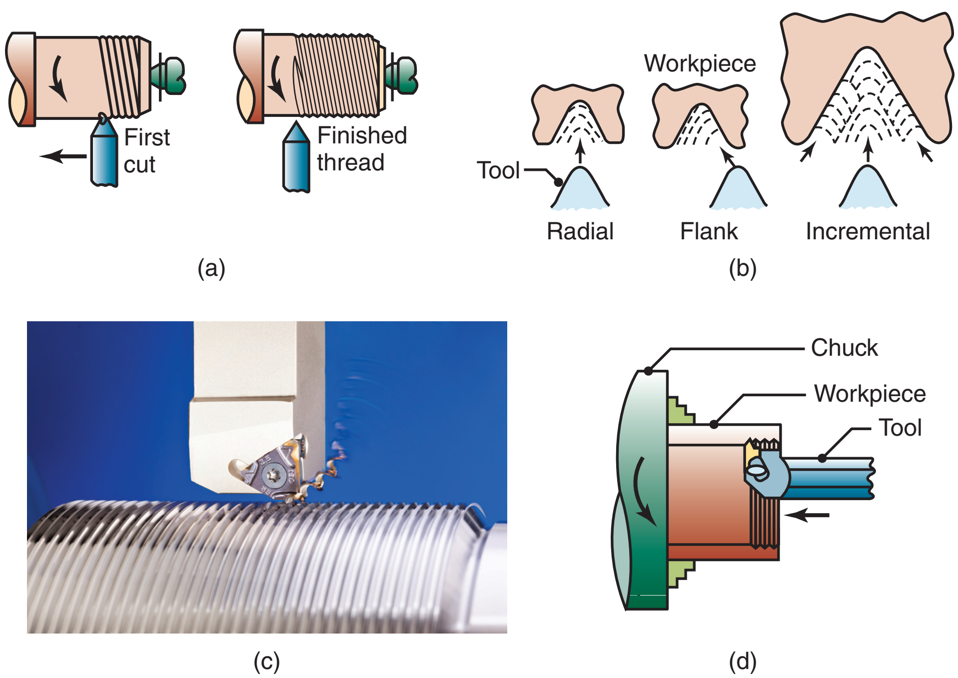


Figure 23.16: (a) Cutting screw threads on a lathe with a single-point cutting tool. (b) Cutting screw threads with a single-point tool in several passes, normally utilized for large threads. The small arrows in the figures show the direction of feed, and the broken lines show the position of the cutting tool as time progresses. In *radial cutting*, the tool is fed directly into the workpiece. In *flank cutting*, the tool is fed into the piece along the right face of the thread. In *incremental cutting*, the tool is fed first directly into the piece at the center of the thread, then at its sides, and finally into the root. (c) A typical coated-carbide insert in the process of cutting screw threads on a round shaft. (d) Cutting internal screw threads with a carbide insert. *Source:* (c) Courtesy of Iscar Metals, Inc.

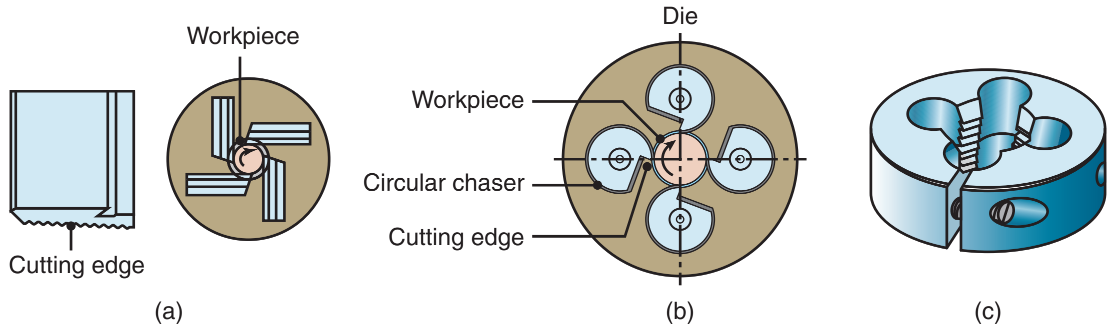


Figure 23.17: (a) Straight chasers for cutting threads on a lathe. (b) Circular chasers. (c) A solid threading die; the screw adjusts the gap.

A number of passes, in the sequence shown in Fig. 23.16b, generally are required to produce threads with good dimensional accuracy and surface finish. Figure 23.16c shows a carbide insert for screw-thread cutting (*threading insert*) for machining threads on a round shaft. Figure 23.16d shows an internal screw-thread cutting process. Except for small production runs, thread cutting largely has been replaced by other methods, such as thread rolling, automatic screw machining, and using CNC lathes.

The production rate in cutting screw threads can be increased with tools called *die-head chasers* (Fig. 23.17a and b), which typically have four cutters with multiple teeth and can be adjusted radially. After the threads are cut, the cutters open automatically (thus the alternative name *self-opening die heads*) by rotating them around their axes to allow the part to be removed. *Solid-threading dies* (Fig. 23.17c) also are available for cutting straight or tapered screw threads.

Design Considerations for Screw Thread Machining. The design considerations to be taken into account to produce high-quality and economical screw threads are the following:

- Rolled threads are generally preferable to cut threads; whenever practical, thread cutting should be avoided.

- Designs should allow for the termination of threads before they reach a shoulder on the part.

- Through-holes are preferable to blind holes when machining threads. The term *blind hole* refers to a hole that does not go through the thickness of the workpiece (see Fig. 23.1i). Internal threads in blind holes should have an unthreaded length at the bottom.

- Shallow, blind tapped holes should be avoided.

- Chamfers should be specified at the ends of threaded sections, to minimize finlike threads with burrs.

- Threaded sections should not be interrupted with slots, holes, or other discontinuities.

- Standard threading tooling and inserts should be used as much as possible.

- Thin-walled parts should have sufficient thickness and strength to resist clamping and cutting forces. A common rule of thumb is that the minimum engagement length of a fastener should be 1.5 times its diameter.

- Parts should be designed so that all cutting operations can be completed in one setup.

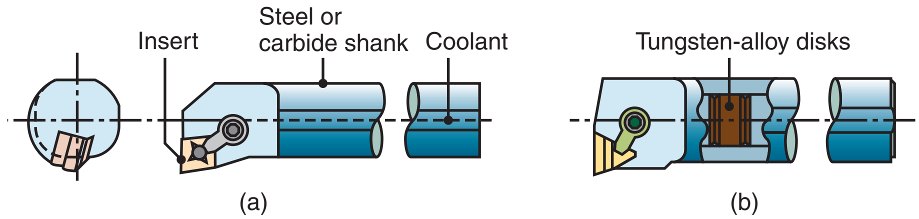


Figure 23.18: (a) Schematic illustration of a steel boring bar with a carbide insert; note the passageway in the bar for cutting fluid application. (b) Schematic illustration of a boring bar with tungsten-alloy *inertia disks*, sealed within the bar to counteract vibration and chatter while boring. This system has been found to be effective for boring-bar length-to-diameter ratios of up to 6.

23.4 Boring and Boring Machines

Boring enlarges a hole previously made by another process, or it produces circular internal profiles in hollow workpieces (Fig. 23.1h). The cutting tools are similar to those used in turning, and are mounted on a *boring bar* (Fig. 23.18a) in order to reach the full length of the bore. It is essential that the boring bar be sufficiently stiff to minimize tool deflection and vibrations, thus maintaining dimensional accuracy and surface finish. For this reason, a material with a high elastic modulus, such as tungsten carbide, is desirable. Boring bars have been designed and built with capabilities for damping vibration (Fig. 23.18b).

Boring operations on relatively small workpieces can be carried out on lathes, whereas large workpieces are machined on **boring mills**. These machine tools are either horizontal or vertical, and are capable of performing such operations as turning, facing, grooving, and chamfering.

In **horizontal boring machines**, the workpiece is mounted on a table that can move horizontally, in both the axial and radial directions. The cutting tool is mounted on a spindle that rotates in the headstock, and is capable of both vertical and longitudinal movements. Drills, reamers, taps, and milling cutters also can be mounted on the machine spindle. A **vertical boring mill** (Fig. 23.19) is similar to a lathe, has a vertical axis of workpiece rotation, and can accommodate workpieces as large as 2.5 m in diameter.

The cutting tool is usually a single point, made of M2 or M3 high-speed steel, or P10 (C7) or P01 (C8) carbide. It is mounted on the tool head, which is capable of vertical movement (for boring and turning) and

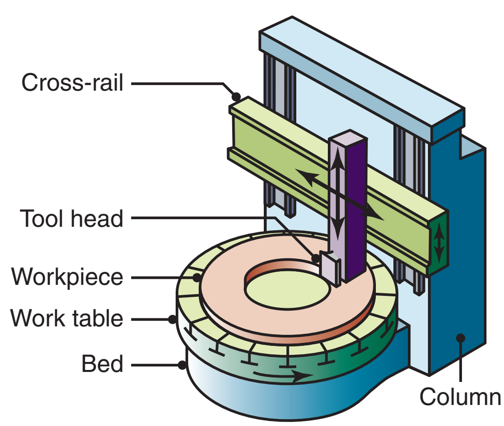


Figure 23.19: Schematic illustration of a vertical boring mill. Such a machine can accommodate workpiece sizes as large as 2.5 m in diameter.

radial movement (for facing), guided by the cross-rail. The head can be swiveled to make conical (tapered) holes. Cutting speeds and feeds for boring are similar to those for turning (see Table 23.9).

Boring machines are available with a variety of features. Machine capacities range up to 150 kW, and are available with CNC, allowing all movements of the machine to be programmed.

Design Considerations for Boring. Guidelines for economical boring operations are similar to those for turning; additionally, the following factors should be considered:

- Whenever possible, through holes rather than blind holes should be specified.
- The greater the length-to-bore-diameter ratio, the more difficult it is to hold dimensions, due to the deflections of the boring bar under cutting forces and the greater tendency for vibration and chatter.
- Interrupted internal surfaces, such as internal splines or radial holes that go through the thickness of the part, should be avoided.

23.5 Drilling, Drills, and Drilling Machines

Hole making is among the most important operations in manufacturing, and **drilling** is a common hole-making process. The cost of hole making, for example, is among the highest machining costs in automotive engine manufacturing. Other basic processes for making holes are punching (Section 16.2) and a variety of advanced machining processes (Chapter 27).

23.5.1 Drills

Drills typically have high length-to-diameter ratios (Fig. 23.20), hence they are capable of producing relatively deep holes. However, high ratios make drills somewhat flexible and prone to fracture or making inaccurate holes. Moreover, the chips produced can present significant difficulties in their disposal from the hole being drilled.

Drills generally leave a *burr* on the bottom surface of the part upon their breakthrough, often necessitating subsequent deburring operations (Section 26.8). Also, because of its rotary motion, drilling produces holes with walls with *circumferential* marks; in contrast, punched holes have *longitudinal* marks (see Fig. 16.5a). This difference can be significant in terms of the hole's fatigue properties (see Section 33.2).

The diameter of a hole produced by drilling is slightly larger than the drill diameter (*oversize*), as one can note by observing that a drill can easily be removed from the hole it has just produced, assuming temperature effects are not present. Hole oversize depends on the quality of the drill, the equipment used, and on the machining practices employed. Depending on their thermal properties, some metals and nonmetallic materials expand significantly due to the heat produced during drilling, thus the final hole diameter could be smaller than the drill diameter after the part cools down. For better surface finish and dimensional accuracy, drilled holes may be subjected to subsequent reaming and honing. The capabilities of drilling and boring operations are shown in Table 23.11.

Twist Drill. The most common drill is the conventional *standard-point twist drill* (Fig. 23.20a). The geometry of the drill point is such that the normal rake angle and velocity of the cutting edge vary with the distance from the center of the drill. The main features of this drill are, with typical ranges given in parentheses: (a) *point angle* (118° to 135°), (b) *lip-relief angle* (7° to 15°), (c) *chisel-edge angle* (125° to 135°), and (d) *helix angle* (15° to 30°). Two spiral grooves, called *flutes*, run the length of the drill, and the chips produced are guided upward through these grooves. The grooves also serve as passageways to enable the cutting fluid to reach the cutting edges. Some drills have internal longitudinal holes (see Fig. 23.23a), through which cutting fluids are forced, thus improving lubrication and cooling and washing away the chips. Drills are available with a **chip-breaker** feature on the cutting edges. This feature is important in automated machinery, where continuous removal of long chips without operator assistance is essential.

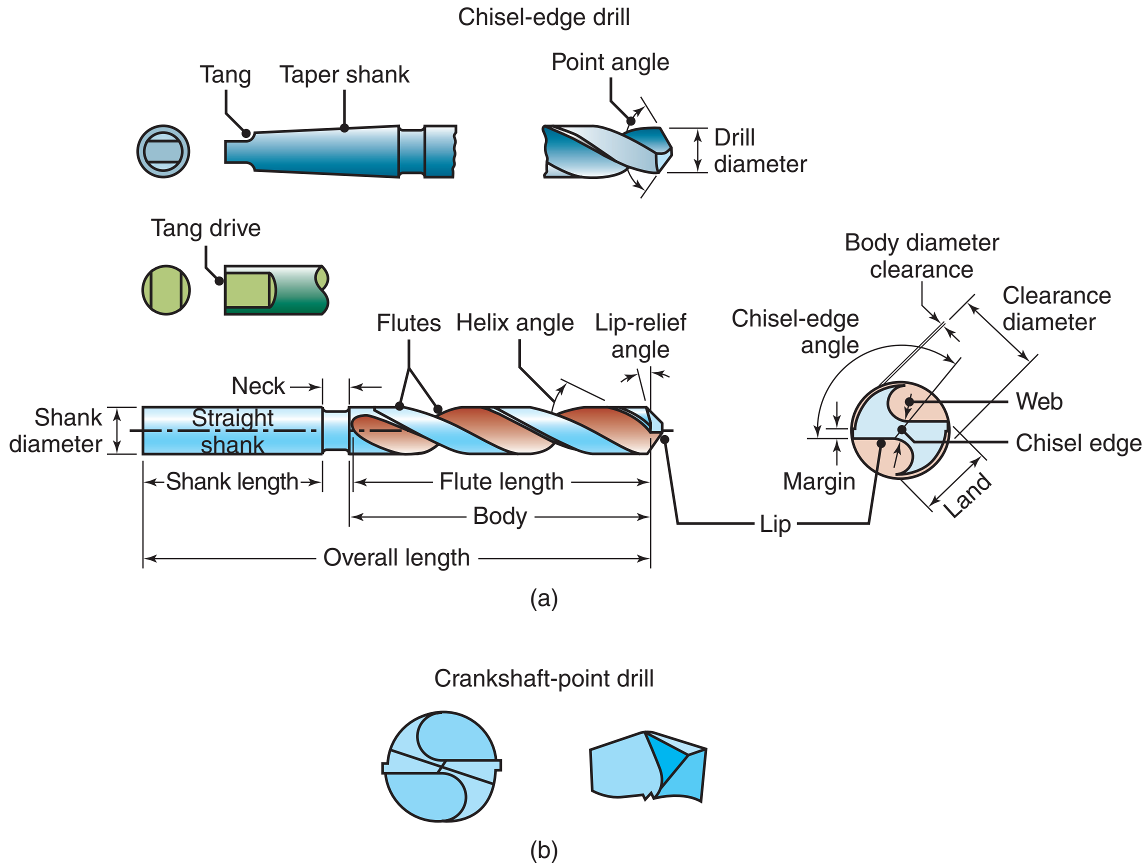


Figure 23.20: Two common types of drills: (a) Chisel-edge drill. The function of the pair of margins is to provide a bearing surface for the drill against walls of the hole as it penetrates the workpiece. Drills with four margins (*double-margin*) are available for improved guidance and accuracy. Drills can have chip-breaker features. (b) Crankshaft drill. These drills have good centering ability, and because the chips tend to break up easily, crankshaft drills are suitable for producing deep holes.

The various angles on a drill have been designed to produce accurate holes, minimize drilling forces and torque, and optimize drill life. Small changes in drill geometry can have a significant effect on a drill's performance, particularly in the chisel-edge region, which accounts for about 50% of the thrust force in drilling. Too small a lip relief angle (Fig. 23.20a) increases the thrust force, generates excessive heat, and increases drill wear. By contrast, too large an angle can cause chipping or breaking of the cutting edge.

Table 23.11: General Capabilities of Drilling and Boring Operations.

		Hole depth/diameter	
Cutting tool	**Diameter range (mm)**	**Typical**	**Maximum**
Twist drill	0.5–150	8	50
Spade drill	25–150	30	100
Gun drill	2–50	100	300
Trepanning tool	40–250	10	100
Boring tool	3–1200	5	8

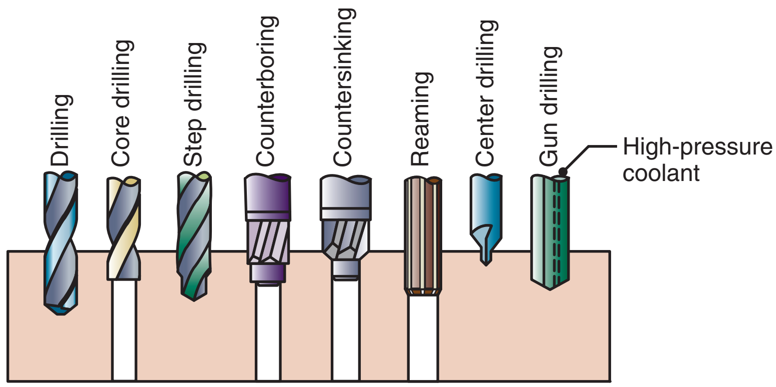


Figure 23.21: Various types of drills and drilling and reaming operations.

Several other drill-point geometries have been developed to improve drill performance and increase the penetration rate required for high rate production. Special grinding techniques and equipment are used to produce these geometries.

Other Types of Drills. Several types of drills are shown in Fig. 23.21. A *step drill* produces a hole with two or more different diameters. A *core drill* is used to make an existing hole larger. *Counterboring* and *countersinking drills* produce depressions on the workpiece surface to accommodate the heads of screws and bolts below the surface. A *center drill* is short and is used to produce a hole at one end of a round stock, so that it can be mounted between the centers of the headstock and the tailstock on a lathe (Fig. 23.2). A *spot drill* is used to start a hole at the desired location on a surface.

Spade drills (Fig. 23.22a) have removable tips or bits, and are used to produce large-diameter deep holes. Because of the absence of flutes in the body of the drill, these drills have the advantages of higher stiffness, ease of grinding the cutting edges, and lower cost. A similar drill is the *straight-flute drill* (Fig. 23.22b).

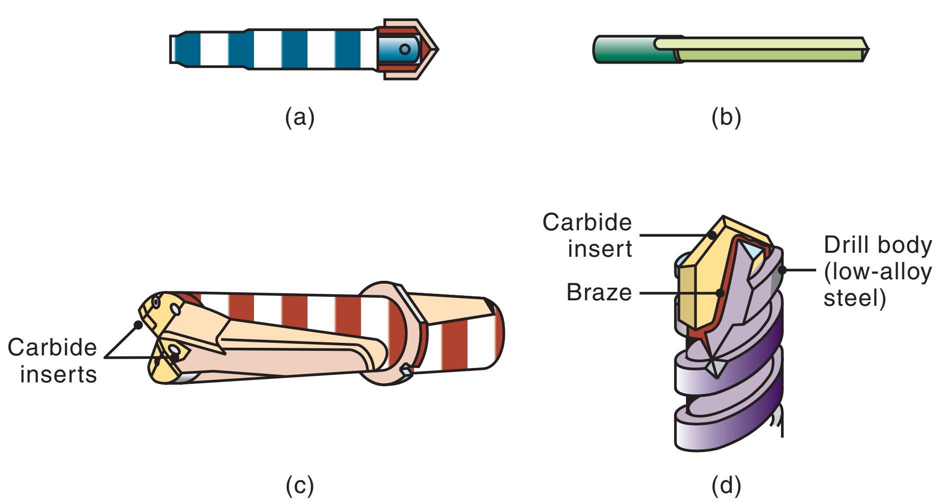


Figure 23.22: Various types of drills. (a) Spade drill; (b) straight-flute drill; (c) drill with indexable carbide inserts; (d) drill with brazed-carbide tip.

Solid carbide and carbide-tipped drills (Fig. 23.22c and d) are made for drilling hard materials, such as cast irons, high-temperature metals, abrasive materials such as concrete and brick (called *masonry drills*), and composite materials containing abrasive fiber reinforcements, such as glass and graphite.

Gun Drilling. Gun drilling is used for drilling deep holes; it requires a special drill, as shown in Fig. 23.23. The depth-to-diameter ratios of holes produced can be more than 300:1. Thrust force (the radial force that tends to deflect the drill sideways) is balanced by *bearing pads* on the drill that slide along the inside surface of the hole. Consequently, a gun drill is *self-centering*, an important feature in drilling straight, deep holes. A variation of this process is **gun trepanning** (described below), which uses a cutting tool similar to a gun drill, except that the tool has a central hole.

Cutting speeds in gun drilling are usually high, and feeds are low, and tolerances are typically 0.025 mm. The cutting fluid is forced, under high pressure, through a longitudinal hole in the body of the drill (Fig. 23.22a). In addition to cooling and lubricating the workpiece, the fluid flushes out chips that otherwise would be trapped in the deep hole being drilled.

Trepanning. In *trepanning* (from the Greek *trypanon*, meaning boring a hole). The cutting tool (Fig. 23.24a) produces a hole by removing a disk-shaped piece (*core*), usually from flat plates. A hole is thus produced without reducing all of the material to chips, as is the case in drilling. The trepanning process can be used to make disks up to 250 mm in diameter, from flat sheets, plates, or structural members such as I-beams. It also can be used to make circular grooves as seats for O-rings (similar to Fig. 23.1f). Trepanning can be carried out on lathes, drill presses, or other machine tools, using single-point or multipoint tools, as shown in Fig. 23.24b.

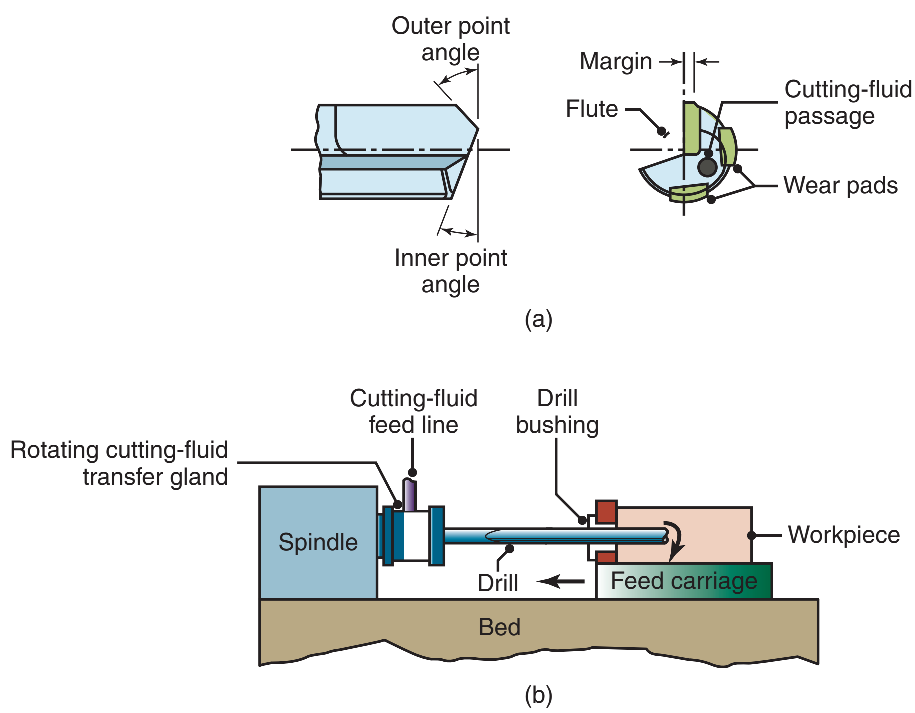


Figure 23.23: (a) A gun drill, showing various features. (b) Schematic illustration of the gun-drilling operation.

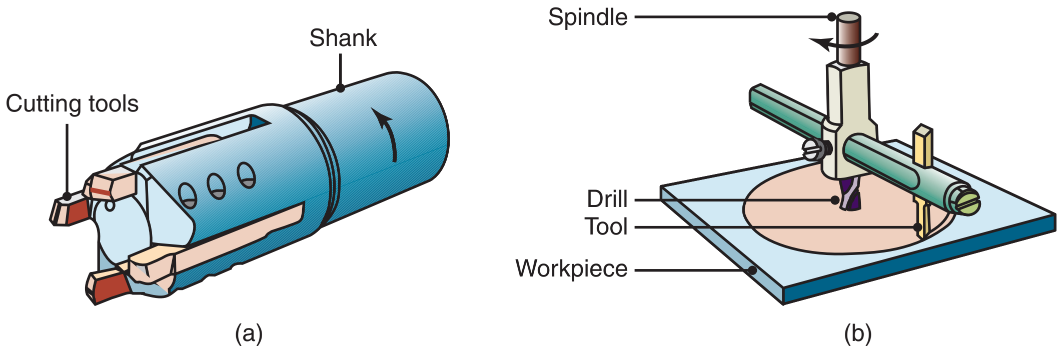


Figure 23.24: (a) Trepanning tool. (b) Trepanning with a drill-mounted single cutter.

23.5.2 Material-removal Rate in Drilling

The *material-removal rate* (MRR) in drilling is the volume of material removed per unit time; thus for a drill with diameter D, the cross-sectional area of the drilled hole is $\pi D^2/4$. The velocity of the drill perpendicular to the workpiece is the product of the feed, f (the distance the drill penetrates per unit revolution), and the rotational speed, N, where $N = V/\pi D$. Thus,

$$\text{MRR} = \left(\frac{\pi D^2}{4}\right) fN. \tag{23.4}$$

23.5.3 Thrust Force and Torque

The *thrust force* in drilling acts perpendicular to the hole axis; if this force is excessive, it can cause the drill to bend or break. An excessive thrust force also can distort the workpiece itself, particularly if it does not have sufficient stiffness, as is the case for thin sheet-metal structures (Chapter 16).

The thrust force depends on factors such as (a) the strength of the workpiece material, (b) feed, (c) rotational speed, (d) drill diameter, (e) drill geometry, and (f) cutting fluids. Forces typically range from a few newtons, for small drills, to as high as 100 kN, for drilling high-strength materials using large drills.

Torque. The *torque* in drilling, essential for estimating the power requirement, is difficult to calculate because of the many factors involved. It can be estimated from the data given in Table 21.2 by noting that the power dissipated during drilling is the product of the torque and the rotational speed. It is also equal the product of specific energy and material removal rate. Torque in drilling can be as high as 4000 N-m.

Example 23.4 Material-removal Rate and Torque in Drilling

Given: A hole is being drilled in a block of magnesium alloy with a 10-mm drill bit at a feed of 0.2 mm/rev and with the spindle running at $N = 800$ rpm.

Find: Calculate the material-removal rate and the torque on the drill.

Solution: The material-removal rate is calculated from Eq. (23.4):

$$\text{MRR} = \left[\frac{(\pi)(10)^2}{4}\right](0.2)(800) = 12,570\ \text{mm}^3/\text{min} = 210\ \text{mm}^3/\text{s}.$$

Referring to Table 21.2, an average unit power of 0.5 W-s/mm^3 is used for magnesium alloys. The power required is then

$$\text{Power} = (210)(0.5) = 105 \text{ W}.$$

Power is the product of the torque on the drill and the rotational speed, which in this case is $(800)(2\pi)60 = 83.8$ radians per second. Noting that W = J/s and J = N-m,

$$T = \frac{105}{83.8} = 1.25 \text{ N-m}.$$

23.5.4 Drill Materials and Sizes

Drills are usually made of high-speed steels (M1, M7, and M10), solid carbides, or with carbide tips [typically made of K20 (C2) carbide] and brazed over a steel shank (Fig. 23.22c and d). Drills are commonly coated with titanium nitride or titanium carbonitride for increased wear resistance (see Section 22.5). Polycrystalline-diamond-coated drills are used for producing holes for fasteners in fiber-reinforced plastic structures; several thousand holes can be drilled with little damage to the workpiece.

Although there are continued developments, standard twist-drill sizes consist basically of the following series:

- **Numerical:** No. 97 (0.0059 in.) to No. 1 (0.228 in.)
- **Letter:** A (0.234 in.) to Z (0.413 in.)
- **Fractional:** Straight shank from $\frac{1}{64}$ to $1\frac{1}{4}$ (in $\frac{1}{64}$-in. increments) to $1\frac{1}{2}$ in. (in $\frac{1}{32}$-in. increments), and larger drills in larger increments. Taper shank from $\frac{1}{8}$ to $1\frac{3}{4}$ (in $\frac{1}{64}$ increments) to 3.5 in. (in $\frac{1}{16}$-in. increments)
- **Millimeter:** From 0.05 mm in increments of 0.01 mm.

23.5.5 Drilling Practice

Drills and similar hole making tools usually are held in *drill chucks*, which may be tightened with keys or different mechanisms. Special chucks and collets, with various quick-change features that do not require stopping the spindle, are available for use on production machinery.

Because it does not have a centering action, a drill tends to *walk* on the workpiece surface at the beginning of an operation, a problem particularly severe with small-diameter long drills that can bend and break. To start a hole properly, the drill bit should be guided, using fixtures such as a bushing, to keep it from excessively deflecting laterally. A small starting dimple or hole can be made with a punch or center drill or the drill point may be ground to an S shape (called *helical* or *spiral point*). This shape has a self-centering characteristic, thus eliminating the need for center drilling, and produces accurate holes and with improved drill life. These factors are particularly important in automated production with CNC machines, in which the usual practice is to use a spot drill. To keep the drill more centered, the point angles of the spot drill and of the drill are matched.

Drilling Recommendations. Recommended ranges for drilling speeds and feeds are given in Table 23.12. The speed is the *surface speed* of the drill at its periphery; thus, a 12.7-mm drill rotating at 300 rpm has a surface speed of

$$V = \left(\frac{12.7}{2} \text{ mm}\right)(300 \text{ rev/min})(2\pi \text{ rad/rev})\left(\frac{1}{1000} \text{ m/mm}\right) = 12 \text{ m/min}.$$

Table 23.12: General Recommendations for Speeds and Feeds in Drilling.

Workpiece material	Surface speed m/min	Drill diameter Feed, mm/rev 1.5 mm	12.5 mm	Speed, rpm 1.5 mm	12.5 mm
Aluminum alloys	30–120	0.025	0.30	6400–25,000	800–3000
Magnesium alloys	45–120	0.025	0.30	9600–25,000	1100–3000
Copper alloys	15–60	0.025	0.25	3200–12,000	400–1500
Steels	20–30	0.025	0.30	4300–6400	500–800
Stainless steels	10–20	0.025	0.18	2100–4300	250–500
Titanium alloys	6–20	0.010	0.15	1300–4300	150–500
Cast irons	20–60	0.025	0.30	4300–12,000	500–1500
Thermoplastics	30–60	0.025	0.13	6400–12,000	800–1500
Thermosets	20–60	0.025	0.10	4300–12,000	500–1500

Note: As hole depth increases, speeds and feeds should be reduced. The selection of speeds and feeds also depends on the specific surface finish required.

In drilling holes smaller than 1 mm, rotational speeds can range up to 30,000 rpm, depending on the workpiece material. The *feed* in drilling is the distance the drill travels into the workpiece per revolution. For example, Table 23.11 recommends that, for most workpiece materials, a 1.5 mm drill should have a feed of 0.025 mm/rev. If the speed column in the table indicates that the drill should rotate at, say, 2000 rpm, then the drill should travel into the workpiece at a linear speed of (0.025 mm/rev)(2000 rev/min) = 50 mm/min.

Chip removal in drilling can be difficult, especially for deep holes in soft, ductile workpieces. Chips generally are removed by being forced up the flutes, and the force and the torque can become excessive if the flutes become loaded with chips. The drill should therefore be retracted periodically (called *pecking*), to remove chips that may have accumulated in the flutes; otherwise, the drill may break. A general guide to the probable causes of problems in drilling operations is given in Table 23.13.

Drill Reconditioning. Drills are *reconditioned* by grinding them either manually or with special fixtures. Proper **reconditioning** of drills is important, particularly with automated manufacturing on CNC machines. Hand grinding is difficult, and requires considerable skill in order to produce symmetric cutting edges. Grinding on fixtures is accurate and is done on special computer-controlled grinders. Worn and dull coated drills also can be recoated in special facilities.

Measuring Drill Life. Drill life, as well as tap life (see Section 23.7), is typically measured by the number of holes drilled before they become dull and have to be reconditioned or replaced. *Drill life* can be determined experimentally by first clamping a block of material on a suitable dynamometer or force transducer. Then, a number of holes are drilled while the torque or thrust force is monitored; drill life is the number of holes drilled until these two quantities begin to increase, indicating that the drill is becoming dull. Monitoring vibration and acoustic emissions (Section 21.5.4) also can be used to determine drill life.

Table 23.13: General Troubleshooting Guide for Drilling Operations.

Problem	Probable causes
Drill breakage	Dull drill, drill seizing in hole because of chips clogging flutes, feed too high, lip relief angle too small
Excessive drill wear	Cutting speed too high, ineffective cutting fluid, rake angle too high, drill burned and strength lost when drill was sharpened
Tapered hole	Drill misaligned or bent, lips not equal, web not central
Oversize hole	Same as previous entry, machine spindle loose, chisel edge not central, side force on workpiece
Poor hole surface finish	Dull drill, ineffective cutting fluid, welding of workpiece material on drill margin, improperly ground drill, improper alignment

23.5.6 Drilling Machines

The most common machine is the **drill press**, the major components of which are illustrated in Fig. 23.25a. The workpiece is placed on an adjustable table, either by clamping it directly into the slots and holes on the table or by using a vise, which itself is clamped to the table. The drill is lowered manually by a handwheel or by power feed at preset rates.

Drill presses are usually designated by the largest workpiece diameter that can be accommodated on the table, and typically range from 150 to 1250 mm. In order to maintain proper cutting speeds at the cutting edges of drills, the spindle speed of the machines has to be adjustable for different drill sizes. Adjustments are made by means of pulleys, gearboxes, or variable-speed motors.

The types of drilling machines range from simple bench-type drills, used to drill small holes, to large *radial drills* (Fig. 23.25b) that can accommodate large workpieces. The distance between the column and the spindle center can be as much as 3 m. The drill head of *universal drilling machines* can also be swiveled to drill holes at an angle. Modern drilling machines include numerically controlled three-axis machines, in which all operations are performed automatically and in their desired sequence, with the use of a turret (Fig. 23.26), which can hold several different drilling tools.

Also used for boring and counterboring operations, drilling machines with multiple spindles (**gang drilling**) are used for high-production-rate operations, and are capable of drilling holes of varying sizes, depths, and locations in one cycle. *Numerical-control turret drilling machines* are also available.

Workholding devices for drilling are essential to ensure that the workpiece is located and clamped properly to keep it from slipping or rotating during drilling. These devices are available in a variety of designs, with important features such three-point locating, for accuracy, and three-dimensional work holding, for secure fixturing (see also Section 37.8).

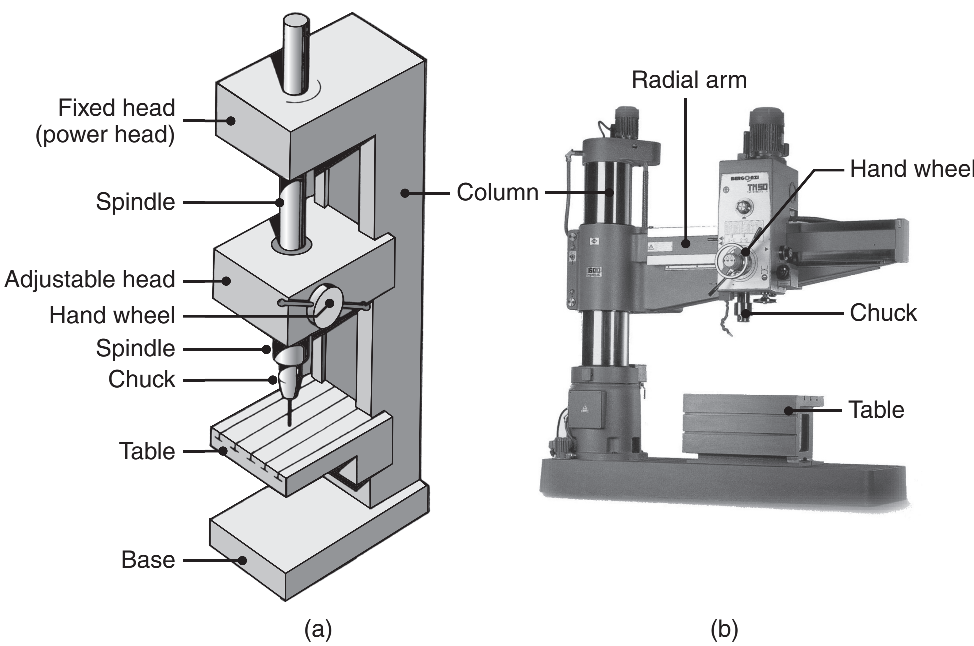


Figure 23.25: (a) Schematic illustration of the components of a vertical drill press. (b) A radial drilling machine. *Source:* (b) Courtesy of Willis Machinery and Tools.

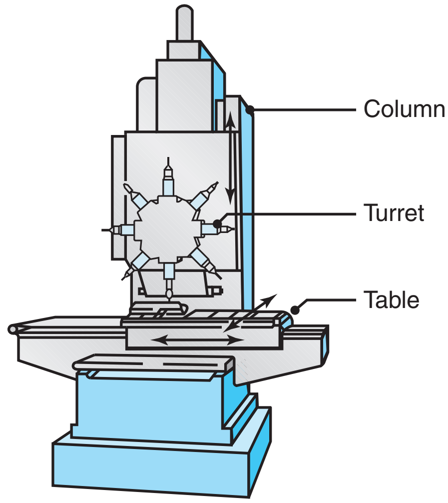


Figure 23.26: A three-axis computer-numerical-control drilling machine. The turret holds as many as eight different tools, such as drills, taps, and reamers.

23.5.7 Design Considerations for Drilling

The basic design guidelines for drilling are:

- Designs should allow holes to be drilled preferably on flat surfaces and perpendicular to the drill motion; otherwise, the drill tends to deflect and the hole will not be located accurately. The exit surfaces for the drill also should be flat.
- Design of hole bottoms should match standard drill-point angles, whenever possible; thus, flat bottoms or odd shapes should be avoided.
- When multiple holes are required, they should all have the same diameter, whenever practical, to avoid unnecessary tool changes.
- Excessively deep holes should be avoided, and length-to-diameter ratios of three or less should be specified whenever possible, although a ratio of 8:1 ratio is feasible.
- Through holes are preferred over blind holes.
- If holes with large diameters are specified, the workpiece should have a preexisting hole, preferably made during fabrication of the part itself, such as by casting, powder metallurgy, or forming.
- Dimples should be provided when preexisting holes are not practical to make, to reduce the tendency for the drill to walk.
- Parts should be designed so that all drilling can be performed with a minimum of fixturing and without the need to reposition the workpiece.
- Blind holes must be drilled deeper than subsequent reaming or tapping operations that may have to be performed, typically by an amount at least one-fourth of the hole diameter.

23.6 Reaming and Reamers

Reaming is an operation used for (a) making an existing hole dimensionally more accurate than can be achieved by drilling alone and (b) improving its surface finish. The most accurate holes in workpieces generally are produced by the following sequence of operations:

1. Centering
2. Drilling
3. Boring
4. Reaming

For even better accuracy and surface finish, holes may be *burnished* or internally *ground* and *honed* (Sections 26.4 and 26.7).

A *reamer* (Fig. 23.27a) is a multiple-cutting-edge tool with straight or helically fluted edges that remove very little material. For soft metals, a reamer typically removes a minimum of 0.2 mm on the diameter of a drilled hole; for harder metals, about 0.13 mm is removed. Attempts to remove smaller layers can be detrimental, as the reamer may be damaged or the hole surface may become burnished (see also Fig. 21.22 as an analogy); in this case, honing would be preferred. In general, reamer speeds are one-half those of the same-size drill and three times the feed rate.

Hand reamers are straight or have a tapered end in the first third of their length. Various *machine reamers*, also called *chucking reamers* because they are mounted in a chuck and machine-operated, are available in two types: (a) *Rose reamers* have cutting edges with wide margins and no relief (Fig. 23.27a). (b) *Fluted reamers* have small margins and relief, with a rake angle of about 5°; they usually are used for light cuts.

Shell reamers are hollow and mounted on an arbor, and are generally used for holes larger than 20 mm. *Expansion reamers* are adjustable for small variations in hole size; they also compensate for wear of the reamer's cutting edges. *Adjustable reamers* (Fig. 23.27b) can be set for specific hole diameters, and are therefore versatile.

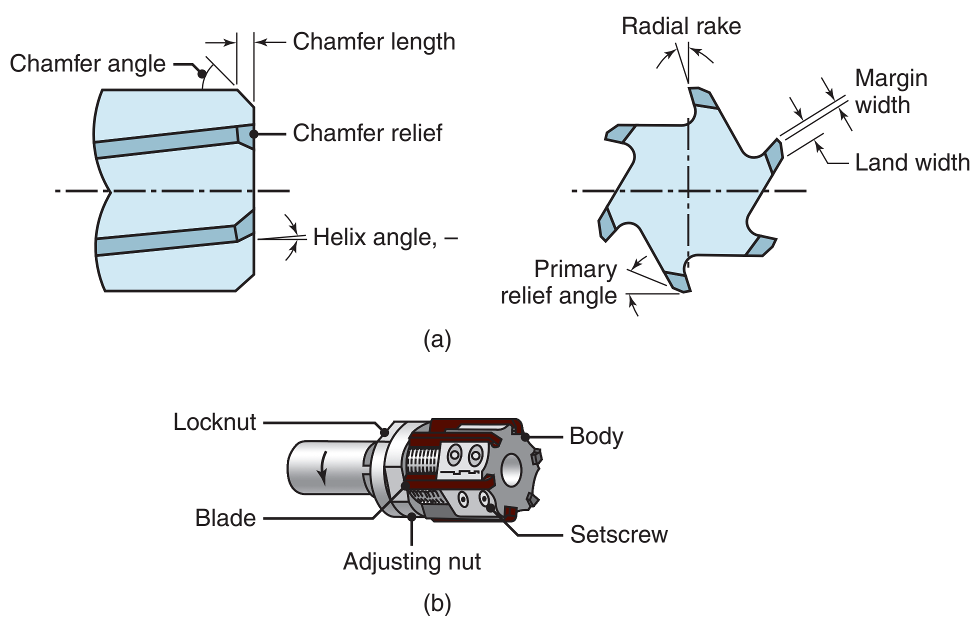


Figure 23.27: (a) Terminology for a helical reamer. (b) Inserted-blade adjustable reamer.

Reamers may be held rigidly, as in a chuck, or they may *float* in their holding fixtures, to ensure alignment or to be *piloted* in guide bushings placed above and below the workpiece. A further development in reaming consists of the *dreamer*, a tool that combines drilling and reaming. The tip of the tool first produces a hole by drilling; the rest of the same tool then performs a reaming operation. A similar development involves drilling and tapping in one stroke, using a single tool.

Reamers are typically made of high-speed steels (M1, M2, and M7) or solid carbides (K20, C2), or have carbide cutting edges. Reamer maintenance and reconditioning are important for hole accuracy and surface finish.

23.7 Tapping and Taps

Internal threads can be produced by *tapping*, a *tap* being a chip-producing threading tool with multiple cutting teeth (Fig. 23.28a). Taps generally are available with two, three, or four flutes. The most common production tap is the two-flute spiral-point tap; it forces the chips into the hole so that the tap needs to be retracted only at the end of the cut. Three-fluted taps are stronger, because more material is available in the flute. Tap sizes range up to 100 mm; larger threads can be machined in a milling machine or a machining center (see Fig. 24.2f).

Tapered taps are designed to reduce the torque required for the tapping of through holes. *Bottoming taps* are for tapping blind holes to their full depth. *Collapsible taps* are used in large-diameter holes; after tapping has been completed, the tap is collapsed mechanically and is removed from the hole without having to rotate them in the hole, as do regular taps.

Chip removal can be a significant problem during tapping, because of the small clearances in the tap. If chips aren't removed properly, the torque increases significantly and can break the tap. The use of a cutting fluid and the periodic reversal and removal of the tap from the hole are effective means of chip removal. For higher tapping productivity, drilling and tapping can be combined in a single operation (*drapping*) in a single tool. The tool has a drilling section at its tip, followed by a tapping section.

Tapping may be done by hand or on machines, such as (a) drilling machines, (b) lathes, (c) automatic screw machines, and (d) vertical CNC milling machines, which combine the correct relative rotation and the longitudinal feed. Special tapping machines are available, with features for multiple tapping operations. Multiple-spindle tapping heads are used extensively, particularly in the automotive industry, where 30% to 40% of machining operations involve tapping holes. One simple method of automatic tapping of nuts is shown in Fig. 23.28b.

Tap life can be determined with the same technique for measuring drill life. With proper lubrication, tap life may be as high as 10,000 holes. Taps usually are made of high-speed steels (M1, M2, M7, and M10). Productivity in tapping operations can be improved by *high-speed tapping*, with surface speeds as

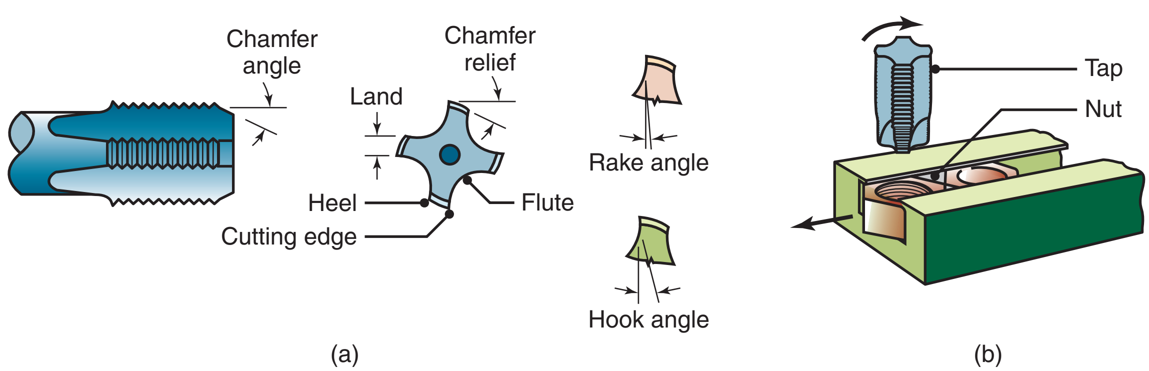


Figure 23.28: (a) Terminology for a tap. (b) Tapping of steel nuts in production.

high as 100 m/min. *Self-reversing* tapping systems are now in use with modern computer-controlled machine tools. Operating speeds can be as high as 5000 rpm, although actual cutting speeds in most applications are considerably lower. Cycle times typically are on the order of 1 to 2 seconds.

Some tapping systems have capabilities for directing the cutting fluid to the cutting zone through the spindle and a hole in the tap, which also helps flush the chips out of the hole being tapped. **Chipless tapping** is a process of internal thread rolling using a forming tap (Section 13.5).

Case Study 23.2 Bone Screw Retainer

A cervical spine implant is shown in Fig. 23.29a. In the event that a patient requires cervical bone fusion at one or more vertebral levels, this implant can act as an internal stabilizer by decreasing the amount of motion in the region, and thereby help promote a successful fusion. The plate is affixed to the anterior aspect of the spine, using bone screws that go through the plate and into the bone. The undersurface of the plate has a very rough surface that helps hold the plate in place while the bone screws are being inserted. One concern with this type of implant is the possibility of the bone screws loosening with

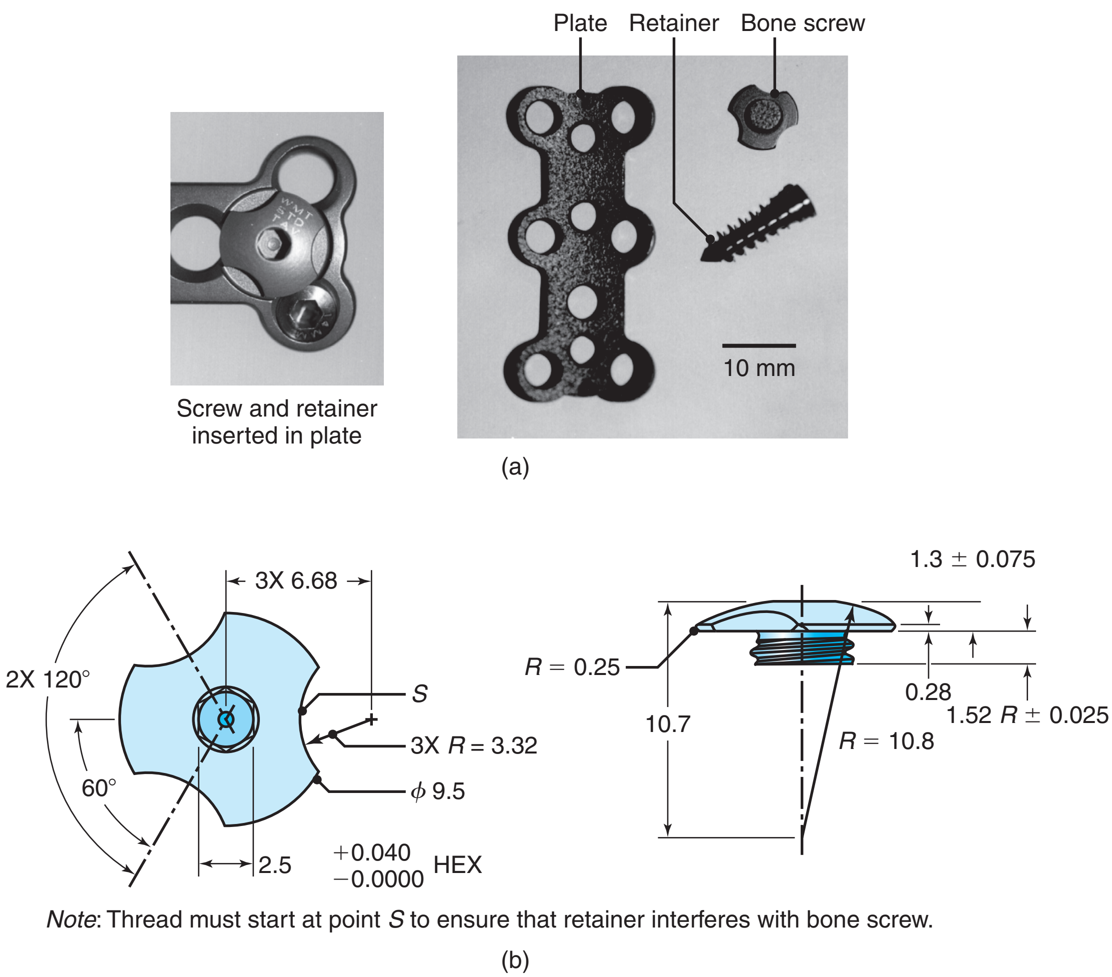


Figure 23.29: A cervical spine implant. (All dimensions in mm)

time, due to normal and repetitive loading from the patient. In extreme cases, this can result in a screw backing out, with the head of the screw no longer being flush with the plate, a condition that obviously is undesirable. The implant described here uses a retainer to prevent the bone screw from backing out away from the plate, as shown in the left half of Fig. 23.29b.

The retainer has several design features that are essential for it to function correctly, and without complicating the surgical procedure. To ease its use in surgery, the plate is provided with the retainers already in place, with the circular notches aligned with the bone screw holes. This arrangement allows the surgeon to insert the bone screws without interference from the retainer. Once the screws are inserted, the surgeon turns the retainer a few degrees so that each screw head is then captured. In order to ensure the retainer's proper orientation in the plate, the thread of its shank must start in the same axial location as point S in Fig. 23.29b.

The manufacturing steps followed to produce this part are shown in Fig. 23.29b. First, a 12.7-mm diameter Ti-6Al-4V rod is placed in a CNC lathe and faced. Then the threaded area is turned to the diameter necessary to machine the threads. The thread is turned on the shank, but over a longer length than is ultimately required, because of difficulties in obtaining high-quality threads at the start of machining. The cap then is turned to the required diameter, and a 2.5-mm radius is machined on the underside of the head. The part is removed, inspected, and placed in another CNC lathe, where it is faced to the specified length. The spherical radius in the cap is then machined, the center hole is drilled, and the hex head is broached. The cap is removed and inspected, and if the desired length has not been achieved, the cap is lapped (Section 26.7) to the final dimension.

At this point, the retainer is placed in a CNC milling machine, using a specially designed fixture that consists basically of a tapered and threaded hole. By carefully applying a predetermined torque on the retainer when placing it into the fixture, the starting location of the threads can be controlled accurately. Once the cap is located in the fixture, the three circular notches are machined as per the drawing. The retainer is then deburred by tumbling to remove all sharp corners, and the bottom is grit blasted to match that of the underside of the plate. Finally, the parts are anodized (Section 34.10) and passivated to obtain the desired biocompatibility.

Source: Courtesy of J. Mankowski and B. Pyszka, Master Metal Engineering Inc., and C. Lyle and M. Handwerker, Wright Medical Technology, Inc.

Summary

- Machining processes that typically produce external and internal circular profiles are turning, boring, drilling, and tapping. Because of the three-dimensional nature of these operations, chip movement from the cutting zone and its control are important considerations. Chip removal can be a significant problem, especially in drilling and tapping, and can lead to tool breakage.

- Optimization of each machining operation requires an understanding of the interrelationships among design parameters (such as part shape, dimensional accuracy, and surface finish) and process parameters (cutting speed, feed, and depth of cut), tool material and shape, the use of cutting fluids, and the sequence of operations to be performed.

- The parts to be machined may have been produced by casting, forging, extrusion, or powder metallurgy. The closer the blank to be machined to the final shape desired (near-net shape), the fewer the number and extent of the subsequent machining processes required.

Key Terms

Automatic bar machine
Back rake angle
Bed
Boring
Boring mill
Carriage
Chip management
Chuck
Collet
Cutting-edge angle
Drilling
Drill life
Drill press
Engine lathe
Face plate
Facing
Feed force
Feed rod
Finishing cuts
Form tools
Gun drilling
Headstock
Hole making
Knurling
Lathes
Lead screw
Mandrel
Material-removal rate
Nose radius
Parting
Power chuck
Rake angle
Reamer
Reaming
Reconditioning
Relief angle
Roughing cuts
Screw threads
Side rake angle
Tailstock
Tapping
Threading
Trepanning
Turning
Turret lathe
Twist drill

Bibliography

ASM Handbook, Vol. 16: **Machining**, ASM International, 1989.
Byers, J.P. (ed.), **Metalworking Fluids**, 2nd ed., CRC Press, 2006.
Hoffman, E.G., **Jigs and Fixture Design**, 5th ed., Industrial Press, 2003.
Joshi, P.H., **Machine Tools Handbook**, McGraw-Hill, 2008.
Krar, S.F., Gill, A.R., and Smid, P., **Technology of Machine Tools**, 7th ed., McGraw-Hill, 2010.
Liang, S., and Shih, A.J., **Analysis of Machining and Machine Tools**, Springer, 2015.
Lopez, L.N., and Lamikiz, A. (eds.), **Machine Tools for High Performance Machining**, Springer, 2009.
Machinery's Handbook, various editions, Industrial Press.
Walsh, R.A., **McGraw-Hill Machining and Metalworking Handbook**, 3rd ed., McGraw-Hill, 2006.

Review Questions

23.1. Describe the types of machining operations that can be performed on a lathe.

23.2. What is turning? What kind of chips are produced by turning?

23.3. What is the thrust force in turning? What is the cutting force? Which is used to calculate the power required?

23.4. What are the components of a lathe?

23.5. What is a tracer lathe?

23.6. Describe the operations that can be performed on a drill press.

23.7. Why were power chucks developed?

23.8. Explain why operations such as boring on a lathe and tapping are difficult.

23.9. What is an automatic bar machine?

23.10. Why are turret lathes typically equipped with more than one turret?

23.11. Describe the differences between boring a workpiece on a lathe and boring it on a horizontal boring mill.

23.12. How is drill life determined?

23.13. What is the difference between a conventional drill and a gun drill?

23.14. Why are reaming operations performed?

23.15. Explain the functions of the saddle on a lathe.

23.16. Describe the relative advantages of (a) self-opening and (b) solid-die heads for threading.

23.17. Explain how external threads are cut on a lathe.

23.18. What is the difference between a blind hole and a through hole? What is the significance of that difference?

Qualitative Problems

23.19. Explain the reasoning behind the various design guidelines for turning.

23.20. Note that both the terms “tool strength” and “tool-material strength” have been used in the text. Do you think there is a difference between them? Explain.

23.21. List and explain the factors that contribute to poor surface finish in the processes described in this chapter.

23.22. List the advantages and disadvantages of turning or cold extruding a shaft.

23.23. Explain why the sequence of drilling, boring, and reaming produces a hole that is more accurate than drilling and reaming it only.

23.24. Why would machining operations be necessary even on net-shape or near-net-shape parts made by precision casting, forming, or powder-metallurgy products, as described in preceding chapters? Explain.

23.25. A highly oxidized and uneven round bar is being turned on a lathe. Would you recommend a small or a large depth of cut? Explain.

23.26. Describe the difficulties that may be encountered in clamping a workpiece made of a soft metal in a three-jaw chuck.

23.27. Does the force or torque in drilling change as the hole depth increases? Explain.

23.28. Drills usually have two flutes. Explain why.

23.29. Explain the similarities and differences in the design guidelines for turning and for boring.

23.30. Describe the advantages and applications of having a hollow spindle in the headstock of a lathe.

23.31. Assume that you are asked to perform a boring operation on a large-diameter hollow workpiece. Would you use a horizontal or a vertical boring mill? Explain.

23.32. Explain the reasons for the major trend that has been observed in producing threads by thread rolling as opposed to thread cutting. What would be the differences, if any, in the types of threads produced and in their performance characteristics?

23.33. Describe your observations concerning the contents of Tables 23.2 and 23.4, and explain why those particular recommendations are made.

23.34. The footnote to Table 23.12 states that as the hole diameter increases, speeds and feeds in drilling should be reduced. Explain why.

23.35. In modern manufacturing, which types of metal chips would be undesirable and why?

23.36. Sketch the tooling marks you would expect if a part was (a) turned; (b) reduced in diameter with a straight form tool; (c) extruded.

23.37. What concerns would you have in turning a powder metal part, such as a shaft made from the Osprey process (see Fig. 17.20)?

23.38. The operational severity for reaming is much lower than that for tapping, even though they both are internal machining processes. Why?

23.39. Review Fig. 23.6, and comment on the factors involved in determining the height of the zones (cutting speed) for various tool materials.

23.40. Explain how gun drills remain centered during drilling. Why is there a hollow, longitudinal channel in a gun drill?

23.41. Comment on the magnitude of the wedge angle on the tool shown in Fig. 23.4.

23.42. If inserts are used in a drill bit (see Fig. 23.22), how important is the shank material? If so, what properties are important? Explain.

23.43. Refer to Fig. 23.11b, and in addition to the tools shown, describe other types of cutting tools that can be placed in toolholders to perform other machining operations.

Quantitative Problems

23.44. Calculate the same quantities as in Example 23.1 for high-strength titanium alloy and at $N = 700$ rpm.

23.45. Estimate the machining time required to rough turn a 0.75-m-long annealed copper-alloy round bar from a 75-mm diameter to a 73-mm diameter, using a high-speed steel tool (see Table 23.4). Estimate the time required for an uncoated carbide tool.

23.46. A high-strength cast-iron bar 200 mm in diameter is being turned on a lathe at a depth of cut of $d = 1.25$ mm. The lathe is equipped with a 12 kW electric motor and has a mechanical efficiency of 80%. The spindle speed is 500 rpm. Estimate the maximum feed that can be used before the lathe begins to stall.

23.47. A 7.5-mm-diameter drill is used on a drill press operating at 300 rpm. If the feed is 0.125 mm/rev, what is the MRR? What is the MRR if the drill diameter is doubled?

23.48. In Example 23.4, assume that the workpiece material is high-strength aluminum alloy and the spindle is running at $N = 750$ rpm. Estimate the torque required for this operation.

23.49. For the data in Problem 23.48, calculate the power required.

23.50. A 150-mm-diameter aluminum cylinder 250 mm in length is to have its diameter reduced to 115 mm. Using the typical machining conditions given in Table 23.4, estimate the machining time if a TiN-coated carbide tool is used.

23.51. A lathe is set up to machine a taper on a bar stock 150-mm in diameter; the taper is 1 mm per 10 mm. A cut is made with an initial depth of cut of 4 mm at a feed rate of 0.300 mm/rev and at a spindle speed of 200 rpm. Calculated the average metal removal rate.

23.52. Assuming that the coefficient of friction is 0.25, calculate the maximum depth of cut for turning a hard aluminum alloy on a 15 kW lathe (with a mechanical efficiency of 80%) at a width of cut of 6 mm, rake angle of 0°, and a cutting speed of 90 m/min. What is your estimate of the material's shear strength?

23.53. A 75-mm-diameter gray cast iron cylindrical part is to be turned on a lathe at 500 rpm. The depth of cut is 6 mm and the feed is 0.5 mm/rev. What minimum horsepower is required for this operation?

23.54. Assume that you are an instructor covering the topics described in this chapter and you are giving a quiz on the numerical aspects to test the understanding of the students. Prepare two quantitative problems and supply the answers.

Synthesis, Design, and Projects

23.55. Drill life could be greatly increased if an effective means of cooling and lubrication were developed. Design methods of delivering a cutting fluid to the cutting zone, and discuss the advantages and shortcomings of your design.

23.56. Would you consider the machining processes described in this chapter as net-shape processes, thus requiring no further processing? Near-net-shape processing? Explain with appropriate examples.

23.57. Would it be difficult to use the machining processes described in this chapter on various soft non-metallic or rubberlike materials? Explain your thoughts, commenting on the role of the physical and mechanical properties of such materials with respect to the machining operation and any difficulties that may be encountered in producing the desired shapes and dimensional accuracies.

23.58. If a bolt breaks in a hole, it typically is removed by first drilling a hole in the bolt shank and then using a special tool to remove the bolt. Inspect such a tool and explain how it functions.

23.59. An important trend in machining operations is the increased use of flexible fixtures. Conduct a search on the Internet regarding these fixtures, and comment on their design and operation.

23.60. Review Fig. 23.8d, and explain if it would be possible to machine eccentric shafts, such as that shown in Fig. 23.13c, on the setup illustrated. What if the part is long compared with its cross section? Explain.

23.61. Boring bars can be designed with internal damping capabilities to reduce or eliminate vibration and chatter during machining (see Fig. 23.18). Referring to the technical literature, describe details of designs for such boring bars.

23.62. A large bolt is to be produced from extruded hexagonal bar stock by placing the hex stock into a chuck and machining the shank of the bolt by turning it on a lathe. List and explain the difficulties that may be involved in this operation.

23.63. Make a comprehensive table of the process capabilities of the machining operations described in this chapter. Using several columns, describe the machine tools involved, type of cutting tools and tool materials used, shapes of parts produced, typical maximum and minimum sizes, surface finish, dimensional tolerances, and production rates.

Chapter 24

Machining Processes: Milling, Broaching, Sawing, Filing, and Gear Manufacturing

- This chapter begins with milling, one of the most versatile and important machining processes, in which a rotating cutter removes material while traveling along a specified path.
- Described next are the processes of planing, shaping, and broaching, in which either the cutting tool or the workpiece travels along a straight path, producing flat or profiled machined surfaces.
- Next described are sawing and filing, including tool design and machinery involved in these processes.
- The chapter ends with descriptions of gear-manufacturing by machining, the special cutters used, the automated equipment involved, the quality and properties of the gears produced.

Typical parts made: Parts with complex external and internal features, splines, and gears.

Alternative processes: Die casting, precision casting, precision forging, additive manufacturing, powder metallurgy, powder-injection molding, creep-feed grinding, electrical discharge machining, and fabrication.

24.1 Introduction

In addition to producing parts with various external or internal round profiles, as described in Chapter 23, machining operations also can produce many other complex shapes (Fig. 24.1). Although processes such as die casting, precision forging, and powder metallurgy can produce parts with close tolerances and fine surface finish, it is often necessary to perform complex machining operations to respond to various design requirements and specifications. In this chapter, several important machining processes and machine tools capable of producing complex shapes, using single-point, multitooth, and profiled cutting tools, are described (see also Table 23.1).

24.2 Milling and Milling Machines

Milling includes a number of highly versatile machining operations taking place in a variety of configurations (Fig. 24.2), with the use of a **milling cutter**, a *multitooth* tool that produces a number of chips in one revolution (Fig. 24.3).

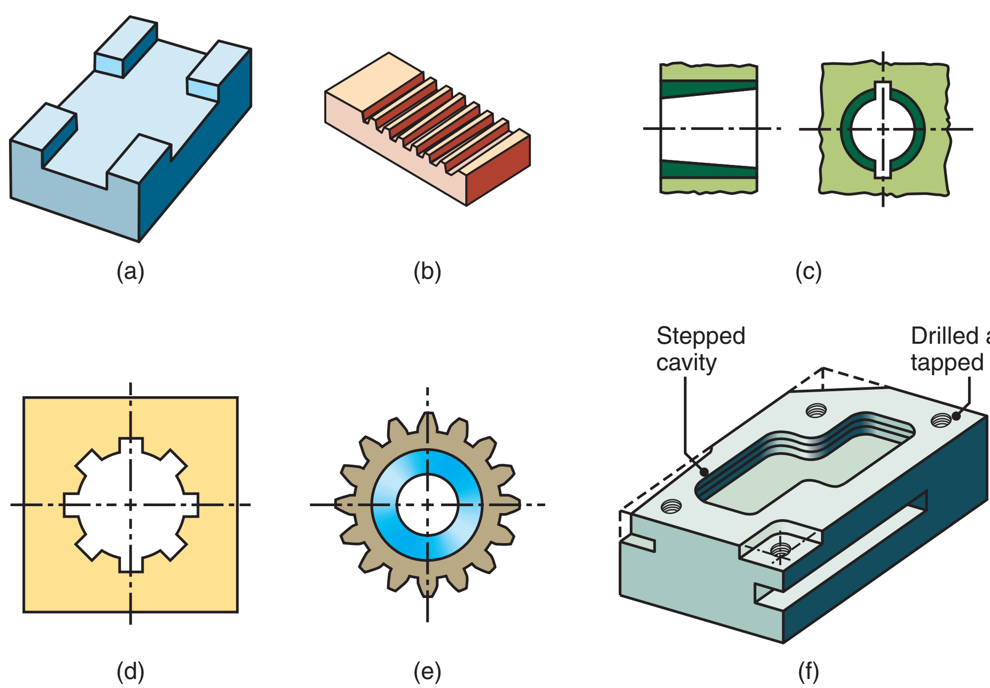


Figure 24.1: Typical parts and shapes that can be produced with the machining processes described in this chapter.

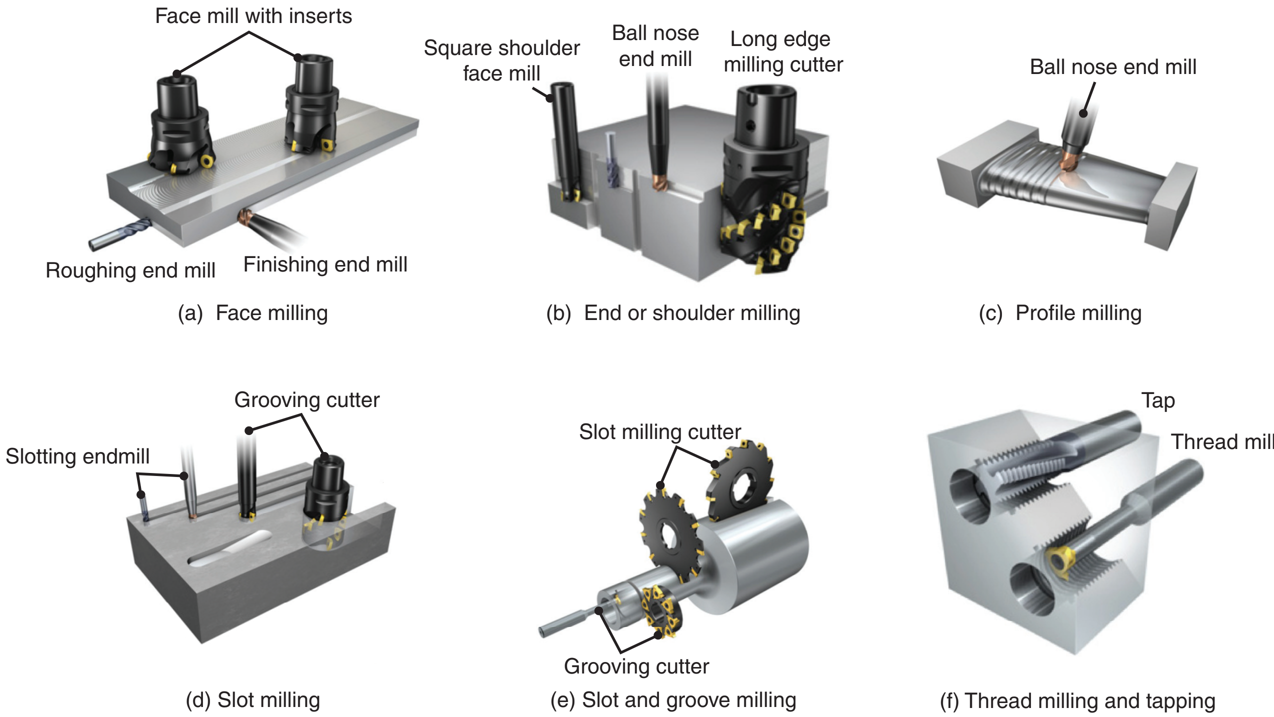


Figure 24.2: Some basic types of milling cutters and milling operations. (a) Face milling; (b) end or shoulder milling; (c) profile milling; (d) slot milling; (e) slot and groove milling; (f) thread milling and tapping. *Source:* Courtesy of Sandvik Coromant.

Figure 24.3: Photograph of the cutting action of a milling cutter that uses a number of inserts to remove metal in the form of chips. *Source:* Courtesy of Sandvik Coromant.

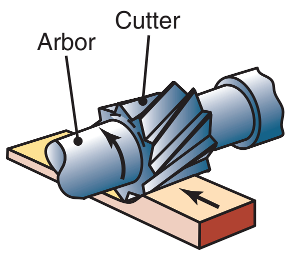


Figure 24.4: Schematic illustration of peripheral milling.

24.2.1 Peripheral Milling

In *peripheral milling*, also called *plain milling*, the axis of cutter rotation is parallel to the workpiece surface, as shown in Fig. 24.4. The cutter body, which generally is made of high-speed steel (Section 22.2), has a number of teeth along its circumference; each tooth acts like a single-point cutting tool. When the cutter is longer than the width of the cut, the operation is called **slab milling**.

Cutters for peripheral milling may have either *straight* or *helical teeth*, resulting in an orthogonal or oblique cutting action, respectively. Helical teeth generally are preferred over straight teeth, because each tooth is always partially engaged with the workpiece as the cutter rotates. Consequently, the cutting force and the torque on the cutter are lower, resulting in a smoother milling operation and reduced chatter.

Conventional Milling and Climb Milling. Note in Fig. 24.5a that the cutter rotation can be either clockwise or counter-clockwise; this is significant in the milling operation. In *conventional milling*, also called *up milling*, the maximum chip thickness is at the *end* of the cut as the tooth leaves the workpiece surface. Thus, contaminants and scale (oxide layer) on the surface do not adversely affect tool life. This is the more common method of milling, where the cutting operation is smooth. However, the cutter teeth must be sharp, as otherwise the tooth will rub against the surface being milled and smear it for some distance before it begins to engage and cut. There may also be a tendency for the cutter to chatter (Section 25.4) and for the workpiece to be lifted *upward*, because of the cutter rotation direction. Proper clamping of the workpiece on the table of the machine is thus important.

In *climb milling*, also called *down milling*, cutting starts at the *surface* of the workpiece where the chip is thickest. The advantage of this method is that the direction of rotation of the cutter will push the workpiece

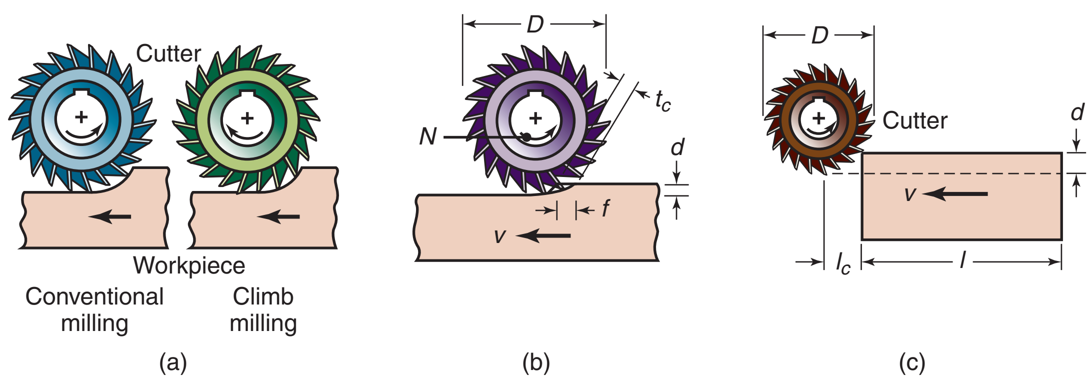


Figure 24.5: (a) Schematic illustration of conventional milling and climb milling. (b) Slab-milling operation showing depth of cut, d; feed per tooth, f; chip depth of cut, t_c, and workpiece speed, v. (c) Schematic illustration of cutter travel distance, l_c, to reach full depth of cut.

downward, thus holding the workpiece in place, a factor particularly important for slender parts. Because of the resulting impact force when a tooth first engages the workpiece, however, this operation must have a rigid workholding setup, and gear backlash in the table feed mechanism must be eliminated. Climb milling is not suitable for machining workpieces having surface scale, such as metals that have been hot worked, forged, or cast. Scale is hard and abrasive, and thus causes excessive wear and damage to the cutter teeth, shortening their life.

Milling Parameters. The cutting speed, V, in peripheral milling is the surface speed of the cutter, or

$$V = \pi D N, \tag{24.1}$$

where D is the cutter diameter and N is the rotational speed of the cutter (Fig. 24.6).

Note from Fig. 24.5b that the thickness of the chip in slab milling will vary along its length because of the relative longitudinal motion between the cutter and the workpiece. For a straight-tooth cutter, the approximate *undeformed chip thickness* (also called *chip depth of cut*), t_c, can be calculated from the equation

$$t_c = 2f\sqrt{\frac{d}{D}}, \tag{24.2}$$

where f is the feed per tooth of the cutter (the distance the workpiece travels per tooth of the cutter, in mm/tooth), and d is the depth of cut. As t_c becomes larger, the force on the cutter tooth will increase.

Feed per tooth is determined from the equation

$$f = \frac{v}{Nn}, \tag{24.3}$$

where v is the linear speed (also called *feed rate*) of the workpiece and n is the number of teeth on the cutter periphery.

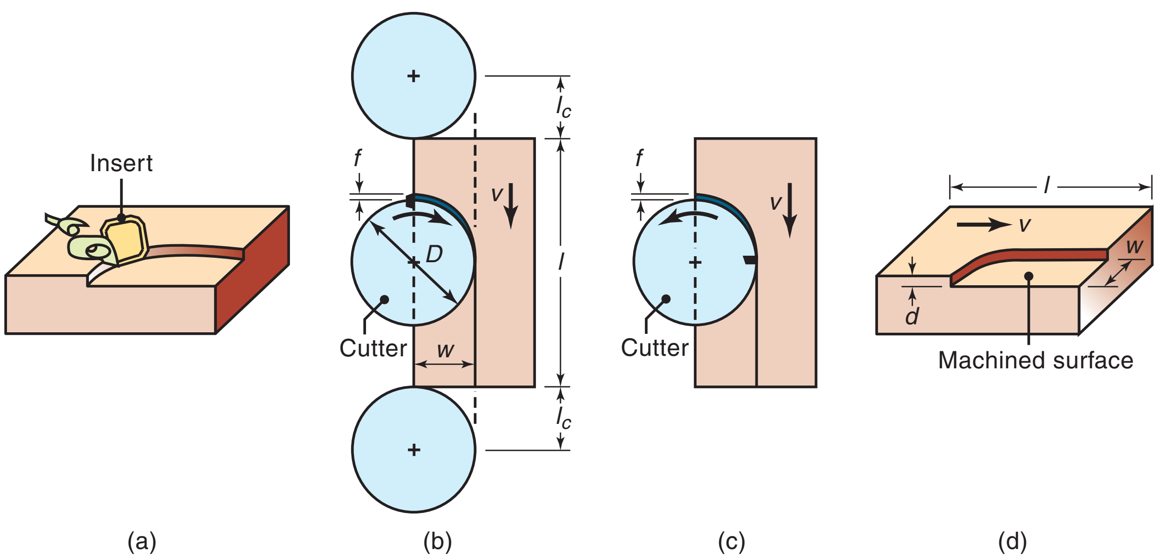


Figure 24.6: (a) Face-milling operation with cutter removed, showing the action of a single insert; (b) climb milling; (c) conventional milling; and (d) dimensions in face milling. The width of cut, w, is not necessarily the same as the cutter radius.

Table 24.1: Summary of Peripheral Milling Parameters and Formulas.

N = Rotational speed of the milling cutter, rpm
F = Feed, mm/tooth
D = Cutter diameter, mm
n = Number of teeth on cutter
v = Linear speed of the workpiece or feed rate, mm/min
V = Surface speed of cutter, m/min
$= DN$
f = Feed per tooth, mm/tooth
$= v/Nn$
l = Length of cut, mm
t = Cutting time, s or min
$= (l + l_c)/v$, where l_c = extent of the cutter's first contact with the workpiece
MRR = mm^3/min
$= wdv$, where w is the width of cut
Torque = N-m
$= F_c D/2$
Power = kW
= (Torque)(ω), where $\omega = 2\pi N$ radians/min

The cutting time, t, is given by the equation

$$t = \frac{l + l_c}{v}, \tag{24.4}$$

where l is the length of the workpiece (Fig. 24.5c) and l_c is the horizontal extent of the cutter's first contact with the workpiece. Based on the assumption that $l_c \ll l$ (although this generally is not the case), the *material-removal rate* (MRR) is

$$\text{MRR} = \frac{lwd}{t} = wdv, \tag{24.5}$$

where w is the width of the cut, which, in slab milling, is equal to the width of the workpiece. As stated in Section 23.2, the distance that the cutter travels in the *noncutting cycle* of the milling operation is an important economic consideration, and should be minimized by such means as faster travel of the machine tool components. The foregoing equations and the terminology used are summarized in Table 24.1.

Although the *power requirement* in peripheral milling can be measured or calculated, the tangential, radial, and axial *forces* on the cutter (see also Fig. 23.5) are difficult to calculate. There are numerous variables involved, many of which pertain to the cutter geometry; these forces can be measured experimentally, while the *torque* on the cutter spindle (the product of the cutter radius and the tangential force) can be calculated from the power (see Example 24.1). The tangential force per tooth will depend on how many teeth are engaged at any moment during the cut.

Example 24.1 Material-removal Rate, Power, Torque, and Cutting Time in Slab Milling

Given: A slab-milling operation is being carried out on a 300 mm-long, 100-mm-wide annealed mild-steel block at a feed $f = 0.25$ mm/tooth and a depth of cut $d = 3$ mm. The cutter is $D = 50$ mm in diameter, has 20 straight teeth, rotates at $N = 100$ rpm, and, by definition, is wider than the block to be machined.

Find: Calculate the material-removal rate, estimate the power and torque required for this operation, and calculate the cutting time.

Solution: From the information given, the linear speed of the workpiece, v, can be calculated from Eq. (24.3):

$$v = fNn = (0.25)(100)(20) = 500 \text{ mm/min.} = 0.00833 \text{ m/s.}$$

From Eq. (24.5), the material-removal rate is calculated to be

$$\text{MRR} = (100)\,(3)\,(500) = 150,000 \text{ mm}^3\text{/min.} = 2500 \text{ mm}^3\text{/s.}$$

Since the workpiece is annealed mild steel, the unit power is estimated from Table 21.2 as 5.5 W-s/mm^3. Therefore, the power required can be estimated as

$$\text{Power} = (5.5)(2500) = 13.75 \text{ kW.}$$

Also,

$$\text{Torque} = \frac{\text{Power}}{\text{Rotational Speed}} = \frac{(13.75 \text{ kW})}{(100 \text{ rpm})(2\pi)} = 21.9 \text{ Nm.}$$

The cutting time is given by Eq. (24.4), in which the quantity l_c can be shown, from simple geometric relationships and for $D \gg d$, to be approximately equal to

$$l_c = \sqrt{Dd} = \sqrt{(50)\,(3)} = 12.24 \text{ mm.}$$

Thus, the cutting time is

$$t = \frac{300 + 12.25}{500} = 0.625 \text{ min} = 37.5 \text{ s.}$$

24.2.2 Face Milling

In *face milling*, the cutter is mounted on a spindle having an axis of rotation perpendicular to the workpiece surface (Fig. 24.2a); it removes material in the manner shown in Fig. 24.6a. The cutter rotates at a rotational speed, N, and the workpiece moves along a straight path, at a linear speed, v. When the direction of cutter rotation is as shown in Fig. 24.6b, the operation is *climb milling*; when it is in the opposite direction (Fig. 24.6c), it is *conventional milling*. The cutting teeth, such as carbide inserts, are mounted on the cutter body, as shown in Fig. 24.7 (see also Fig. 22.3c).

Because of the relative motion between the cutter tooth and the workpiece, face milling leaves *feed marks* on the machined surface (Fig. 24.8), similar to those left by turning operations as shown in Fig. 21.2. Note that the surface roughness of the workpiece depends on the corner geometry of the insert and the feed per tooth.

The *terminology* for a face-milling cutter, as well as for various angles, is shown in Fig. 24.9. As can be seen from the side view of the insert in Fig. 24.10, the *lead angle* of the insert in face milling has a direct influence on the *undeformed chip thickness*, as it does in turning operations (see Fig. 23.3). As the lead angle (positive, as shown in Fig. 24.10b) increases, the undeformed chip thickness decreases, and the length of contact, and hence chip width, increases. Note, however, that the cross-sectional area of the undeformed chip remains constant.

The lead angle also influences the forces in milling. It can be seen that as the lead angle decreases, there is a smaller vertical-force component (that is, the axial force on the cutter spindle). The *lead angles* for most face-milling cutters typically range from 0° to 45°.

A wide variety of milling cutters and inserts are available (Figs. 22.2 and 24.7). The cutter diameter should be chosen so that it will not interfere with fixtures, workholding devices, or other components in

Figure 24.7: A face-milling cutter with indexable inserts. *Source:* Courtesy of Ingersoll Cutting Tool Company.

the setup. In a typical face-milling operation, the ratio of the cutter diameter, D, to the width of cut, w, should be no less than 3:2.

The relationship of cutter diameter to insert angles, and their position relative to the surface to be milled are important, in that they will determine the angle at which an insert *enters* and *exits* the workpiece. Note in Fig. 24.6b for climb milling that, if the insert has zero axial and radial rake angles (see Fig. 24.9), the rake face of the insert engages the workpiece directly. As seen in Fig. 24.11a and b, however, the same insert may engage the workpiece at different angles, depending on the relative positions of the cutter and the workpiece width.

Note in Fig. 24.11a that since the tip of the insert makes the first contact, there is a possibility for the cutting edge to chip off. In Fig. 24.11b, on the other hand, the first contacts (at entry, reentry, and the two exits) are at an angle and away from the tip of the insert. Consequently, there is a lower tendency for the insert to fail, because the forces on the insert vary more slowly. Note from Fig. 24.9 that the radial and axial rake angles also will have an effect on this operation.

Figure 24.11c shows the exit angles for various cutter positions. Note in the first two examples that the insert exits the workpiece at an angle, thus causing the force on the insert to be reduced to zero at a slower rate (desirable for longer tool life) than in the third example, where the insert exits the workpiece abruptly.

Example 24.2 Material-removal Rate, Power Required, and Cutting Time in Face Milling

Given: Refer to Fig. 24.6 and assume that $D = 150$ mm, $w = 60$ mm, $l = 500$ mm, $d = 3$ mm, $v = 0.6$ m/min, and $N = 100$ rpm. The cutter has 10 inserts, and the workpiece material is a high-strength aluminum alloy.

Find: Calculate the material-removal rate, cutting time, and feed per tooth, and estimate the power required.

Solution: First note that the cross section of the cut is $wd = (60)(3) = 180$ mm^2. Then, noting that the workpiece speed, v, is 0.6 m/min $= 600$ mm/min, the material-removal rate (MRR) can be calculated as

$$\text{MRR} = (180)(600) = 108,000 \text{ mm}^3/\text{min}.$$

The cutting time is given by

$$t = \frac{l + 2l_c}{v}.$$

Note from Fig. 24.6 that, for this problem, $l_c = \frac{D}{2} = 75$ mm. The cutting time is therefore

$$t = \frac{500 + 150}{10} = 65 \text{ s} = 1.08 \text{ min}.$$

The feed per tooth can be obtained from Eq. (24.3), where $N = 100$ rpm $= 1.67$ rev/s, and hence

$$f = \frac{10}{(1.67)(10)} = 0.6 \text{ mm/tooth}.$$

For this material, the unit power can be estimated from Table 21.2 to be 1.1 Ws/mm^3. Thus, the power is

$$\text{Power} = (1.1)(1800) = 1980 \text{ W} = 1.98 \text{ kW}.$$

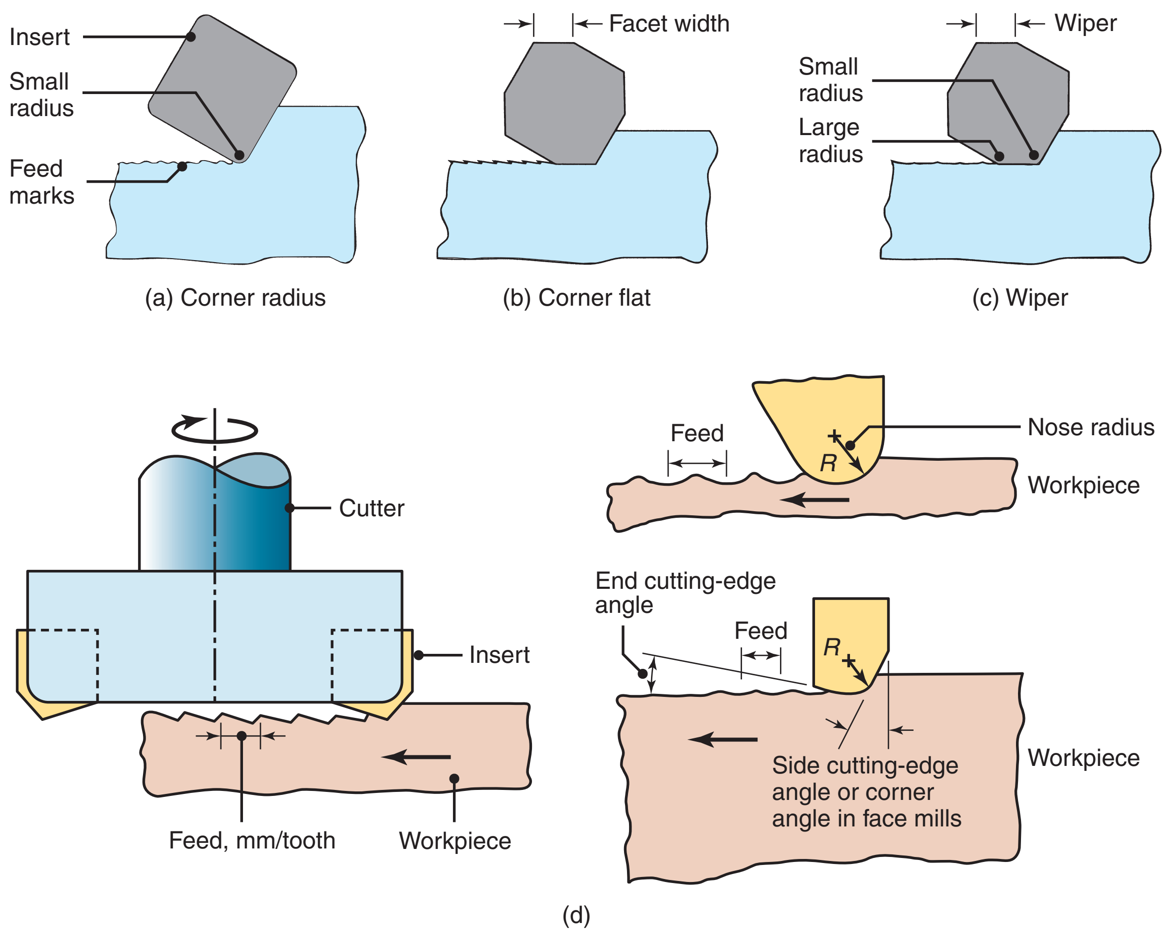


Figure 24.8: Schematic illustration of the effect of insert shape on feed marks on a face-milled surface: (a) small corner radius, (b) corner flat on insert, and (c) wiper, consisting of a small radius followed by a large radius, resulting in smoother feed marks. (d) Feed marks due to various insert shapes.

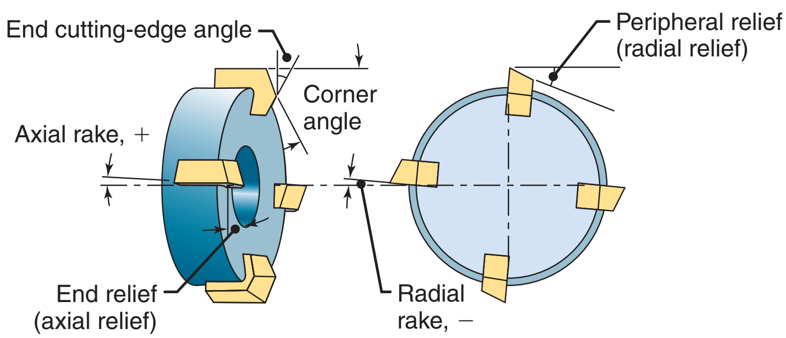


Figure 24.9: Terminology for a face-milling cutter.

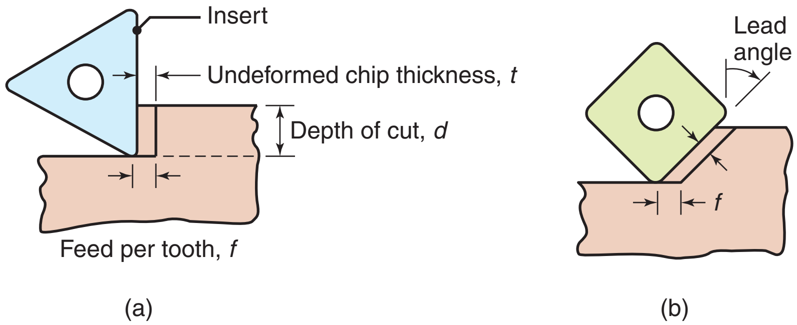


Figure 24.10: The effect of the lead angle on the undeformed chip thickness, t in face milling. Note that as the lead angle increases, the chip thickness decreases, but the length of contact (i.e., chip width) increases. The edges of the insert must be sufficiently large to accommodate the contact length increase.

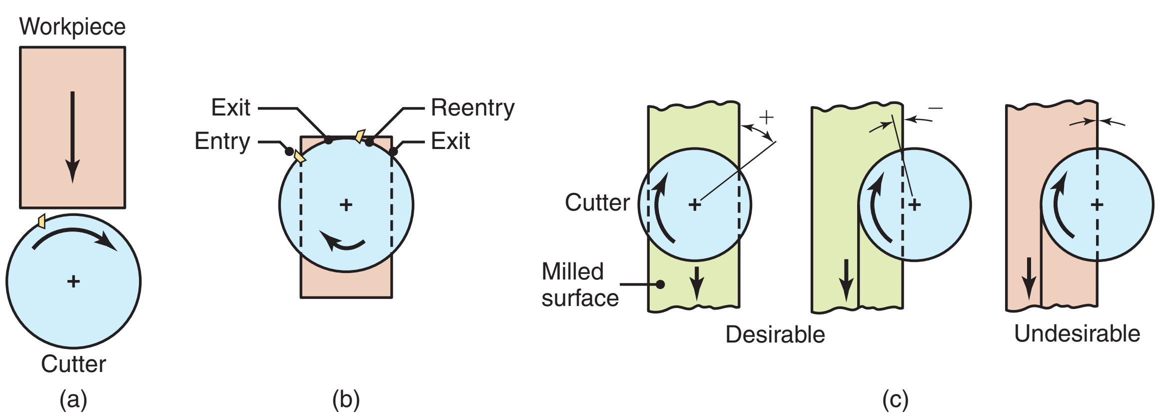


Figure 24.11: (a) Relative position of the cutter and insert as they first engage the workpiece in face milling. (b) Insert positions towards the end of cut. (c) Examples of exit angles of the insert, showing desirable (positive or negative angle) and undesirable (zero angle) positions. In all figures, the cutter spindle is perpendicular to the page and rotates clockwise.

24.2.3 End Milling

End milling is an important and common machining operation, because of its versatility and capability to produce various profiles and curved surfaces. The cutter, called an **end mill** (Fig. 24.12), has either a straight shank (for small cutter sizes) or a tapered shank (for larger sizes), and is mounted into the spindle of the milling machine. End mills may be made of high-speed steels, solid carbide, or with coated or uncoated carbide inserts, similar to those for face milling. The cutter usually rotates on an axis perpendicular to the workpiece surface, but it can be tilted to conform to machine-tapered or curved surfaces.

End mills are available with hemispherical ends (*ball nose mills*) for machining sculptured surfaces, such as in dies and molds. They can also be produced with a specific radius, profile, flat end, or with chamfer. *Hollow end mills* have internal cutting teeth, and are used to machine the cylindrical surfaces of solid, round workpieces. End milling can produce a variety of surfaces at any depth, such as curved, stepped, and pocketed (Fig. 24.2b). The cutter can remove material on both its end and on its cylindrical cutting edges, as can be seen in Fig. 24.2b.

Vertical-spindle and horizontal-spindle milling machines (see Section 24.2.8), as well as machining centers (see Fig. 25.7), can all be used for end milling. The machines can be programmed such that the cutter can follow a complex set of paths so as to optimize the whole machining operation, for higher productivity and minimum cost.

High-speed End Milling. *High-speed end milling* is an important process, with numerous applications such as the milling of large aluminum-alloy aerospace components and honeycomb structures (see also *high-speed machining*, Section 25.5). With spindle speeds up to 80,000 rpm, the machines must have high stiffness, usually requiring hydrostatic or air bearings and high-quality work holding devices. The spindles have a rotational accuracy of 10 μm; thus the surfaces produced have very high dimensional accuracy. At high rates of material removal, chip collection and disposal can be a significant problem, as described in Section 23.3.7.

Machining cavities in dies (called **die sinking**, such as in forging or in sheet-metal forming) also is done by high-speed end milling, often using TiAlN-coated ball-nose end mills (Fig. 24.13). The machines generally have *four-axis* or *five-axis* movement capability (see, for example, Fig. 24.21), but machining centers (Section 25.2) can add more axes for more complex geometries. Such machines can accommodate dies as large as 3 m $\times$ 6 m and weighing 54 metric tons, costing over $2 million. The advantages of five-axis machines are that they (a) are capable of machining very complex shapes, in a single setup, (b) can use shorter tools, thus reducing the tendency for vibration and chatter, and (c) enable drilling of holes at various compound angles.

Figure 24.12: A selection of end mills. The flute depth and helix angle are selected based on whether it is a roughing or finishing cut. Note the variety of geometries of the end of the mill; with the proper cutter, a radius, chamfer, or flat surface can be machined. *Source:* Courtesy of Kennametal, Inc.

Figure 24.13: Ball nose end mills. These cutters are can produce complex contours and are often used in machining dies and molds (see also Fig. 24.2d). *Source:* Courtesy of Dijet, Inc.

24.2.4 Other Milling Operations and Milling Cutters

Several other milling operations and cutters are used to machine workpieces. In **straddle milling**, two or more cutters are mounted on an arbor, and used to simultaneously machine two parallel surfaces on workpieces (Fig. 24.14a). **Form milling** produces curved profiles, using cutters with specially shaped teeth (Fig. 24.14b); such cutters are also used for cutting gear teeth (Section 24.7). **Slotting** and **slitting** operations

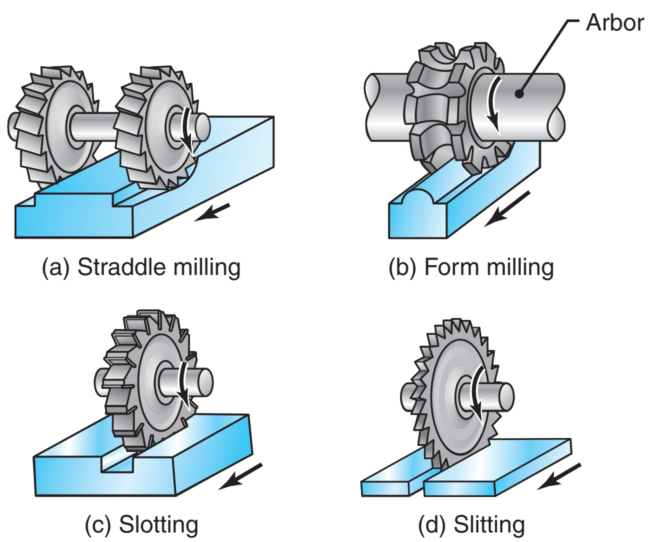


Figure 24.14: Cutters for (a) straddle milling, (b) form milling, (c) slotting, and (d) slitting with a milling cutter.

are performed with *circular cutters*, as shown in Fig. 24.14c and d, respectively. The teeth may be staggered slightly, like those in a saw blade (Section 24.5), to provide clearance for the cutter width when machining deep slots. *Slitting saws* are relatively thin, usually less than 5 mm. *T-slot cutters* are used to mill T-slots, such as those in machine-tool worktables for clamping workpieces. As shown in Fig. 24.15a, a slot is first milled with an end mill, and then the cutter machines the complete profile of the T-slot, all in one pass.

Key seat cutters are used to make semicylindrical (or *Woodruff*) key seats for shafts. *Angle milling cutters*, either single-angle or double-angle, are used to produce tapered surfaces with various angles. *Shell mills* (Fig. 24.15b) are hollow inside, and are mounted on a shank, thus allowing the same shank to be used for different-sized cutters. The uses of shell mills are similar to those for end mills.

Milling with a single cutting tooth, mounted on a spindle, is known as *fly cutting*; it is generally used in simple face-milling and boring operations. The tool can be shaped as a single-point cutting tool, and can be placed in various radial positions on the spindle, in an arrangement similar to that shown in Fig. 23.24b.

24.2.5 Toolholders

The stiffness of toolholders and cutters is important for surface quality and in reducing vibration and chatter during milling operations. **Arbor cutters** are mounted on an *arbor* (see Figs. 24.14 and 24.18a), for operations such as peripheral, face, straddle, and form milling. In **shank-type cutters**, the cutter and the shank are made in one piece, the most common examples being end mills. Small end mills have straight shanks, but larger ones have tapered shanks, for better mounting in the machine spindle in order to resist the high forces and torque involved during cutting. Cutters with straight shanks are mounted in collet chucks or in special end-mill holders; those with tapered shanks are mounted in tapered toolholders.

24.2.6 Milling Process Capabilities

In addition to the various characteristics of the milling processes described thus far, milling process capabilities include such parameters as surface finish, dimensional tolerances, production rate, and cost considerations. Data on process capabilities are presented in Tables 23.1 and 23.9, Figs. 23.13 and 23.15, and in Chapter 40.

The conventional ranges of cutting speeds and feeds for milling are given as guidelines in Table 24.2. Depending on the workpiece material, cutting-tool material and process parameters, cutting speeds are in the range of 30 to 3000 m/min. Feed per tooth typically ranges from about 0.1 mm to 0.5 mm, and depths of cut are usually 1 to 8 mm. For cutting-fluid recommendations, see Table 23.6.

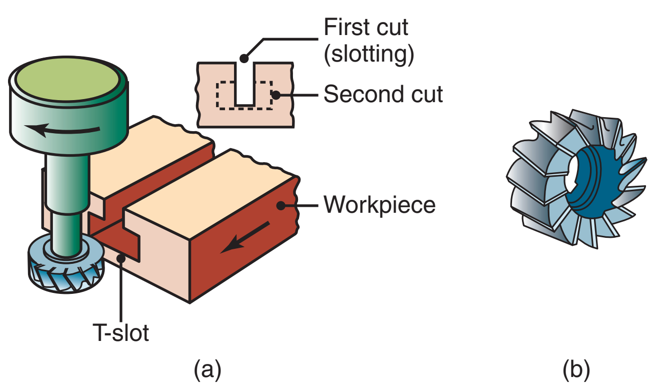


Figure 24.15: (a) T-slot cutting with a milling cutter. (b) A shell mill.

Table 24.2: General Recommendations for Milling Operations. Note that these values are for a particular machining geometry and are often exceeded in practice.

		General-purpose starting conditions		Range of conditions	
Material	**Cutting tool**	**Feed mm/tooth**	**Speed m/min**	**Feed mm/tooth**	**Speed m/min**
Low-carbon and free-machining steels	Uncoated carbide, coated carbide, cermets	0.13-0.20	100-472	0.085-0.38	90-425
Alloy steels					
Soft	Uncoated, coated cermets	0.10–0.18	100–260	0.08–0.30	60–370
Hard	Cermets, PcBN	0.10–0.15	90–220	0.08–0.25	75–460
Cast iron, gray					
Soft	Uncoated, coated, cermets, SiN	0.10–0.20	160–440	0.08–0.38	90–1370
Hard	Cermets, SiN, PcBN	0.10–0.20	120–300	0.08–0.38	90–460
Stainless steel, Austenitic	Uncoated, coated, cermets	0.13–0.18	120–370	0.08–0.38	90–500
High-temperature alloys Nickel based	Uncoated, coated, cermets, SiN, PcBN	0.10–0.18	30–370	0.08–0.38	30–550
Titanium alloys	Uncoated, coated, cermets	0.13–0.15	50–60	0.08–0.38	40–140
Aluminum alloys					
Free machining	Uncoated, coated, PCD	0.13–0.23	1200–1460	0.08–0.46	300–3000
High silicon	PCD	0.13	610	0.08–0.38	370–910
Copper alloys	Uncoated, coated, PCD PCD	0.13–0.23	300–760	0.08–0.46	90–1070
Plastics	Uncoated, coated, PCD PCD	0.13–0.23	270–460	0.08–0.46	90–1370

Source: Based on data from Kennametal, Inc.

Note: Depths of cut, *d*, usually are in the range of 1–8 mm. PcBN: polycrystalline cubic-boron nitride. PCD: polycrystalline diamond. See also Table 23.4 for range of cutting speeds within tool material groups.

A general **troubleshooting guide** for milling operations is given in Table 24.3; the last four items in this table are illustrated in Figs. 24.16 and 24.17. *Back striking* involves double feed marks, which are made by the trailing edge of the cutter. Note from Table 24.3 that some recommendations (such as changing milling parameters or cutting tools) are easier to accomplish than others (such as changing tool angles, cutter geometry, and the stiffness of spindles and work holding devices).

24.2.7 Design and Operating Guidelines for Milling

The guidelines for turning and boring, given in Sections 23.3.6 and 23.4, are also generally applicable to milling operations. Additional factors relevant to milling include the following:

- Standard milling cutters should be used as much as possible, depending on part design features; costly special cutters should be avoided.

Table 24.3: General Troubleshooting Guide for Milling Operations.

Problem	Probable causes
Tool breakage	Tool material lacks toughness, improper tool angles, machining parameters too high
Excessive tool wear	Machining parameters too high, improper tool material, improper tool angles, improper cutting fluid
Rough surface finish	Feed per tooth too high, too few teeth on cutter, tool chipped or worn, built-up edge, vibration and chatter
Tolerances too broad	Lack of spindle and work holding device stiffness, excessive temperature rise, dull tool, chips clogging cutter
Workpiece surface burnished	Dull tool, depth of cut too low, radial relief angle too small
Back striking	Dull cutting tools, tilt in cutter spindle, negative tool angles
Chatter marks	Insufficient stiffness of system; external vibrations; feed, depth of cut, and width of cut too large; select stable processing parameters
Burr formation	Dull cutting edges or too much honing, incorrect angle of entry or exit, feed and depth of cut too high, incorrect insert shape
Breakout	Lead angle too low, incorrect cutting-edge geometry, incorrect angle of entry or exit, feed and depth of cut too high

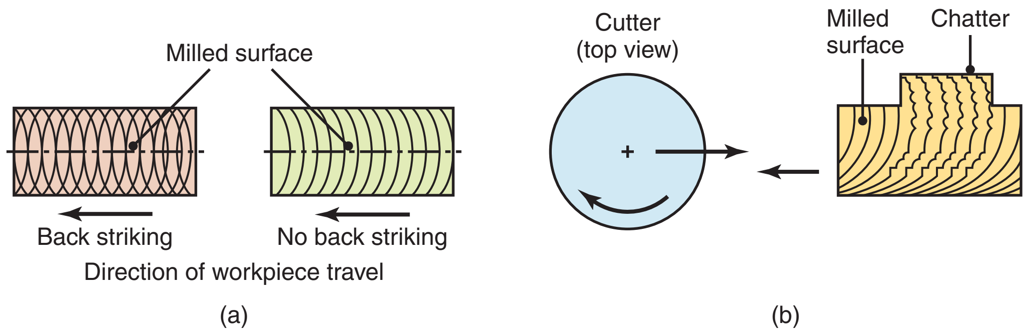


Figure 24.16: Machined surface features in face milling (see also Fig. 24.8).

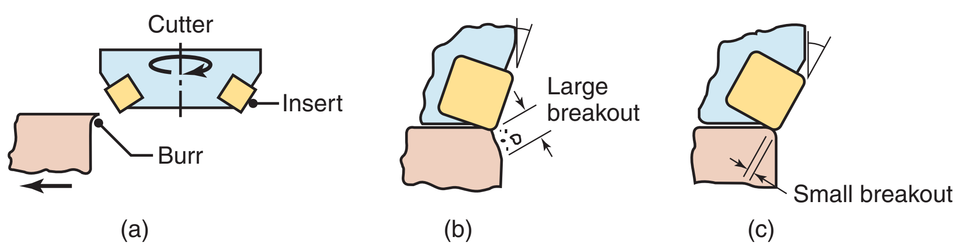


Figure 24.17: Edge defects in face milling: (a) burr formation along workpiece edge, (b) breakout along workpiece edge, and (c) how it can be avoided by increasing the lead angle (see also last row in Table 24.3).

- Internal cavities and pockets with sharp corners should be avoided because of the difficulty of milling them, since cutting teeth or inserts have a finite edge radius. When possible, the corner radius should match the milling cutter geometry.

- Bevels should be preferred over radii, because cutter and setup costs are higher for machining radii. If inner and outer mating surfaces have the same radius, then the transition between them has to be very accurately machined and is very difficult. Specifying a chamfer for the inner surface ensures proper assembly can take place.

- Although small milling cutters can be used for milling any surface, they are less rugged and more susceptible to chatter and to tool breakage than large cutters.
- Workpieces should be sufficiently rigid to minimize deflections that may result from clamping and cutting forces.
- Parts should be designed so that they can be clamped or held in fixtures during machining, and the fixturing should be designed to minimize the number of times that the part needs to be repositioned to complete the milling operation.

Guidelines for avoiding vibration and chatter in milling are similar to those for turning; in addition, the following practices should be considered:

- Cutters should be mounted as close to the spindle base as possible, in order to reduce tool deflections.
- Toolholders and fixturing devices should be as rigid as possible, in order to avoid or minimize vibrations and chatter.
- In cases where vibration and chatter may occur, tool shape and process conditions should be modified, including the use of cutters with fewer teeth or, whenever possible, with random tooth spacing (see Section 25.4).

24.2.8 Milling Machines

Because they are capable of performing a wide variety of cutting operations, milling machines are among the most versatile of all machine tools. The first milling machine was built in 1820 by E. Whitney (1765–1825); a wide selection of milling machines with numerous features is now available, the most common of which are described below. These machines have been and continue to be replaced with *computer numerical-control* (CNC) *machines* and *machining centers*. Modern machines are very versatile and capable of milling, drilling, boring, and tapping, with repeated and high accuracy (Fig. 24.20).

Column-and-knee-type Machines. Used for general-purpose milling operations, *column-and-knee-type machines* have been the most common milling machines. The spindle on which the cutter is mounted may be *horizontal* (Fig. 24.18a), for peripheral milling, or *vertical*, for face and end milling, boring, and drilling operations (Fig. 24.18b). The basic components of these machines are:

- *Worktable:* the workpiece is clamped on the worktable using T-slots; the table moves longitudinally relative to the saddle.
- *Saddle:* supports the table and can move in the transverse direction.
- *Knee:* supports the saddle and gives the table vertical movement so that the depth of cut can be adjusted and workpieces with various heights can be accommodated.
- *Overarm:* used on horizontal machines; it is adjustable to accommodate different arbor lengths.
- *Head:* contains the spindle and cutter holders. In vertical machines, the head may be fixed or it can be adjusted vertically; it can be swiveled in a vertical plane on the column for milling tapered surfaces.

Plain milling machines have at least three axes of movement, with the motion usually imparted manually, either with a power screw actuator or by engaging powered actuators to the drive motor. In **universal column-and-knee milling machines**, the table can be swiveled on a horizontal plane. Complex shapes,

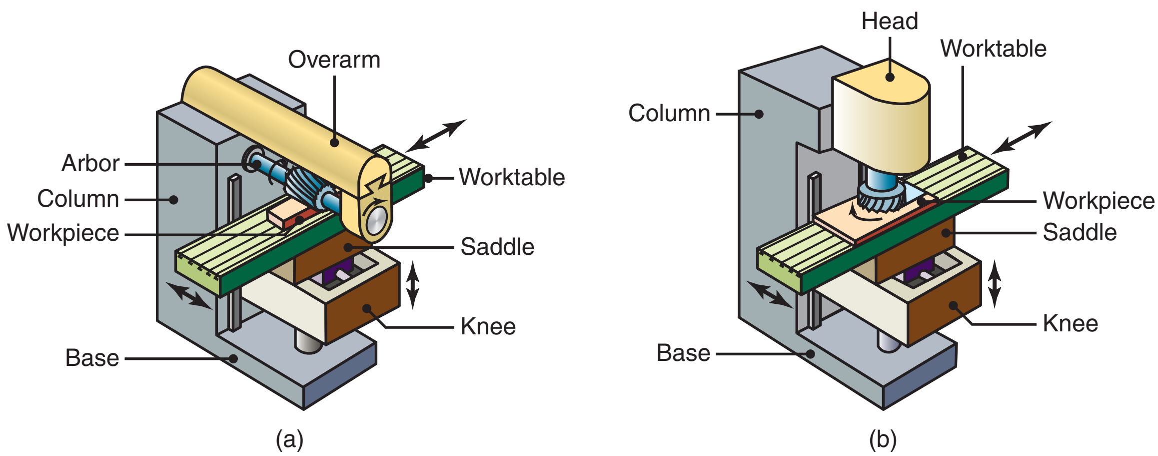


Figure 24.18: Schematic illustration of (a) a horizontal-spindle column-and-knee-type milling machine and (b) vertical-spindle column-and-knee-type milling machine. *Source:* After G. Boothroyd.

such as helical grooves at various angles, can be machined to produce such parts as gears, drills, taps, and cutters.

Bed-type Milling Machines. In *bed-type machines*, the worktable is mounted directly on the bed, which replaces the knee and moves only longitudinally (Fig. 24.19). Although not as versatile as other types, these machines have high stiffness and typically are used for high-production work. The spindles may be horizontal or vertical and of duplex or triplex types (with two or three spindles, respectively), for the simultaneous machining of two or three workpiece surfaces.

Other Types of Milling Machines. Several other types of milling machines are available (see also *machining centers*, Section 25.2). **Planer-type milling machines**, which are similar to bed-type machines, are equipped with several heads and cutters to mill different surfaces. They are typically used for heavy workpieces and are more efficient than simple planers (Section 24.3) when used for similar purposes. **Rotary-table machines** are similar to vertical machines and are equipped with one or more heads for face-milling operations. **Profile milling machines** have five axes of movement (Fig. 24.21); note the three linear and two angular movements of the machine components.

Workholding Devices and Accessories. The workpiece to be machined must be clamped securely to the worktable in order to resist cutting forces and prevent slipping during milling. Various fixtures and vises

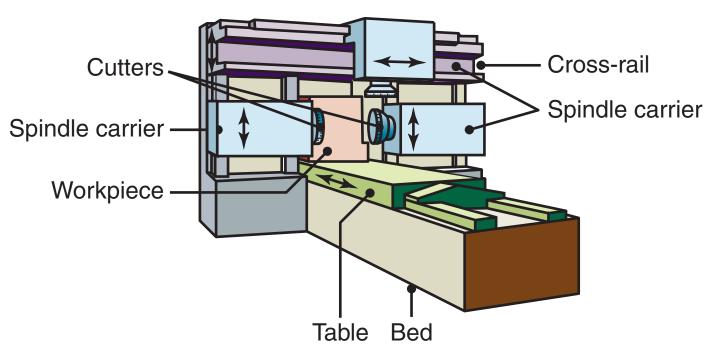


Figure 24.19: Schematic illustration of a bed-type milling machine.

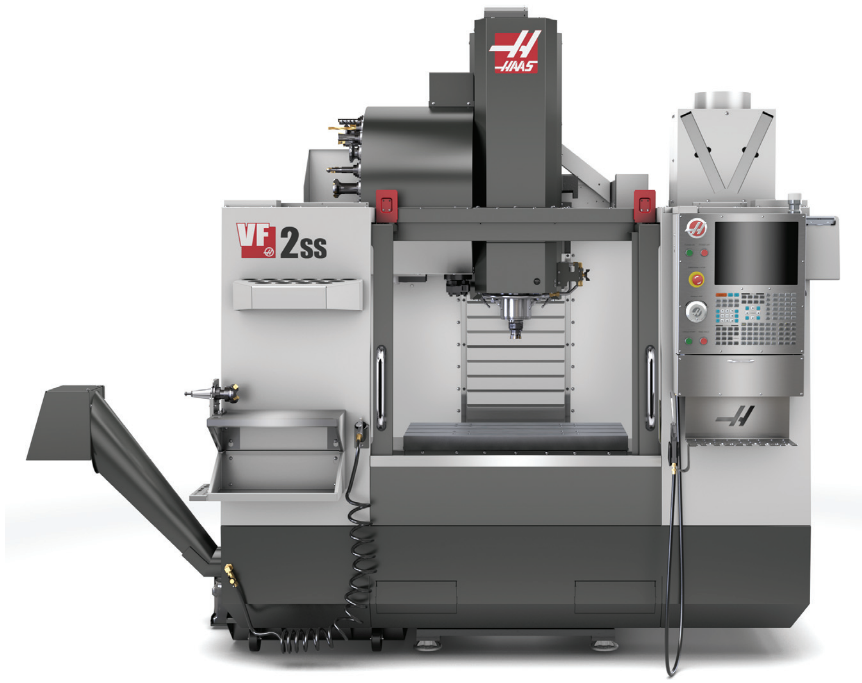


Figure 24.20: A computer numerical-control (CNC) vertical-spindle milling machine. This is one of the most versatile machine tools. *Source:* Haas Automation, Inc.

are generally used for this purpose (see also Section 37.8 on *flexible fixturing*). Mounted and clamped to the worktable using the T-slots seen in Fig. 24.18a and b, vises are used for small production runs on small parts, while *fixtures* are used for higher production runs, and can be automated by various mechanical and hydraulic means.

Accessories for milling machines include various fixtures and attachments for the machine head and the worktable, designed to adapt them to different operations. The accessory that has been used most

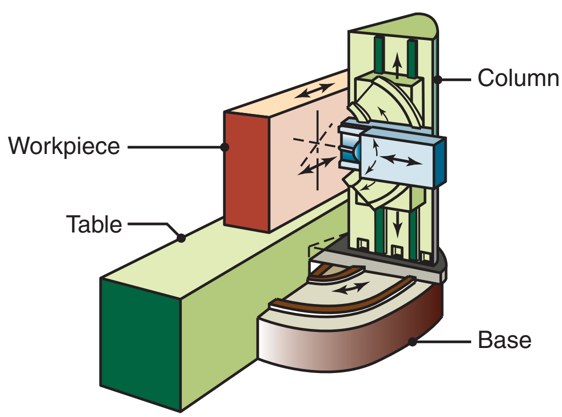


Figure 24.21: Schematic illustration of a five-axis profile milling machine. Note that there are three principal linear and two angular movements of machine components.

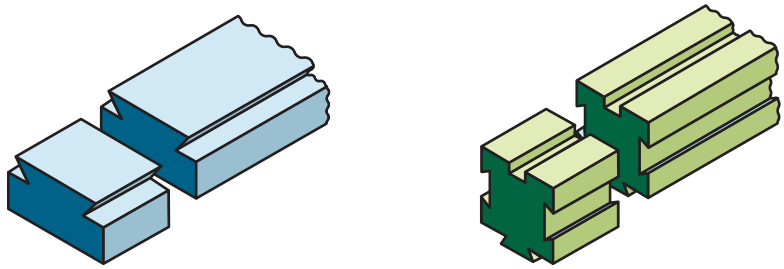


Figure 24.22: Typical parts that can be made on a planer.

commonly is the *universal dividing* (**index**) *head*. Manually operated, this fixture rotates (*indexes*) the workpiece to specified angles between individual machining steps. Typically, it has been used to mill parts with polygonal surfaces and to machine gear teeth.

24.3 Planing and Shaping

Planing is a relatively simple machining operation by which flat surfaces, as well as cross sections with grooves and notches, can be made along the length of the workpiece (Fig. 24.22). Planing is usually done on workpieces as large as 25 m $\times$ 15 m, although 10 m is more typical. In a **planer**, also called a *scalper* when a layer is machined from a cast ingot, the workpiece is mounted on a table that travels back and forth along a straight path. A horizontal *cross-rail*, which moves vertically along the ways of the column, is equipped with one or more tool heads. The cutting tools are mounted on the heads, and machining is done along a straight path. In order to prevent the cutting edges from chipping when tools rub along a workpiece during the return stroke, tools are either tilted or lifted mechanically or hydraulically at each stroke.

Because of the reciprocating motion of the workpiece, the noncutting time elapsed during the return stroke is significant. Consequently, these operations are neither efficient nor economical, except for low-quantity production, which is generally the case for large and long workpieces. The efficiency of the operation can be improved by equipping planers with toolholders and tools that cut in both directions of table travel. Also, because of the length of the workpiece, it is essential to equip cutting tools with chip breakers, as otherwise the chips produced can be very long, thus interfering with the machining operation and becoming a safety hazard.

Shaping. Machining by shaping is basically the same as by planing, except that it is the tool and not the workpiece that travels; the workpieces are smaller, typically less than 1 m $\times$ 2 m of surface area. In a **horizontal shaper**, the cutting tool travels back and forth along a straight path. The tool is attached to the tool head, which is mounted on the ram; the ram has a reciprocating motion.

In most machines, cutting is done during the forward movement of the ram (*push cut*); in others, it is done during the return stroke of the ram (*draw cut*). Vertical shapers (called **slotters**) are used to machine notches, keyways, and dies. Because of low production rates, only special-purpose shapers (such as gear shapers, Section 24.7.2) are in common use today.

24.4 Broaching and Broaching Machines

Broaching is similar to shaping using a long, multiple-tooth cutter, and is used to machine internal and external surfaces, such as holes with circular, square, or irregular cross sections; keyways; the teeth of internal gears; multiple spline holes; and flat surfaces (Fig. 24.23). In a typical **broach** (Fig. 24.24a), the total depth

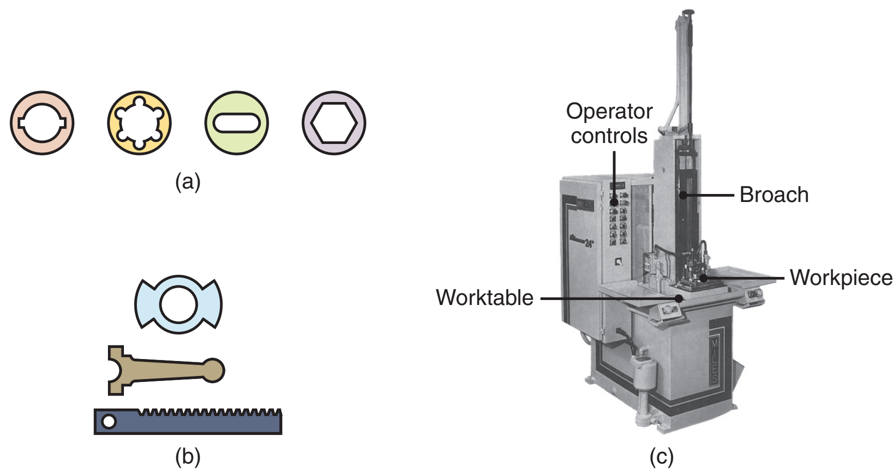


Figure 24.23: (a) Typical parts made by internal broaching. (b) Parts made by surface broaching. (c) Vertical broaching machine. *Source:* (a) and (b) Courtesy of General Broach and Engineering Company, (c) Courtesy of Ty Miles, Inc.

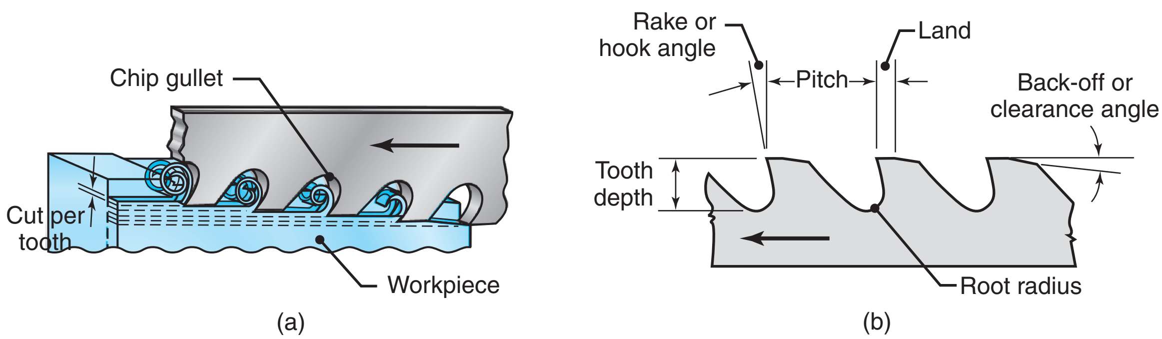


Figure 24.24: (a) Cutting action of a broach, showing various features. (b) Terminology for a broach.

of material removed in one stroke is the sum of the depths of cut of each tooth of the broach. A large broach can remove material as deep as 38 mm in one stroke.

Broaching is an important production process, and can produce parts with good surface finish and dimensional accuracy. It competes favorably with other machining processes, such as boring, milling, shaping, and reaming, to produce similar shapes. Although broaches can be expensive, the cost is justified with high-quantity production runs.

Broaches. The terminology for a typical broach is given in Fig. 24.24b. The *rake* (hook) *angle* depends on the material cut (as it does in turning and other cutting operations) and usually ranges from 0° to 20°. The *clearance angle* is typically 1° to 4°; finishing teeth have smaller angles. Too small a clearance angle causes rubbing of the teeth against the broached surface. The *pitch* of the teeth depends on factors such as the length of the workpiece (length of cut), tooth strength, and size and shape of chips.

The tooth depth and pitch must be sufficiently large to accommodate the chips produced during broaching, particularly for long workpieces. At least two teeth should be in contact with the workpiece at all times. The following formula may be used to obtain the pitch for a broach to cut a surface of length l:

$$\text{Pitch} = k\sqrt{l}, \tag{24.6}$$

where k is a constant, equal to 1.76 when l is in mm. An average pitch for small broaches is in the range from 3.2 to 6.4 mm, and for large ones it is in the range from 12.7 to 25 mm. The depth of cut per tooth depends on the workpiece material and the surface finish required. It is usually in the range from 0.025 to 0.075 mm for medium-sized broaches, but can be larger than 0.25 mm for larger broaches.

Broaches are available with various tooth profiles, including some with *chip breakers* (Fig. 24.25). The variety of *surface* broaches include *slab* (for cutting flat surfaces), *slot*, *contour*, *dovetail*, *pot* (for precision external shapes), and *straddle*. *Internal broach* types include *hole* (for close-tolerance holes, round shapes, and other shapes; Fig. 24.26), *keyway*, *internal gear*, and *rifling* (for gun barrels). Irregular internal shapes usually are broached by starting with a round hole drilled or bored in the workpiece.

Turn Broaching. Turn broaching is a combination of *shaving* and *skiving* (removing a thin layer of material with a specially shaped cutting tool). The process is typically used for broaching the bearing surfaces of crankshafts and similar parts. The shaft is rotated between centers, and the broach, equipped with multiple carbide inserts, passes tangentially across the bearing surfaces and removes material. Straight as well as circular broaches can be used successfully in turn broaching, including simultaneously broaching a number of crankshafts.

Broaching Machines. These machines are relatively simple in construction, as they have linear motions only; they usually are actuated hydraulically, although some are moved by crank, screw, or rack mechanisms. Several styles of machines are available, and sizes range from machines for making needle-like parts to those used for broaching gun barrels, including rifling (producing internal spiral grooves).

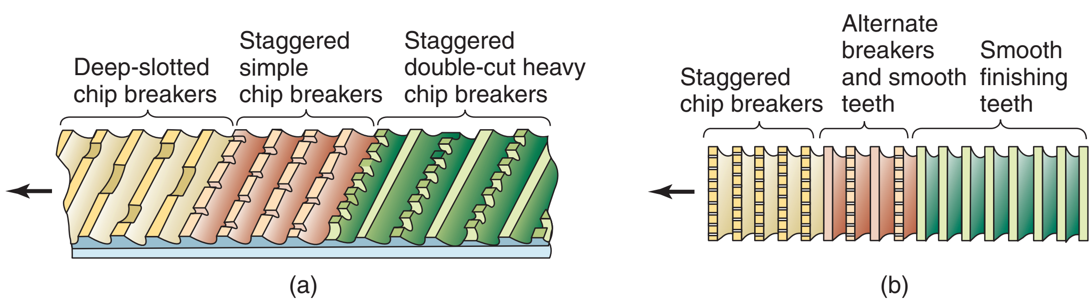


Figure 24.25: Chip breaker features on (a) a flat broach and (b) a round broach.

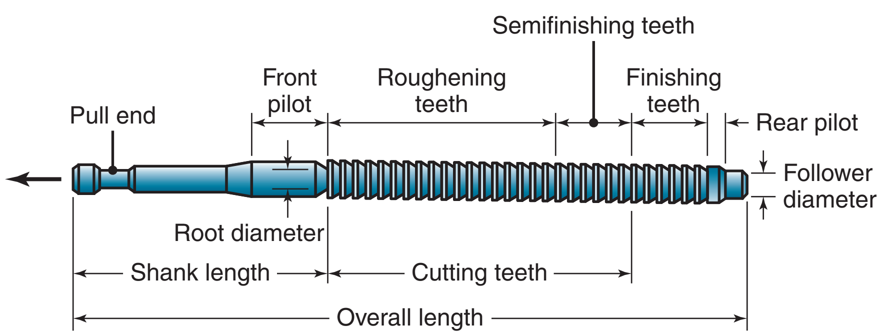


Figure 24.26: Terminology for a pull-type internal broach used for enlarging long holes.

Broaching machines, either pull or push type, are either horizontal or vertical. *Push broaches* usually are shorter, with lengths in the range of 150–350 mm. *Pull broaches* tend to straighten a hole, whereas a push broach permits the broach to follow any irregularity of the leader hole. The *force* required to pull or push a broach depends on the (a) strength of the workpiece material, (b) total periphery of the cut, (c) cutting speed, (d) tooth profile, and (e) type of cutting fluid used. The pulling force capacities of broaching machines are as high as 0.9 MN (90 metric tons).

Process Parameters. Cutting speeds for broaching may range from 1.5 m/min for high-strength alloys to as much as 30 m/min for aluminum and magnesium alloys. The most common broach materials are M2 and M7 high-speed steels and carbide inserts. Smaller, high-speed steel blanks for making broaches can be produced by powder-metallurgy techniques (Chapter 17) for better control of quality. Cutting fluids generally are recommended, especially for internal broaching.

Design Considerations. Broaching, as with other machining processes, requires that certain guidelines be followed in order to obtain economical and high-quality production. The major requirements are:

- Blanks should be designed and prepared so that they can be securely clamped in broaching machines, and should have sufficient structural strength and stiffness to withstand the cutting forces during broaching.
- Keyways, splines, gear teeth, etc., should all have standard sizes and shapes, so as to allow the use of common broaches.
- Balanced cross sections are preferable to keep the broach from drifting laterally, thus maintaining close tolerances.
- Radii are difficult to broach and chamfers are preferred; inverted or dovetail splines should be avoided.
- Blind holes should be avoided whenever possible, but if necessary, there must be a relief at the end of the area to be broached.

Case Study 24.1 Broaching Internal Splines

The part shown in Fig. 24.27 is made of nodular iron (65-45-15; Section 12.3.2), with internal splines, each 50 mm long. The splines have 19 involute teeth, with a pitch diameter of 63.52 mm. An M2 high-speed steel broach, with 63 teeth, a length of 1.448 m, and a diameter the same as the pitch diameter, was used to produce the splines. The cut per tooth was 0.116 mm. The production rate was 63 pieces per hour. The number of parts per grind was 400, with a total broach life of about 6000 parts.

Source: ASM International.

24.5 Sawing

Sawing is a common process dating back to around 1000 B.C. The *blade* or saw has a series of small teeth, each removing a small amount of material with each stroke of the saw. Sawing can be used for all materials and is capable of producing any shape (Fig. 24.28). It is an efficient material-removal process and can produce near-net shapes from blanks. The width of cut (**kerf**) in sawing usually is small, so that the process wastes relatively little material.

Typical saw-tooth and saw-blade configurations are shown in Fig. 24.29, where tooth spacing is generally in the range from 0.08 to 1.25 teeth per mm. A wide variety of sizes, tooth forms, tooth

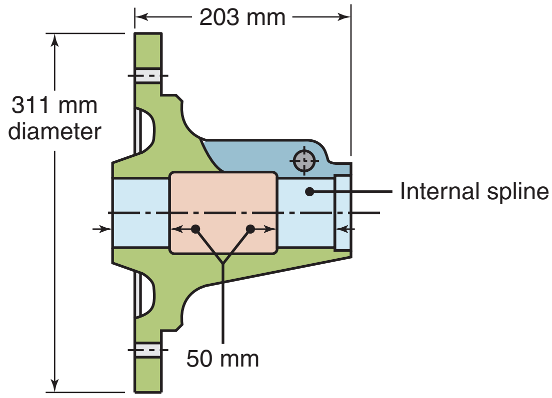


Figure 24.27: Example of a part with internal splines that were produced by broaching.

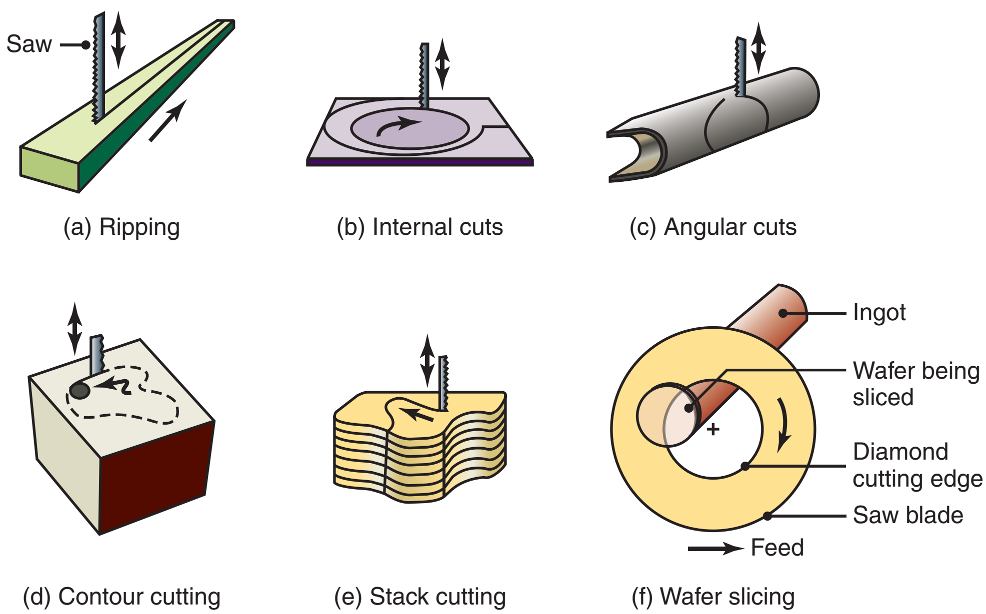


Figure 24.28: Examples of various sawing operations.

spacing, and blade thicknesses and widths are available. Saw blades generally are made from high-carbon and high-speed steels (M2 and M7); carbide or high-speed-steel tipped steel blades are used for sawing harder materials and at higher speeds (Fig. 24.30).

The **tooth set** in a saw (Fig. 24.29b) is important in providing a sufficiently wide kerf for the blade to move freely in the workpiece without binding or excessive frictional resistance and thus reducing the heat generated. Elevated temperatures can have adverse effects on the material cut, especially for thermoplastics, which soften rapidly when temperature rises (see Fig. 7.11). The tooth set also allows the blade to track a path accurately, following the pattern to be cut without wandering. At least two or three teeth should always be engaged with the workpiece, in order to prevent *snagging* (catching of the saw tooth on the workpiece). This is the reason why thin materials, especially sheet metals, can be difficult to saw. The thinner the stock, the finer the saw teeth should be, and the greater the number of teeth per unit length of the saw. Cutting fluids are generally used to improve the quality of the cut and the life of the saw.

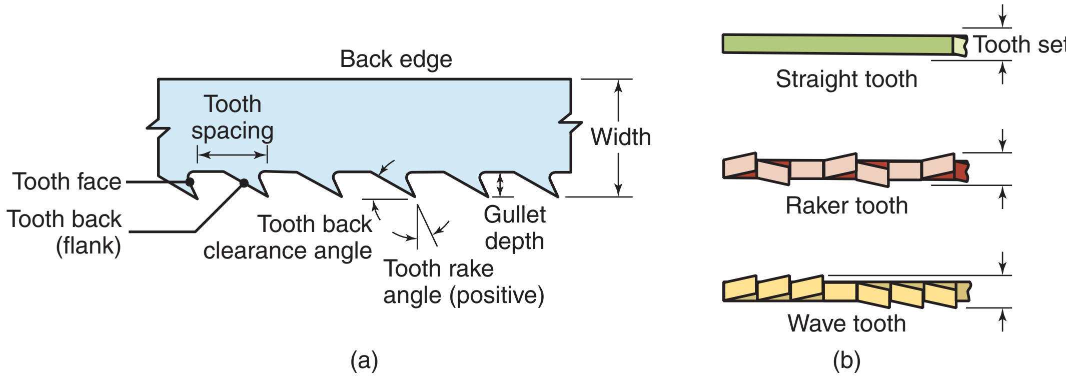


Figure 24.29: (a) Terminology for saw teeth. (b) Types of tooth sets on saw teeth staggered to provide clearance for the saw blade to prevent binding during sawing.

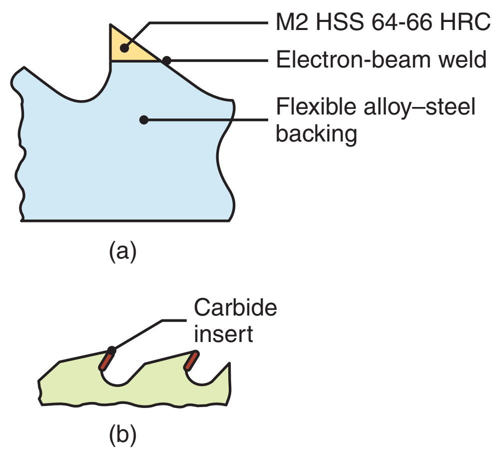


Figure 24.30: (a) High-speed-steel teeth welded onto a steel blade. (b) Carbide inserts brazed to blade teeth.

Types of Saws. *Hacksaws* have straight blades and reciprocating motions. Developed in the 1650s, they generally are used to cut off bars, rods, and structural shapes; they may be manual or power operated. Because cutting takes place during only one of the strokes, hacksaws are not as efficient as band saws (see below). *Power hacksaw* blades are usually 1.2 to 2.5 mm thick, and up to 610 mm long. The rate of strokes ranges from 30 per minute for high-strength alloys to 180 per minute for carbon steels. The hacksaw frame in power hacksaws is weighted by various mechanisms, applying as much as 1.3 kN of force to the workpiece to improve the cutting rate. *Hand hacksaw* blades are thinner and shorter than power hacksaw blades, which have as many as 1.2 teeth per mm for sawing sheet metal and thin tubing.

Circular saws, also called *cold saws* for cutting metals, generally are used for high-production-rate sawing, called *cutting off*. Cutting-off operations also can be carried out with thin, *abrasive* disks, as described in Section 26.4. Cold sawing is common in industry, particularly for cutting off large cross sections. These saws are available with a variety of tooth profiles and sizes. In modern machines, cutting off with circular saws produce relatively smooth surfaces, with good thickness control and dimensional accuracy because of the stiffness of the machines and of the saws. The *inner-diameter-cutting saw*, shown in Fig. 24.28f, is widely used to cut single-crystal silicon wafers in microelectronic devices (Section 28.4).

Band saws have continuous, long, flexible blades. *Vertical band saws* are used for straight as well as *contour cutting* of flat sheets and other parts, supported on a horizontal table (Fig. 24.28d). Also available are *computer-controlled* band saws, with the capability of guiding the contour path automatically. *Power band saws* have higher productivity than power hacksaws, because of their continuous cutting action. Using high-speed-steel blades, speeds for sawing high-strength alloys are up to about 60 m/min, and 120 m/min for carbon steels.

Blades and high-strength wire can be coated with diamond powder (**diamond-edged blades** and **diamond-wire saws**), in which diamond particles act as cutting teeth (*abrasive cutting*); carbide particles also are used for this purpose. The blades and wires are suitable for sawing hard metallic, nonmetallic, and composite materials. Wire diameters range from 13 mm for use in rock cutting to 0.08 mm for precision cutting. Hard materials also can be sawed with thin, *abrasive disks* and with advanced machining processes (Chapter 27).

Friction Sawing. In *friction sawing*, a mild-steel blade or disk rubs against the workpiece, at speeds of up to 7600 m/min. The frictional energy is converted into heat, which rapidly softens a narrow zone in the workpiece. The action of the blade, which can have teeth or notches for higher cutting efficiency, pulls and ejects the softened metal from the cutting zone. The heat generated produces a *heat-affected zone* (Section 30.9) on the cut surfaces; the workpiece properties along the cut edges can thus be affected adversely by this process. Because only a small portion of the blade is engaged with the workpiece at any time, the blade itself cools rapidly as it passes through the air.

The friction-sawing process is suitable for hard ferrous metals and reinforced plastics, but not for non-ferrous metals because of their tendency to stick to the blade. Friction sawing is commonly used to also remove flash from castings. Disks as large as 1.8 m in diameter are used to cut off large steel sections.

24.6 Filing

Filing involves small-scale removal of material from a surface, corner, edge, or hole, including the removal of burrs (see Fig. 16.2). First developed around 1000 B.C., files are usually made of hardened steel, and are available in a variety of cross sections, such as flat, round, half-round, square, and triangular. They can have several tooth forms and coarseness grades. Although filing is usually done by hand, *filing machines*, with automatic features, are available for high production filing, with files reciprocating at up to 500 strokes/min.

Band files consist of file *segments*, each about 75 mm long and riveted to a flexible steel band; they are used in a manner similar to band saws. *Disk-type files* also are available. *Rotary files* and *burs* (Fig. 24.31) are used for such applications as deburring, removing scale from surfaces, producing chamfers on parts, and removing small amounts of material in die making. These tools generally are conical, cylindrical, or spherical in shape, and have various tooth profiles. Their cutting action (similar to that of reamers, Section 23.6) removes small amounts of material at high rates. The rotational speed of burs ranges from 1500 rpm for cutting steels (using large burs) to as high as 45,000 rpm for magnesium (small burs).

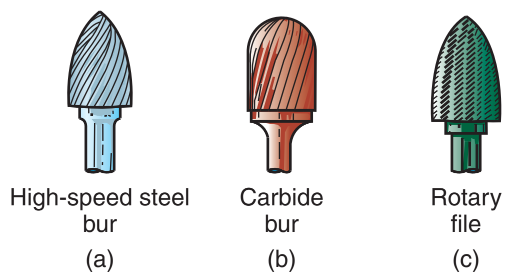


Figure 24.31: Types of burs used in burring operations.

24.7 Gear Manufacturing by Machining

Several processes for making gears or producing individual gear teeth were described in Parts II and III, involving such processes as casting, forging, extrusion, drawing, thread rolling, and powder metallurgy. *Blanking* of sheet metal also can be used for making thin gears, such as those used in mechanical watches, clocks, and similar mechanisms. Plastic gears can be made by casting (Chapter 11) or injection molding (Section 19.3).

Gears may be as small as those used in watches or as large as 9 m in diameter for rotating mobile crane superstructures and mining equipment. The dimensional accuracy and surface finish required for gear teeth depend on their particular intended use. Poor gear-tooth quality contributes to inefficient energy transmission, increased vibration and noise, and adversely affects the gear's friction and wear characteristics. Submarine gears, for example, have to be of extremely high quality so as to reduce noise levels.

The standard nomenclature for an involute spur gear is shown in Fig. 24.32. Starting with a wrought or cast gear blank, there are two basic methods of making gear teeth: *form cutting* and *generating*.

24.7.1 Form Cutting

In *form cutting*, the cutting tool is similar to a form-milling cutter made in the shape of the space between the gear teeth (Fig. 24.33a). The gear-tooth shape is reproduced by machining the gear blank around its periphery. The cutter travels axially along the length of the gear tooth, and at the appropriate depth to produce the gear-tooth profile. After each tooth is cut, the cutter is withdrawn, the gear blank is rotated (**indexed**), and the cutter proceeds to cut another tooth. This process continues in a cycle until all of the teeth are machined. Each cutter is designed to cut a range of numbers of teeth.

The precision of a form-cut tooth profile depends on the cutter accuracy and on the machine and its stiffness. Because the cutter has a fixed geometry, form cutting can be used only to produce gear teeth that have a constant width, that is, on spur or helical gears but not on bevel gears. Internal gears and gear teeth

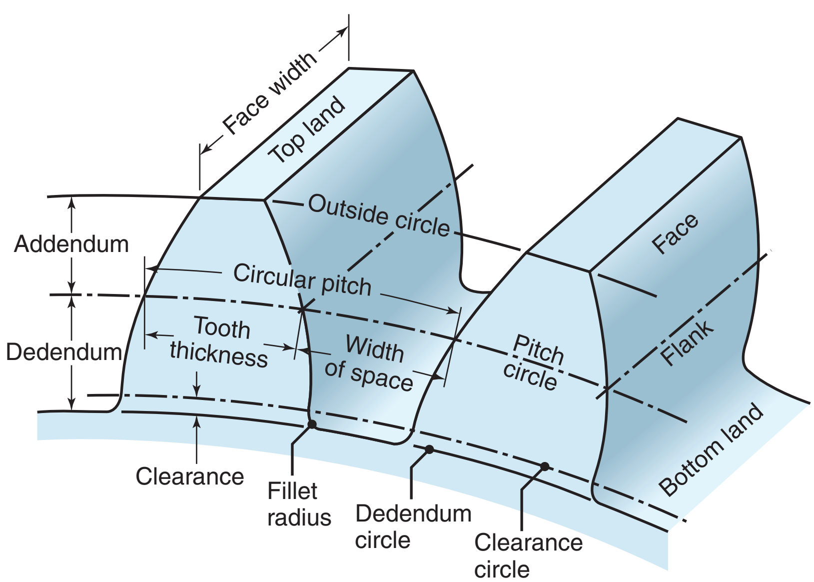


Figure 24.32: Nomenclature for an involute spur gear.

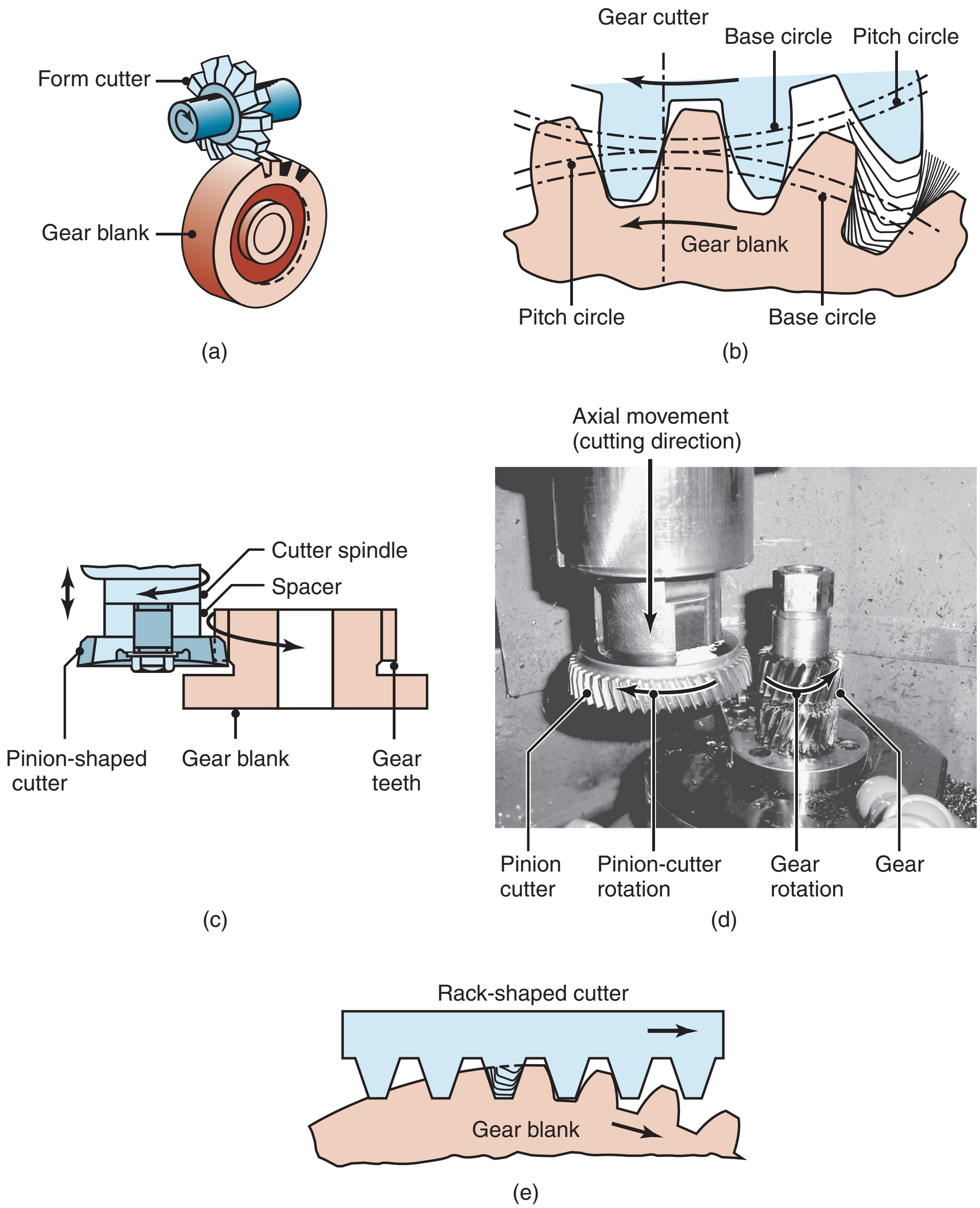


Figure 24.33: (a) Producing gear teeth on a blank by form cutting. (b) Schematic illustration of gear generating with a pinion-shaped gear cutter. (c) and (d) Gear generating in a gear shaper using a pinion-shaped cutter. Note that the cutter reciprocates vertically. (e) Gear generating with rack-shaped cutter. *Source:* (d) Courtesy of Steve Schmid, with permission from Schafer Gear Works.

on straight surfaces, such as those in a rack and pinion, are form cut with a shaped cutter, on a machine similar to a shaper.

Broaching can be used to machine gear teeth and is particularly suitable for producing internal teeth. The process is rapid and produces fine surface finish with high dimensional accuracy. However, because a different broach is required for each gear size and broaches are expensive, this method is suitable almost exclusively for high-quantity production.

Although inefficient, form cutting also can be done on milling machines, with the cutter mounted on an arbor and the gear blank mounted in a dividing head. Gear teeth also may be cut on special machines, with a single-point cutting tool guided by a *template*, made in the shape of the gear-tooth profile. Because the template can be made much larger than the gear tooth itself, dimensional accuracy is improved.

Form cutting is a relatively simple process and can be used for cutting gear teeth with various profiles. Nonetheless, it is a slow operation and some types of machines require skilled labor. Machines with semiautomatic features can be used economically for form cutting on a limited-production basis; generally, however, form cutting is suitable only for low-quantity production.

24.7.2 Gear Generating

The cutting tool used in *gear generating* may be a pinion-shaped cutter, a rack-shaped straight cutter, or a hob. Gear-generating machines also can produce spiral-bevel and hypoid gears. Like most other machine tools, modern gear-generating machines are now *computer controlled*. *Multiaxis* computer-controlled machines are capable of generating several types and sizes of gears, using indexable milling cutters.

1. A **pinion-shaped cutter** can be considered as one of the two gears in a conjugate pair, with the other being the gear blank (Fig. 24.33b). This type of cutter is used on vertical *gear shapers* (Fig. 24.33c and d). The cutter has an axis parallel to that of the gear blank, and rotates slowly with the blank at the same pitch–circle velocity and in an axial-reciprocating motion. A train of gears provides the required relative motion between the cutter shaft and the gear-blank shaft. Cutting may take place at either the downstroke or the upstroke of the machine.

   Because the clearance required for the cutter travel is small, gear shaping is suitable for gears that are located close to obstructing surfaces, such as the flange in the gear blank shown in Fig. 24.33c and d. This process can be used for low-quantity as well as high-quantity production.

2. On a **rack shaper**, the generating tool is a *segment of a rack* (Fig. 24.33e), which reciprocates parallel to the axis of the gear blank. Because it is not practical to have more than 6 to 12 teeth on a rack cutter, the cutter must be disengaged at suitable intervals, and returned to the starting point. The gear blank remains fixed during this operation.

3. A **hob** (Fig. 24.34) is basically a gear-cutting worm, or screw, made into a gear-generating tool by a series of longitudinal slots or gashes machined into it to form the cutting teeth. When hobbing a spur gear, the angle between the hob and gear-blank axes is 90° minus the lead angle at the hob threads. All motions in hobbing are rotary, and the hob and gear blank rotate continuously, much as two gears mesh, until all of the teeth are cut.

Hobs are available with one, two, or three threads. For example, if the hob has a single thread and the gear is to have 40 teeth, the hob and the gear spindle must be geared together such that the hob makes 40 revolutions while the gear blank makes 1 revolution. Similarly, if a double-threaded hob is used, the hob would make 20 revolutions to the gear blank's 1 revolution. In addition, the hob must be fed parallel to the gear axis, for a distance greater than the face width of the gear tooth (Fig. 24.32), in order to produce straight teeth on spur gears. The same hobs and machines can be used to cut helical gears, by tilting the axis of the hob spindle.

Because it produces a variety of gears at high rates and with good dimensional accuracy, gear hobbing is used extensively in industry. Although the process is suitable also for low-quantity production, it is most economical for medium- to high-quantity production.

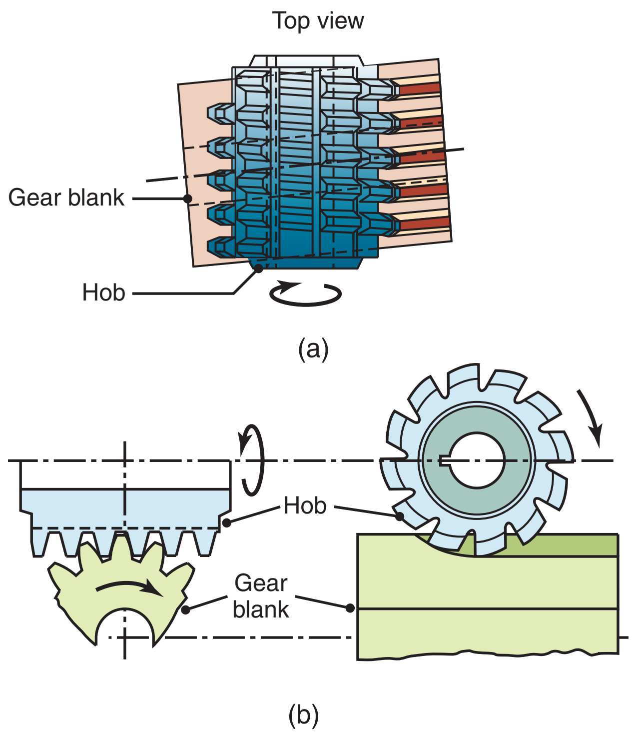


Figure 24.34: (a) Hobs, used to machine gear teeth; (b) schematic illustration of gear cutting with a hob. *Source:* (a) Courtesy of Sandvik Coromant.

24.7.3 Cutting Bevel Gears

Straight bevel gears generally are roughed out in one cut with a form cutter, on machines that index automatically; the gear is then finished on a gear generator. The cutters reciprocate across the face of the bevel gear, as does the tool on a shaper (Fig. 24.35a). The machines for *spiral bevel gears* operate essentially on the

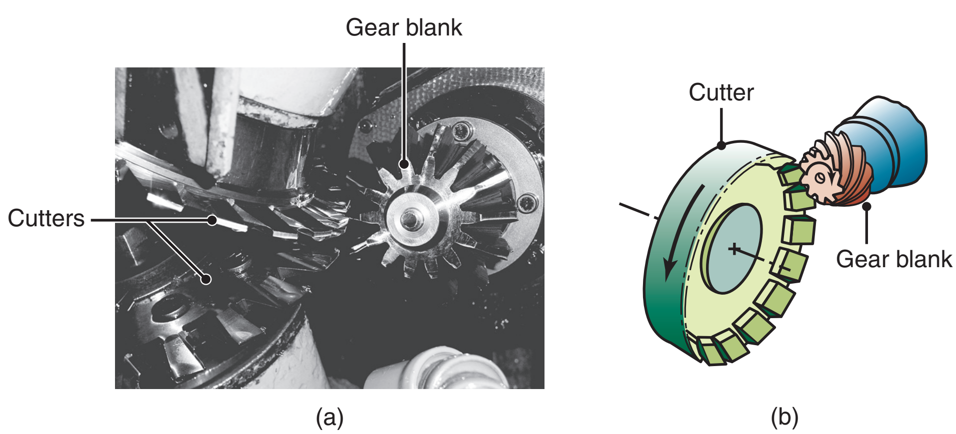


Figure 24.35: (a) Cutting a straight bevel-gear blank with two cutters. (b) Cutting a helical bevel gear. *Source:* Courtesy of Steve Schmid, with permission from Schafer Gear Works.

same principle, in which the spiral cutter is basically a face-milling cutter, with a number of straight-sided cutting blades protruding from its periphery (Fig. 24.35b).

24.7.4 Gear-finishing Processes

As produced by any of the processes described, the surface finish and dimensional accuracy of gear teeth may still be insufficient for some applications. Several *finishing processes* are available to improve the surface quality of the gears, the choice being dictated by the method of gear manufacture, the desired performance, and whether the gears have been hardened by heat treatment. As described in Chapter 4, heat treating can cause distortion of parts; consequently, for a precise gear-tooth profile, heat-treated gears typically are subjected to finishing operations.

Shaving. The gear-shaving process involves a cutter made in the exact shape of the finished tooth profile, which removes very small amounts of material from the surface of the gear teeth (see also Fig. 16.9). The cutter, which has a reciprocating motion, has teeth that are *slotted* or gashed at several points along its width, making the process similar to fine broaching. Shaving and burnishing (described next) can be performed only on gears with a hardness of 40 HRC or lower.

Although the tooling is expensive and special machines are required, shaving is rapid and is the most commonly used process for gear finishing. It produces gear teeth with improved surface finish and good dimensional accuracy of the tooth profile. Shaved gears may subsequently be heat treated and then ground, for improved hardness, wear resistance, and a more accurate tooth profile.

Burnishing. The surface finish of gear teeth also can be improved by burnishing. Introduced in the 1960s, burnishing is basically a surface plastic-deformation process (see Section 34.2), using a gear-shaped burnishing and specially hardened die, that subjects the tooth surfaces to a surface-rolling action, called **gear rolling**. The resulting cold working of the tooth surfaces not only improves the surface finish, but also induces compressive residual stresses on the surfaces of the gear teeth, thus improving their fatigue life. It has been shown, however, that burnishing does not significantly improve the dimensional accuracy of the gear tooth. With powder-metallurgy gears, burnishing leads to surface densification, with a significant improvement in performance (see Section 17.5).

Grinding, Honing, and Lapping. For the highest dimensional accuracy in tooth spacing and form, and for superior surface finish, gear teeth may be ground, honed, and lapped (Chapter 26). Specially dressed grinding wheels are used for either forming or for generating gear-tooth surfaces. There are several types of grinders, with the single-index form grinder being the most commonly used. In **form grinding**, the shape of the grinding wheel is identical to that of the tooth spacing (Fig. 24.36a). In **generating**, the grinding wheel acts in a manner similar to a gear-generating cutter, described previously (Fig. 24.36b).

The honing process is faster than grinding, and is used to improve surface finish. The **honing** tool is a plastic gear impregnated with fine abrasive particles. To further improve the finish, ground gear teeth are **lapped**, using abrasive compounds either with (a) a gear-shaped lapping tool made of cast iron or bronze, or (b) a pair of mating gears that are run together. Although production rates are lower and costs are higher, these finishing operations are particularly suitable for making hardened gears of very high quality, long life, and quiet operation.

24.7.5 Design Considerations and Economics of Gear Machining

Design considerations for gear-cutting operations are summarized as follows:

- Gears should preferably be machined prior to their assembly on shafts; wide gears are more difficult to machine than narrow ones.
- Sufficient clearance should be provided between gear teeth and flanges, shoulders, and other features of the part, so that the cutting tool can machine without any interference.

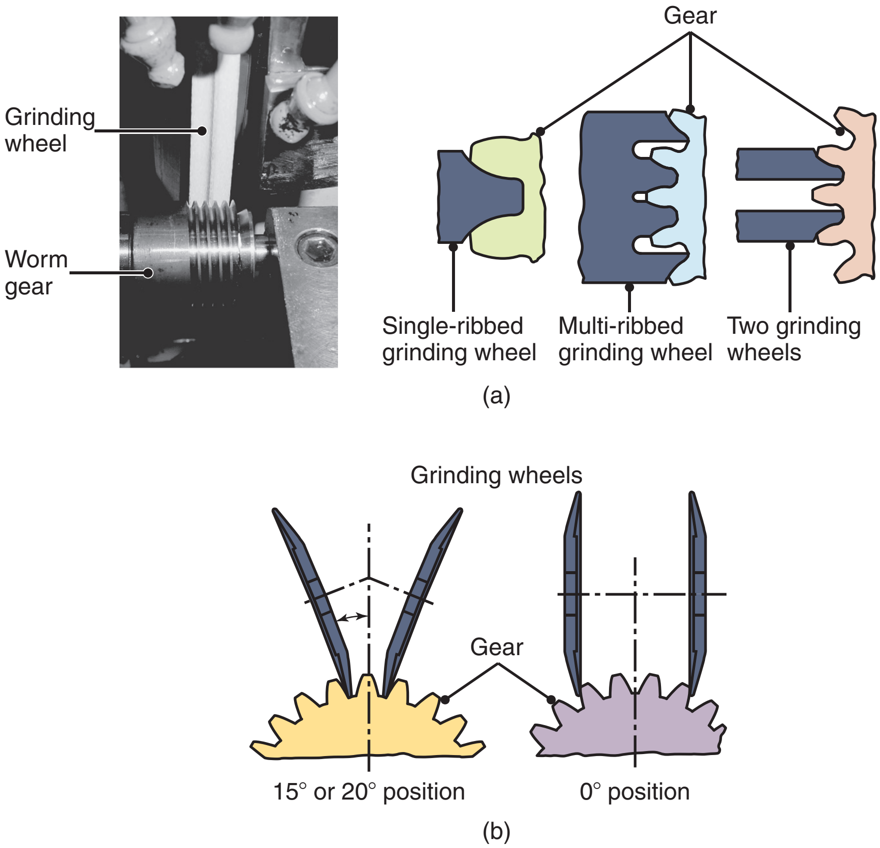


Figure 24.36: Finishing gears by grinding: (a) form grinding with shaped grinding wheels; (b) grinding by generating, using two wheels.

- Blank design is important for proper fixturing and to simplify cutting operations. Machining allowances must be provided in the blanks, and if machining is to be followed by finishing operations, the part must still be oversized after machining; that is, it must have a finishing allowance after being machined.

- Spur gears are easier to machine than helical gears, which, in turn, are easier to machine than bevel gears or worm gears.

- Dimensional tolerances and standardized gear shapes are specified by industry standards. A gear quality number should be selected so that the gear has as wide a tolerance range as possible, while still meeting performance requirements in service.

Economics. As in all machining operations, the cost of gears increases rapidly with improved surface finish and gear quality. Figure 24.37 shows the relative manufacturing cost of gears as a function of quality, as specified by the American Gear Manufacturers Association (AGMA). The higher the number, the higher is the dimensional accuracy of the gear teeth. As noted in this figure, the manufacturing cost can vary by an order of magnitude, depending on dimensional tolerances.

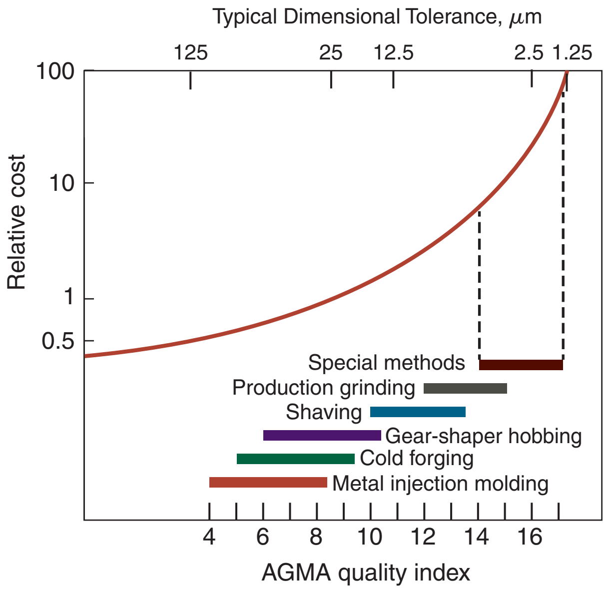


Figure 24.37: Gear manufacturing cost as a function of gear quality.

Case Study 24.2 Ping Golf Putters

In their efforts to develop high-end, top performing putters, engineers at Ping Golf, Inc., utilized advanced machining practices in their design and production processes for a new style of putter, called the Anser® series, shown in Fig. 24.38. In response to a unique set of design constraints, they had the task and goal of creating putters that would be practical for production quantities and also meet specific functional and aesthetic requirements.

One of the initial decisions concerned the selection of a proper material for the putter to meet its functional requirements. Four types of stainless steel (303, 304, 416, and 17-4 precipitation hardening; see Section 5.6) were considered for various property requirements, including machinability, durability, and the sound or feel of the particular putter material, another requirement that is unique to golf equipment. Among the materials evaluated, 303 stainless steel was chosen because it is a free-machining steel (Section 21.7), indicating that in machining it produces smaller chips, lower power consumption, better surface finish, and improved tool life, thus allowing for increased machining speeds and higher productivity.

The next step of the project involved determining the optimum blank shape and the sequence of operations to be performed during its production. For this case, engineers chose to develop a slightly oversized forged blank (Chapter 14). A forging was chosen because it provided a favorable internal grain structure as opposed to a casting, which could result in porosity and an inconsistent surface finish after machining. The blank incorporated a machining allowance, whereby dimensions were specified approximately 1.25–1.9 mm larger in all directions than that of the final part.

The most challenging, and longest, task was developing the necessary programming and designing fixtures for each part of the putter. Beyond the common requirements of typical machined parts, including tight tolerances and repeatability, putters require an additional set of aesthetic specifications. In this case, both precise machining and the right overall appearance of the finished part were imperative. A machining technique known as surfacing or contouring (commonly used in making injection molds) was used to machine most of the finished geometry. Although this operation required additional machining,

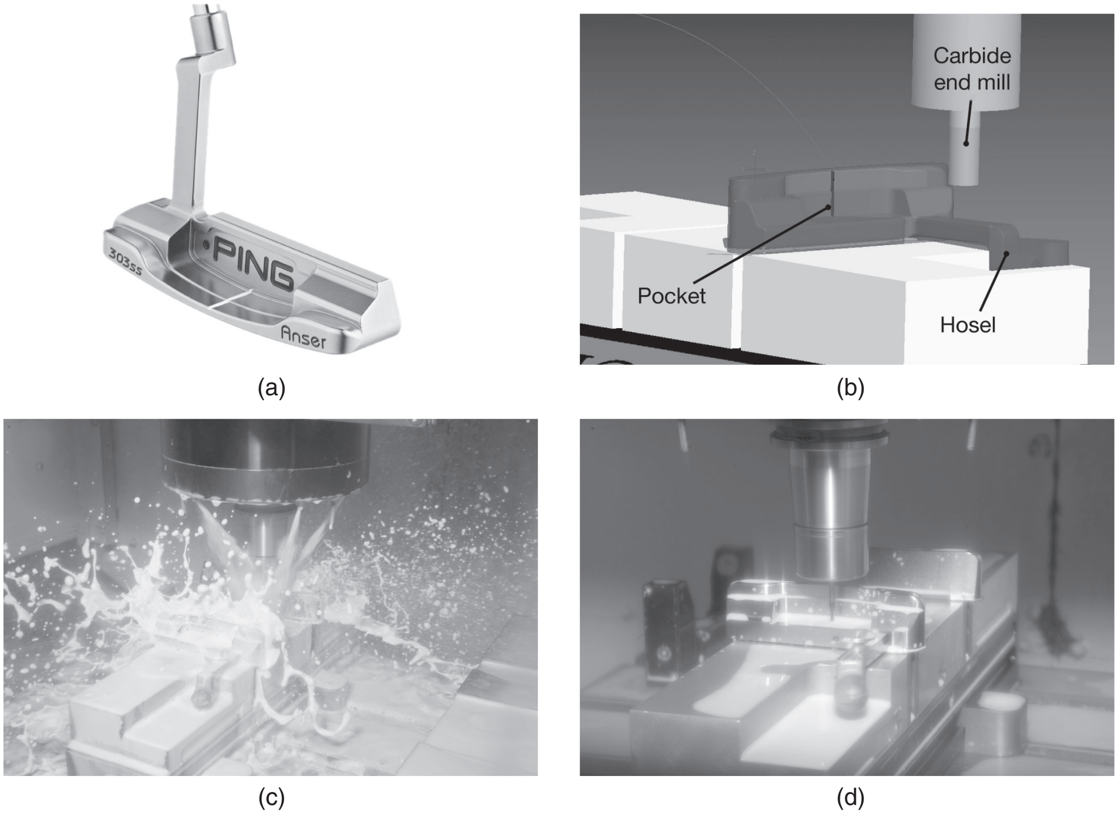


Figure 24.38: (a) The Ping Anser® golf putter; (b) CAD model of rough machining of the putter outer surface; (c) rough machining on a vertical machining center; (d) machining of the lettering in a vertical machining center; the operation was paused to take the photo, as normally the cutting zone is flooded with a coolant. *Source:* Courtesy of Ping Golf, Inc., Phoenix, Arizona.

it provided a superior finish on all surfaces and allowed machining of more complex geometries, thus adding value to the finished product.

As for all high-volume machined parts, repeatability was essential. Each forged blank was designed with a protrusion across the face of the putter, allowing for the initial locating surfaces, for ease of fixturing. A short machining operation removed a small amount of material around the bar and produced three flat, square surfaces, as a reference location for the first primary machining operation.

Each putter required six different operations in order to machine all of its surfaces, and each operation was designed to provide locating surfaces for the next step in the manufacturing process. Several operations were set up using a tombstone loading system (Section 37.8) on a horizontal-spindle CNC milling machine. This method allowed machine operators to load and unload parts while other parts were being machined, thus significantly increasing the efficiency of the operation.

Modular fixturing and using tungsten-carbide cutting tools coated with TiAlN (Section 22.5) allowed for the quick changeover between right- and left-handed parts, as well as different putter models. After the initial locating operation was complete, the parts were transferred to a three-axis vertical machining center (see, for example, Fig. 25.7) to cut the putter cavity. Since the forged blanks were near net shape, the maximum radial depth of cut on most surfaces was 1.9 mm, but the axial depth of cut of 37.5 mm inside the cavity of the putter was the most demanding milling operation (see Fig. 24.38b and c). The putter has small inside radii with a comparatively long depth (7× the diameter or greater).

A four-axis horizontal machining center (see, for example, Fig. 25.2) was used to reduce the number of setups in this operation. The rotary axis was used for creating the relatively complex geometry of the hosel (the socket for the shaft of the golf club). Since the hosel is relatively unsupported, chatter was the most complex challenge to overcome. Several iterations of spindle speeds were attempted in conjunction with upfront guidance from a simulation model. Modal analyses were conducted on the fixtured parts, in an attempt to identify and avoid the natural frequencies of the part or the fixture (see Section 25.4). The machines had spindle speeds ranging from 12,000 to 20,000 rpm, each having 22 kW. With the near-net-shape forging, the milling operations were designed to have low depths of cut, but high speed.

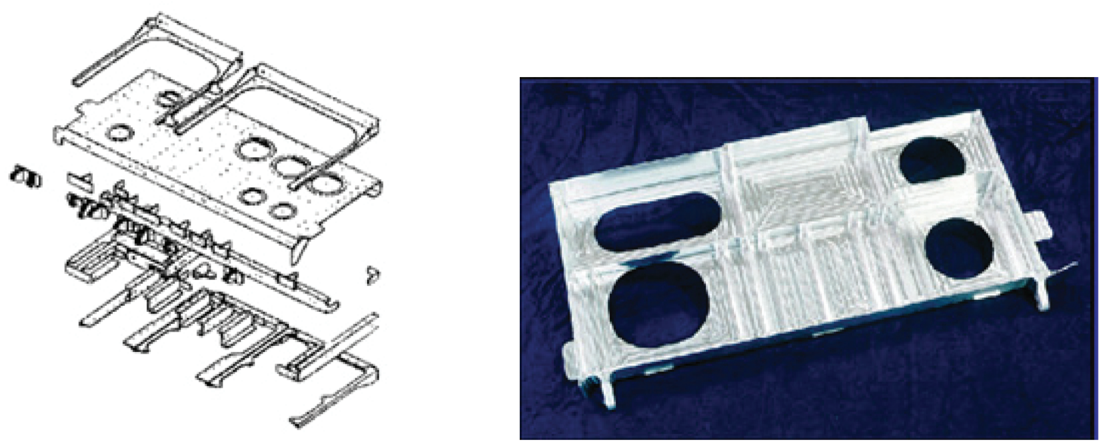
After each machining operation was completed, a small amount of hand finishing was necessary in order to produce a superior surface appearance. The putters were then lightly shot blasted (with glass bead media, Section 34.2) for the purpose of achieving surface consistency. A black, nickel–chrome plating (Section 34.9) was then applied to all parts to enhance aesthetic appeal and protect the stainless steel from small dings and dents and from corrosion from specific chemicals that might be encountered on a golf course.

Source: Courtesy of D. Jones and D. Petersen, Ping Golf, Inc.

Case Study 24.3 Machining of Aerospace Structures from Monolithic Extruded Aluminum

It is well known that the aircraft industry, and increasingly the automotive industry, places a premium on lightweight designs, using materials that reduce the weight of aircraft. It is also understood that aircraft components must be extremely reliable, as failure in service is potentially catastrophic.

Figure 24.39a shows a typical component, in this case an avionics tray for an F/A-18 fighter aircraft. It is produced from a number of sheet metal pieces that are then assembled using a variety of fasteners



(a) (b)

Figure 24.39: Production of avionics tray by (a) riveting of multiple stamped parts, and (b) machining from a monolithic block of aluminum.

to form the component. This part is approximately 1000 mm x 500 mm x 100 mm. While quite elaborate shapes can be produced using this strategy, the design has a number of shortcomings:

- Each of the multiple components requires its own set of tooling. The tooling cost must be amortized over limited production, so that each component has a very high cost.
- The tooling must be stored when not in use and readily available when needed.
- The sheet material is very reliable and has a repeatable strength. However, the use of assembly operations to join the sheets introduces a variety of failure modes. The fasteners pass through drilled holes which cause stress concentrations, and must be sealed.
- It may be advisable from a lightweight or stiff design standpoint to have varying section thicknesses, but sheet metal assemblies are most easily built up from preforms with a constant metal thickness.
- Assembled components have two layers of sheet metal, the weight of the fastener, and any shims and sealing material, so they are heavy.
- The assembly operations are almost always manual (low volume limits the options for automation).
- Tooling creates a tracking problem, and tooling has to be maintained, and can get lost.
- To enable assembly, very tight tolerances have to be maintained, especially between the ribs and the mounting holes. It is not unusual to require tolerances of 0.125 mm for 2.5 m long parts. This is very challenging, and assemblies often require match drilling; parts then need to be dissembled, deburred, and then reassembled.

An improved design strategy is shown in Fig. 24.39b, and demonstrates the approach of machining monolithic aluminum plate stock into functional components. To produce this component through machining, the following has to be noted:

- Aluminum plate stock contains residual stresses. To prevent distortion of the part during and after machining, the residual stresses are reduced before machining, either by stretching or compression (Section 2.11).
- A typical machining strategy of roughing to almost the finished dimension, and then finish machining does not work well for this kind of design. Thin parts are difficult to clamp (they deform under clamping forces), and they are difficult to machine (they deflect away from the tool or exhibit chatter). Instead, these parts are made using a "picture frame" strategy. The workpiece is the fixture, and the plate stock is simply clamped to the table. The part is machined layer by layer, using a relieved shank tool. This special tool prevents contact with the thin areas already machined. At each layer, the part is machined to the finished dimension, and then never machined again. The workpiece is cut where it is stiff, and the thin parts are left behind. As the part is almost finished, it is held in place with a few small tabs, which are removed last.
- While this kind of part can be made with a conventional machine, the machining time may be long. It is more common to use a high-speed spindle. Solid carbide tools can withstand the melting temperature of aluminum without exhibiting a high wear rate. With aluminum components, and using the stability lobes approach discussed in Section 25.4, chatter can be avoided at even at very high material removal rates.

The machined design is able to meet the stringent tolerances easily, and can achieve higher stiffness at lighter weight than the assembled component. This can be achieved, for example, by making high load sections thicker than the adjacent sections, eliminating the weight of the fasteners, and eliminating the weight of the double layers.

These parts are significantly less expensive, even though 90% or more of the purchased plate stock is converted into chips. The parts are less expensive because:

- The chips can be easily recycled.
- The special tooling is eliminated. Only the part program must be stored.
- The hand assembly operations are almost completely eliminated.
- A significant portion of the cost is in the weight (in the fuel required to carry the part in during its service life), and the monolithic parts are lighter for the same function.

A number of benefits were achieved with the machined components, as summarized in Table 24.4. Note the significant reduction in the number of parts in the structure and associated assembly time. The part cost was substantially reduced because of the ability to avoid tooling and assembly costs.

Machining of monolithic aluminum components has become standard for high performance components, and has disseminated into a wide variety of applications including commercial aerospace, automotive, and personal electronics. Many laptop computer housings, for example, are machined monolithic aluminum.

Source: Courtesy of K. Scott Smith, University of North Carolina at Charlotte.

Table 24.4: Comparison of Stamped and Riveted Assembly vs. Monolithic Machined Components Depicted in Fig. 24.39.

	Stamped and riveted	**Machined**	**Percent reduction**
Number of parts	44	6	84
Pan stock	445	108	76
Weight (kg)	4.34	3.88	4.98
Assembly time (hrs)	50	5.3	89

Summary

- A variety of complex shapes can be machined by the processes described in this chapter. Milling is one of the most common machining processes, because it is capable of economically producing a variety of shapes.
- Although these processes are similar to turning, drilling, and boring, and involve similar cutting mechanics, tool materials, and cutting fluids, most of the processes described utilize multi-tooth tools and cutters at various axes with respect to the workpiece.
- Machine tools used to produce complex shapes are now mostly computer controlled, having various dedicated features, and imparting much more flexibility in their application than traditional machine tools.

- Broaching is a method of accurately enlarging a round hole or other profile in a workpiece. Sawing is the gradual removal of material by small teeth spaced on a saw, and is very versatile. Filing involves small-scale removal of material from a surface, especially the removal of burrs and sharp profiles.
- In addition to being produced by various forming and shaping processes, gears also are produced by machining, either by form cutting or generating; the latter produces gears with better surface finish and higher dimensional accuracy. The quality of the gear-tooth profile is further improved by such process as shaving, burnishing, grinding, honing, and lapping.

Key Terms

Arbor
Broaching
Bur
Burnishing
Climb milling
Die sinking
End milling
Face milling
Filing
Fly cutting
Form cutting
Friction sawing
Gear generating
High-speed milling
Hob
Honing
Indexing
Kerf
Lapping
Milling
Planing
Pull broach
Push broach
Rack shaper
Sawing
Shaping
Shaving
Slab milling
Tooth set
Turn broaching
Workholding

Bibliography

Brown, J., **Advanced Machining Technology Handbook**, McGraw-Hill, 1998.
Davim, J.P. (ed.), **Machining: Fundamentals and Recent Advances**, Springer, 2010.
Davis, J.R., (ed.), **Gear Materials, Properties and Manufacture**, ASM International, 2006.
Joshi, P.H., **Machine Tools Handbook**, McGraw-Hill, 2008.
Krar, S.F., Gill, A.R., and Smid, P., **Technology of Machine Tools**, 7th ed., McGraw-Hill, 2010.
Lopez, L.N., and Lamikiz, A. (eds.), **Machine Tools for High Performance Machining**, Springer, 2009.
Machinery's Handbook, various editions, Industrial Press.
Radzevich, S.P., **Gear Cutting Tools: Science and Engineering**, 2nd ed., CRC Press, 2017.
—, Dudley's Handbook of Practical Gear Design and Manufacture, 3rd ed., CRC Press, 2016.
Stephenson, D.A., and Agapiou, J.S., **Metal Cutting: Theory and Practice**, 2nd ed., Marcel Dekker, 2006.
Walsh, R.A., **McGraw-Hill Machining and Metalworking Handbook**, 3rd ed., McGraw-Hill, 2006.

Review Questions

24.1. Explain why milling is such a versatile machining operation.

24.2. Describe a milling machine. How is it different from a drill press?

24.3. Describe the different types of cutters used in milling operations and give an application of each type.

24.4. Define the following: face milling, peripheral milling, shoulder milling, slot milling, thread milling.

24.5. Can threads be machined on a mill? Explain.

24.6. What is the difference between feed and feed per tooth? Can they ever be the same?

24.7. Explain the relative characteristics of climb milling and up milling.

24.8. Describe the geometric features of a broach and explain their functions.

24.9. What is a pull broach? A push broach?

24.10. Why is sawing a commonly used process? Why do some saw blades have staggered teeth? Explain.

24.11. What advantages do bed-type milling machines have over column-and-knee-type machines for production operations?

24.12. Explain why the axis of a hob is tilted with respect to the axis of the gear blank.

24.13. What is a shell mill? Why is it used?

24.14. Why is it difficult to saw thin sheet metals?

24.15. Of the processes depicted in Fig. 24.2, which is the most similar to hobbing?

24.16. Describe the tool motion during gear shaping.

24.17. When is filing necessary?

Qualitative Problems

24.18. Would you consider the machining processes described in this chapter to be near-net or net-shape processing? Explain with appropriate examples.

24.19. Why is end milling such an important versatile process? Explain with examples.

24.20. List and explain factors that contribute to poor surface finish in the processes described in this chapter.

24.21. Are the feed marks left on the workpiece by a face-milling cutter true segments of a true circle? Explain with appropriate sketches.

24.22. Explain why broaching crankshaft bearings is an attractive alternative to other machining processes.

24.23. Several guidelines are presented in this chapter for various cutting operations. Discuss the reasoning behind these guidelines.

24.24. What are the advantages of helical teeth over straight teeth on cutters for slab milling?

24.25. Explain why hacksaws are not as productive as band saws.

24.26. What similarities and differences are there in slitting with a milling cutter and with a saw?

24.27. Why do machined gears have to be subjected to finishing operations? Which of the finishing processes are not suitable for hardened gear teeth? Why?

24.28. How would you reduce the surface roughness shown in Fig. 24.8? Explain.

24.29. Why are machines such as the one shown in Fig. 24.20 so useful?

24.30. Comment on your observations concerning the designs illustrated in Fig. 24.23b and on the usefulness of broaching operations.

24.31. Explain how contour cutting could be started in a band saw, as shown in Fig. 24.28d.

24.32. In Fig. 24.30a, high-speed steel cutting teeth are welded to a steel blade. Would you recommend that the whole blade be made of high-speed steel? Explain your reasons.

24.33. Describe the parts and conditions under which broaching would be the preferred method of machining.

24.34. With appropriate sketches, explain the differences between and similarities among shaving, broaching, and turn-broaching operations.

24.35. Explain the reason that it is difficult to use friction sawing on nonferrous metals.

24.36. Would you recommend broaching a keyway on a gear blank before or after machining the gear teeth? Why?

Quantitative Problems

24.37. In milling operations, the total cutting time can be significantly influenced by (a) the magnitude of the noncutting distance, l_c, shown in Figs. 24.5 and 24.6, and (b) the ratio of width of cut, w, to the cutter diameter, D. Sketch several combinations of these parameters, give dimensions, select feeds and cutting speeds, etc., and determine the total cutting time. Comment on your observations.

24.38. A slab-milling operation is being performed at a specified cutting speed (surface speed of the cutter) and feed per tooth. Explain the procedure for determining the table speed required.

24.39. Show that the distance l_c in slab milling is approximately equal to $\sqrt{Dd}$ for situations where $D \gg d$ (see Fig. 24.5c).

24.40. In Example 24.1, which of the quantities will be affected when the feed is increased to $f = 0.75$ mm/tooth?

24.41. In Example 24.1, if the feed is increased to 0.75 mm/tooth, how are the chip depth of the cut and torque affected?

24.42. Estimate the time required to face mill a 250-mm-long, 25-mm-wide brass block with a 150-mm-diameter cutter with 10 high-speed steel inserts.

24.43. A 300-mm-long, 25-mm-thick plate is being cut on a band saw at 45 m/min. The saw has 0.5 teeth per 25 mm. If the feed per tooth is 0.075 mm, how long will it take to saw the plate along its length?

24.44. A single-thread hob is used to cut 40 teeth on a spur gear. The cutting speed is 35 m/min and the hob is 75 mm. in diameter. Calculate the rotational speed of the spur gear.

24.45. Assume that in the face-milling operation shown in Fig. 24.6 the workpiece dimensions are 100 mm by 250 mm. The cutter is 150 mm in diameter, has eight teeth, and rotates at 300 rpm. The depth of cut is 3 mm and the feed is 0.125 mm/tooth. Assume that the specific energy requirement for this material is and that only 75% of the cutter diameter is engaged during cutting. Calculate (a) the power required and (b) the material-removal rate.

24.46. A slab-milling operation will take place on a part 350 mm long and 45 mm wide. A helical cutter 75 mm in diameter with 10 teeth will be used. If the feed per tooth is 0.2 mm/tooth and the cutting speed is 0.8 m/s, find the machining time and metal-removal rate for removing 8 mm from the surface of the part.

24.47. Repeat Problem 24.46 if the same cutter diameter, number of inserts, and machining parameters are used, but the part is machined in a face milling operation.

24.48. A slab-milling operation is being carried out on a 0.75-m-long, 75-mm-wide high-strength-steel block at a feed of 0.3 mm/tooth and a depth of cut of 4 mm. The cutter has a diameter of 75 mm, has six straight cutting teeth, and rotates at 150 rpm. Calculate the material removal rate and the cutting time, and estimate the power required.

24.49. Explain whether the feed marks left on the workpiece by a face-milling cutter (as shown in Fig. 24.16a) are segments of true circles. Describe the parameters you consider in answering this question.

24.50. In describing the broaching operations and the design of broaches, equations regarding feeds, speeds, and material-removal rates have not been given as was been done in turning and milling operations. Review Fig. 24.24 and develop such equations.

Synthesis, Design, and Projects

24.51. The parts shown in Fig. 24.1 are to be machined from a rectangular blank. Suggest the machine tool(s) required, the fixtures needed, and the types and sequence of operations to be performed. Discuss your answer in terms of the workpiece material, such as aluminum versus stainless steel.

24.52. Would you prefer to machine the part in Fig. 24.1f from a preformed blank (near-net shape) rather than a rectangular blank? If so, how would you prepare such a blank? How would the number of parts required influence your answer?

24.53. If expanded honeycomb panels (see Section 16.13) were to be machined in a form-milling operation, what precautions would you take to keep the sheet metal from buckling due to tool forces? Think up as many solutions as you can.

24.54. Assume that you are an instructor covering the topics described in this chapter and you are giving a quiz on the numerical aspects to test the understanding of the students. Prepare two quantitative problems and supply the answers.

24.55. Suggest methods whereby milling cutters of various designs (including end mills) can incorporate carbide inserts.

24.56. Prepare a comprehensive table of the process capabilities of the machining processes described in this chapter. Using several columns, list the machines involved, types of tools and tool materials used, shapes of blanks and parts produced, typical maximum and minimum sizes, surface finish, dimensional tolerances, and production rates.

24.57. On the basis of the data developed in Problem 24.56, describe your thoughts regarding the procedure to be followed in determining what type of machine tool to select when machining a particular part.

24.58. Make a list of all the processes that can be used in manufacturing gears, including those described in Parts II and III of this text. For each process, describe the advantages, limitations, and quality of gears produced.

24.59. List the concerns you would have if you needed to economically pocket mill carbon-fiber-reinforced polymers or metal matrix composites with graphite fibers in an aluminum matrix.

Chapter 25

Machining Centers, Machine-tool Structures, and Machining

- This chapter presents the characteristics, types, and advantages of machining centers.
- Emphasis is placed on the importance of understanding the performance of machine tools and their modules and components, particularly with regard to their stiffness, vibration, chatter, and damping characteristics.
- Presented next are the topics of high-speed machining, hard machining, and ultraprecision machining operations, all strongly tied to the economics of machining.
- The chapter ends with a simple method of cost analysis for determining the conditions under which machining parameters can be selected, so that machining cost per piece or machining time per piece can be minimized.

25.1 Introduction

The preceding four chapters described machining operations and machine tools, but have not emphasized the integration of advanced computer technology and the flexibility it allows in manufacturing operations.

Computers have dramatically improved the capabilities of machine tools, whereby they now have the capability of rapidly and repeatedly producing very complex part geometries economically. The programs controlling a machine tool can incorporate changes in machining conditions, compensate for tool wear, automatically change tools, and machine a workpiece without refixturing or having to transfer it to another machine tool.

In addition to implementation of advanced computer technologies, vibration and chatter and their avoidance, high-speed machining, hard machining, and advanced analysis of machining economics are now highly developed and have revolutionized machining operations.

25.2 Machining Centers

In describing individual machining processes and machine tools in the preceding chapters, it was noted that each machine, regardless of how highly it is automated, is designed to perform basically the same specific operation, such as turning, boring, drilling, milling, broaching, planing, or shaping. It was also shown that most parts manufactured by the methods described throughout this book require additional operations on their various surfaces before they are completed.

Note, for example, that the parts shown in Fig. 25.1 have a variety of complex geometric features, and the surfaces on these parts require a different type of machining operation to meet a set of specific requirements concerning shapes, features, dimensional tolerances, and surface finish. Note also the following observations:

- Some possibilities exist in *net-shape* or *near-net shape* production of these parts, depending on specific constraints on shapes, dimensional tolerances, detailed surface features, surface finish, and various mechanical and other properties to meet service requirements. Shaping processes that are candidates for such parts are precision casting, powder metallurgy, powder-injection molding, and precision forging. Even then, however, it is very likely that the parts will still require additional finishing operations for small-diameter deep holes, threaded holes, flat surfaces for sealing with gaskets, parts with very close dimensional tolerances, sharp corners and edges, and flat or curved surfaces, with different surface-finish requirements.

Figure 25.1: Examples of parts that can be machined on machining centers, using processes such as turning, facing, milling, drilling, boring, reaming, and threading. Such parts ordinarily would require the use of a variety of machine tools. Forged motorcycle wheel, finish machined to tolerance and subsequently polished and coated. *Source:* Courtesy of R.C. Components.

- If some machining is required or if it is shown to be more economical to finish machine these parts to their final shapes, it is obvious that none of the machine tools described in Chapters 23 and 24 could *individually* and *completely* produce the parts. Note also that traditionally, machining operations are performed by moving the workpiece from one machine tool to another until all of the required operations are completed.

Machining Centers. The traditional method of machining parts by using different types of machine tools has been, and continues to be, a viable manufacturing method. This method can be highly automated to increase productivity, and in fact it is the principle behind **transfer lines**, also called *dedicated manufacturing lines* (DML) (see Section 37.2.4). Commonly used in *high-volume* or *mass production*, transfer lines consist of several specific (*dedicated*) machine tools, arranged in a logical and efficient sequence. The workpiece, such as an automotive engine block, is moved from one station to another, with a specific operation performed at each station; it is then transferred to the next station for another specific machining operations.

There are situations, however, where transfer lines are not feasible or economical, particularly when the types of products to be processed change rapidly due to factors such as product demand or modifications in product shape or style. It is very costly and time-consuming to rearrange these machine tools to respond to the changed needs for the next and different production cycle. An important concept that addresses *flexibility* in manufacturing, developed in the late 1950s, is that of **machining centers**.

A *machining center* (Fig. 25.2) is an advanced computer-controlled machine tool that is capable of performing a variety of machining operations on different surfaces and orientations of a workpiece, without having to remove it from its workholding device or fixture. The workpiece generally is stationary and the cutting tools rotate, as they do in such operations as milling, drilling, honing, and tapping. Whereas in

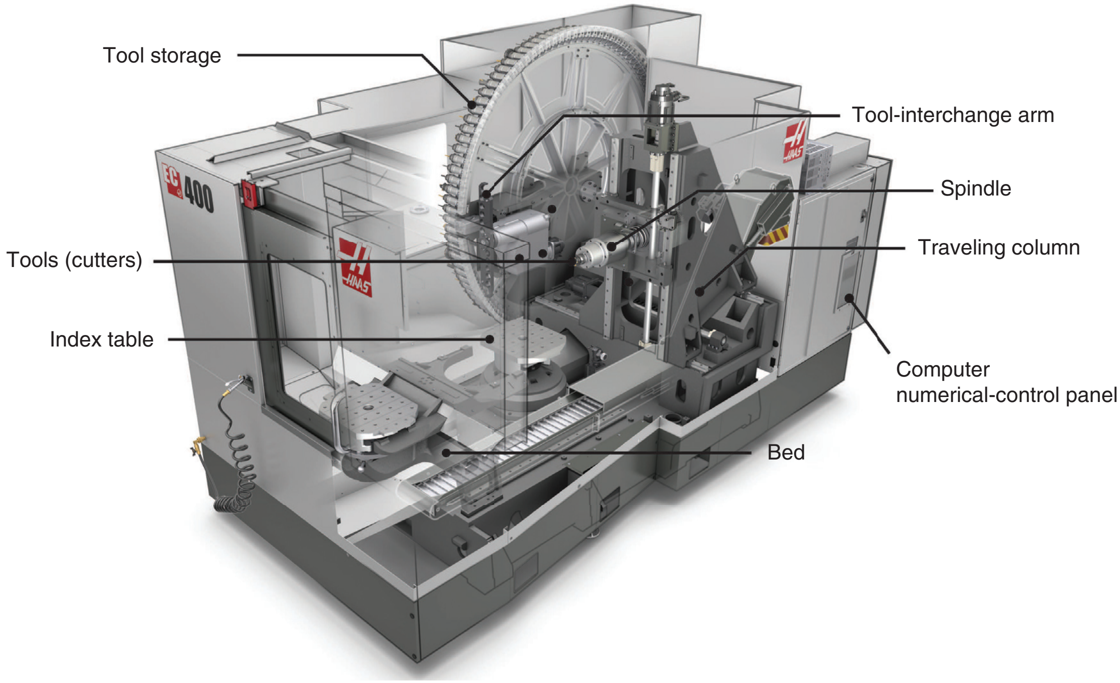


Figure 25.2: A horizontal-spindle machining center equipped with an automatic tool changer. Tool magazines can store up to 200 cutting tools of various functions and sizes. *Source:* Courtesy Haas Automation, Inc.

transfer lines or in traditional shops and factories the workpiece is brought *to the machine*, in machining centers, it is the machining operation that is brought *to the workpiece*.

The development of machining centers is related closely to advances in automation and computer control of machine tools, the details of which are described in Chapter 37. Recall that, as an example of the advances in modern lathes, Fig. 23.10 illustrates a numerically controlled lathe or *turning center*, with two turrets, each carrying several cutting tools.

Components of a Machining Center. The workpiece in a machining center is placed on a **pallet**, or *module*, which can be moved and swiveled (oriented) in various directions (Fig. 25.3). After a particular machining operation has been completed, another operation begins, which may require reindexing of the workpiece on its pallet. After all of the machining operations have been completed, the pallet automatically moves away with the finished part, and another pallet, carrying another workpiece or workpieces to be machined, is brought into position by an **automatic pallet changer** (Fig. 25.4). All movements are computer controlled, with pallet-changing cycle times on the order of only 10 to 30 seconds. Pallet stations are available with several pallets serving the one machining center. The machines also can be equipped with various automatic features, such as *part loading and unloading devices*.

A machining center is equipped with a programmable **automatic tool changer** (ATC). Depending on the particular design, up to 100 cutting tools can be stored in a magazine, drum, or chain (*tool storage*). *Auxiliary* tool storage also is available on some special and large machining centers, raising the tool capacity to 200. The cutting tools are selected automatically for the shortest route to the machine spindle. The maximum dimensions that the cutting tools can reach around a workpiece in a machining center is called the **work envelope**, a term that was first used in connection with industrial robots (Section 37.6).

The **tool-exchange arm** shown in Fig. 25.5 is a common design; it swings around to pick up a specific tool and places it in the spindle. Note that each tool has its own toolholder, thus making the transfer of cutting tools to the machine spindle highly efficient. Tools are identified by bar codes, QR codes, or coded tags attached directly to their toolholders. Tool-changing times are typically between 5 and 10 s, but may be up to 30 s for tools weighing up to 110 kg, and less than 1 s for small tools.

Machining centers may be equipped with a **tool-checking** and/or **part-checking station** that feeds information to the machine control system, so that it can compensate for any variations in tool settings or tool wear. **Touch probes** (Fig. 25.6) can be installed into a toolholder to determine workpiece reference surfaces, for selection of tool settings and for online inspection of parts being machined. Note in Fig. 25.6 that several surfaces can be contacted (see also *sensor technology*, Section 37.7), and that their relative positions

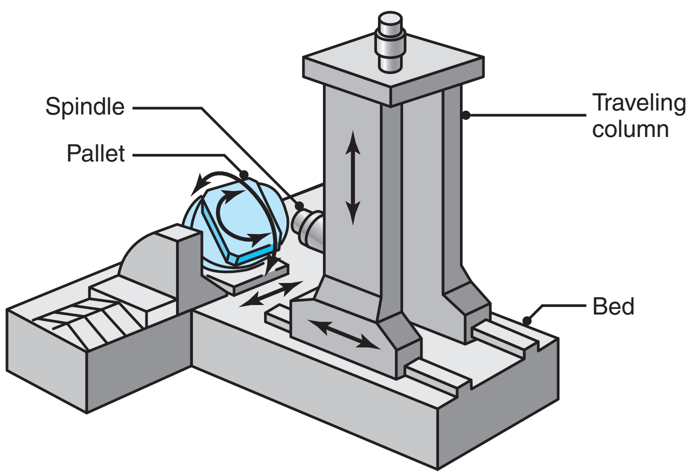


Figure 25.3: Schematic illustration of the principle of a five-axis machining center. The pallet, which supports and transfers the workpiece, has three axes of movement and can be swiveled around two axes (thus a total of five axes), allowing the machining of complex shapes, such as those shown in Fig. 25.1. *Source:* Courtesy of Toyoda Machinery.

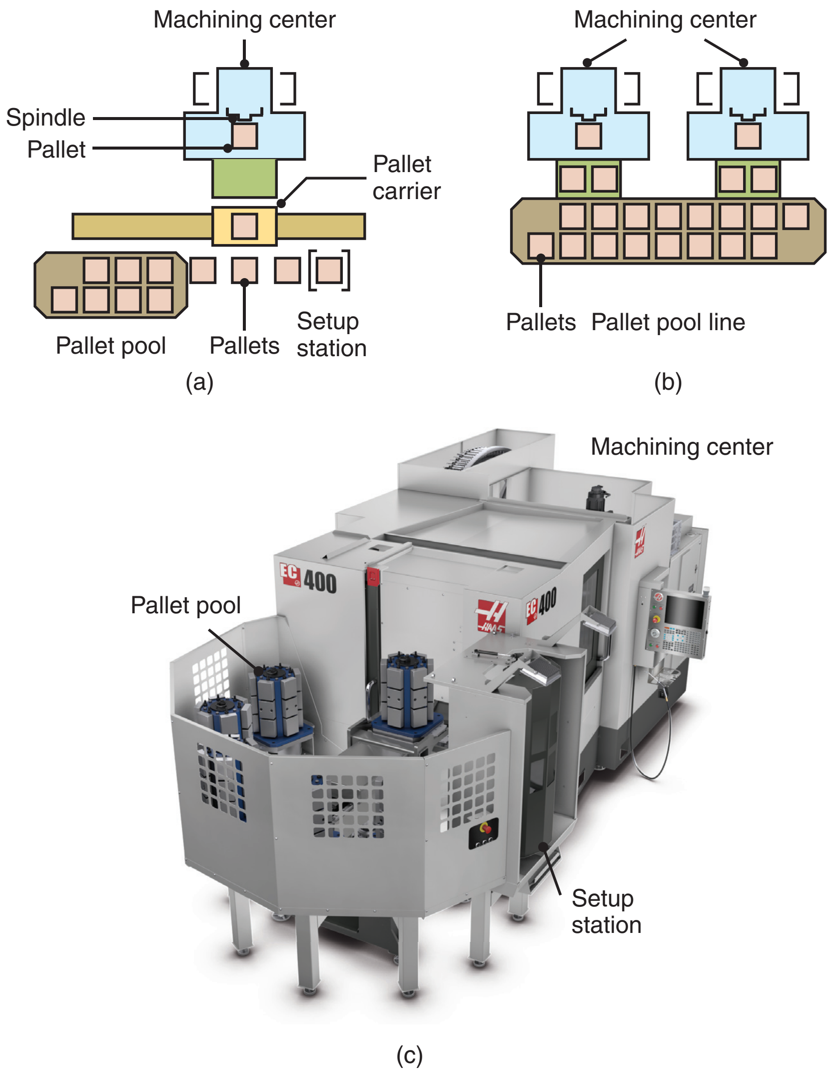


Figure 25.4: (a) Schematic illustration of the top view of a horizontal-spindle machining center, showing the pallet pool, setup station for a pallet, pallet carrier, and an active pallet in operation (shown directly below the spindle of the machine). (b) Schematic illustration of two machining centers, with a common pallet pool. (c) A pallet pool for a horizontal-spindle machining center. Various other pallet arrangements are possible in such systems. *Source:* (a) and (b) Courtesy of Hitachi Seiki Co., Ltd., (c) Courtesy of Haas Automation, Inc.

are determined and stored in the database of the computer software. In more advanced machine tools, the data are then used to program tool paths (see, for example, Fig. 37.12) and to compensate for tool length, tool diameter, and for tool wear. Non-contact probes also can be used, and can measure dimensions, surface roughness, or temperature.

25.2.1 Types of Machining Centers

There are various designs for machining centers. The two basic types are vertical spindle and horizontal spindle, although many machines are capable of operating along both axes.

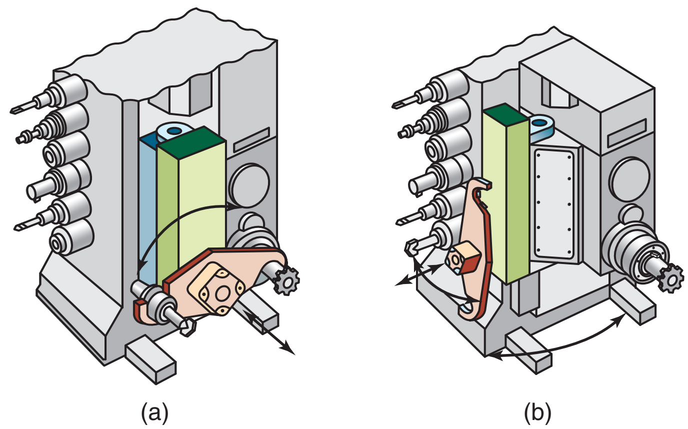


Figure 25.5: Swing-around tool changer on a horizontal-spindle machining center. (a) The tool-exchange arm is placing a toolholder, with a cutting tool, into the machine spindle. Note the axial and rotational movements of the arm. (b) The arm is returning to its home position. Note its rotation along a vertical axis after placing the tool, and the two degrees of freedom in its home position.

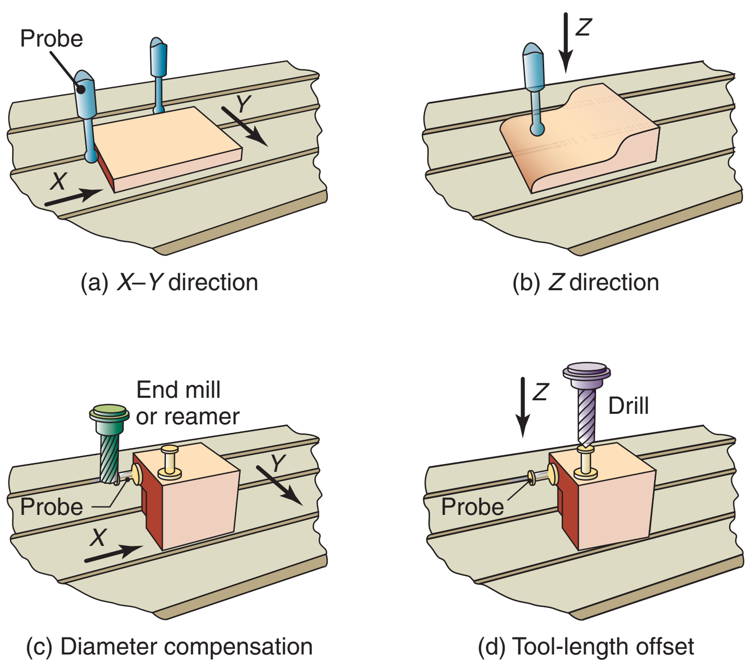


Figure 25.6: Touch probes used in machining centers for determining workpiece and tool positions and surfaces relative to the machine table or column. Touch probe (a) determining the x-y (horizontal) position of a workpiece, (b) determining the height of a horizontal surface, (c) determining the planar position of the surface of a cutter (*e.g.,* for cutter–diameter compensation), and (d) determining the length of a tool for tool-length offset.

Vertical-spindle Machining Centers. Also called *vertical machining centers* (VMC), these machines are capable of performing various machining operations on parts with deep cavities, as in mold and die making (also called *die sinking*). A vertical-spindle machining center, which is similar to a vertical-spindle milling machine, is shown in Fig. 25.7. The tool magazine is on the left of the machine, and all operations and movements are directed and modified through the computer control panel, shown on the right. Because

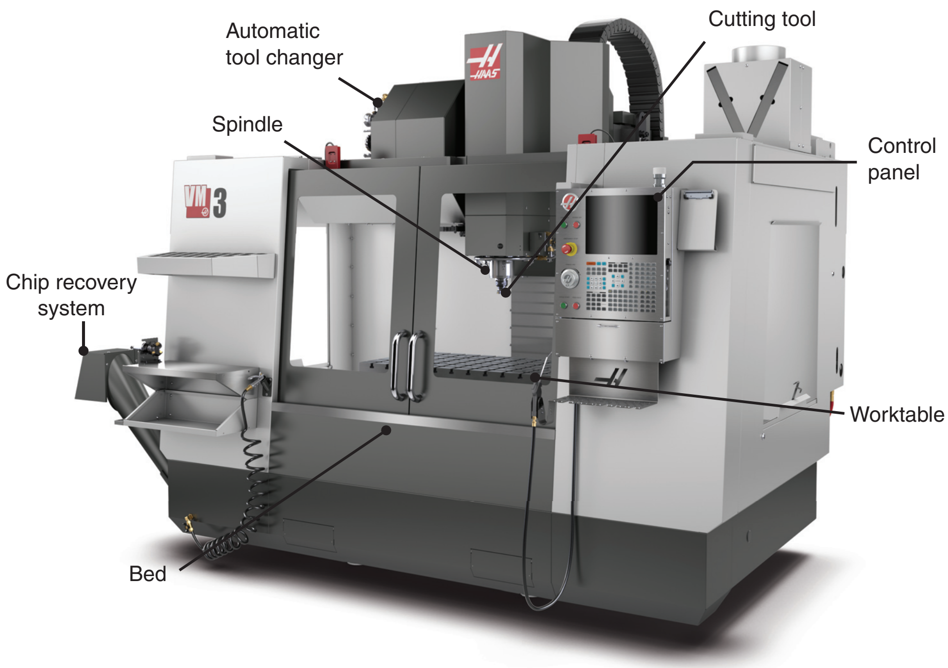


Figure 25.7: A vertical-spindle machining center. The tool changer is on the left of the machine, and has a 40 tool magazine. *Source:* Courtesy of Haas Automation, Inc.

the thrust forces in vertical machining are directed downward, such machines have high stiffness and hence able to produce parts with good dimensional accuracy. VMCs generally are less expensive than horizontal-spindle machines of similar capacity.

Horizontal-spindle Machining Centers. Also called *horizontal machining centers* (HMC), these machines are suitable for large as well as tall workpieces that require machining on a number of surfaces. The pallet can be swiveled on different axes to various angular positions (see Fig. 25.3).

Turning Centers. This is another category of horizontal-spindle machines; basically, they are computer-controlled lathes, with several features. A multi-spindle turning center is shown in Fig. 25.8. It is constructed with two horizontal spindles and two turrets, equipped with a variety of cutting tools used to perform several operations on a rotating workpiece. The turrets can be powered to allow for drilling or milling operations within the CNC turning center, and without the need to refixture the workpiece. For this reason, such machines are often referred to as **CNC Mill-turn Centers**.

Universal Machining Centers. These machines are equipped with both vertical and horizontal spindles. They have a variety of features and are capable of simultaneously machining all surfaces of a workpiece, i.e., vertically, horizontally, and at a wide range of angles.

25.2.2 Characteristics and Capabilities of Machining Centers

The major characteristics of machining centers are as follows:

- Machining centers are capable of efficiently handling a wide variety of part sizes and shapes, economically, repetitively, and with high dimensional accuracy and tolerances on the order of ± 0.0025 mm.
- These machines are versatile and capable of quick changeover from one type of product to another.

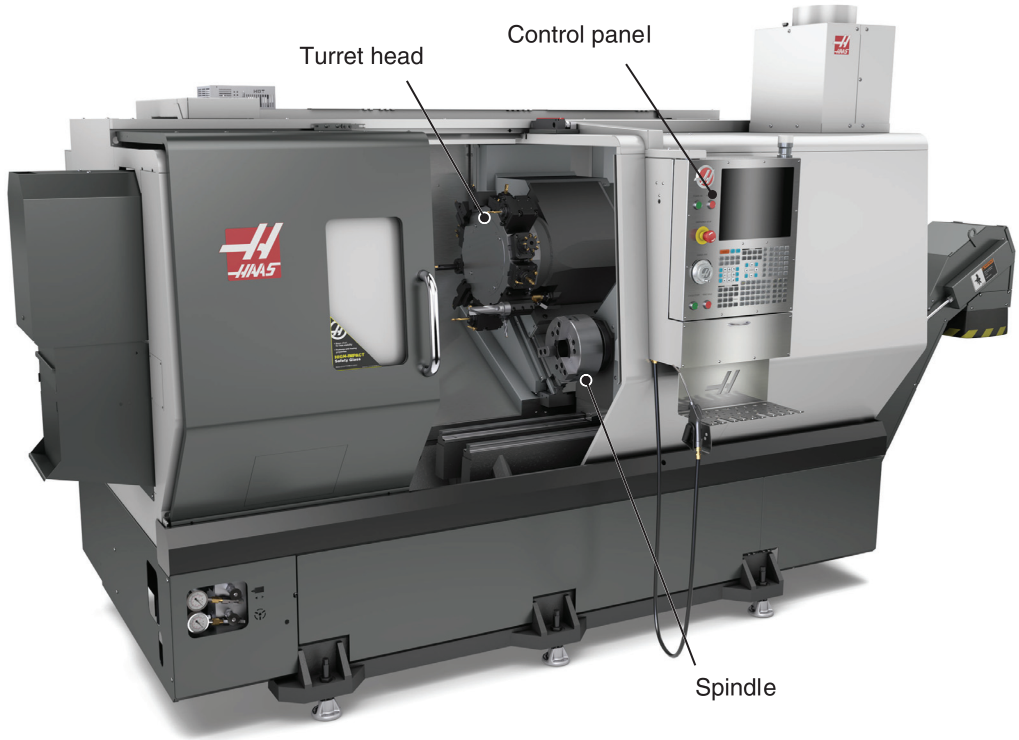


Figure 25.8: The Haas DS-30Y, a computer numerical-controlled turning center. The two spindle heads and tool-driven turret head make the machine very flexible in its machining capabilities. Source: Courtesy of Haas Automation, Inc

- The time required for loading and unloading workpieces, changing tools, gaging of the part being machined, and troubleshooting is reduced. Because of the inherent flexibility in machining centers, the workpiece may not have to be refixtured during machining, referred to as the *one and done* approach. Productivity is improved, labor requirements (particularly skilled labor) are reduced, and production costs are minimized.

- These machines can be equipped with tool-condition monitoring devices for the detection of tool breakage and wear, as well as with probes for tool-wear compensation and tool positioning.

- In-process and postprocess gaging and inspection of machined workpieces are now features of machining centers.

- These machines are highly automated and relatively compact, and have advanced control systems; one operator can attend to two or more machining centers at the same time, thus reducing labor costs.

Because of the high productivity of machining centers, large amounts of chips are produced and must be collected and disposed of properly (see *chip management*, Section 23.3.7). Several system designs are available for *chip collection*, with one or more chain or spiral (screw) conveyors; they collect the chips along troughs in the machine and deliver them to a collecting point (Fig. 25.7).

Machining centers are available in a wide variety of sizes and features; typical capacities range up to 75 kW. Maximum spindle speeds are usually in the range of 4000 to 8000 rpm, and some are as high as 75,000 rpm for special applications, using small-diameter cutters. Modern spindles can accelerate to a speed of 20,000 rpm in 1.5 s. Some pallets are capable of supporting workpieces weighing as much as 7000 kg, although even higher capacities are available for special applications. The cost of machining centers ranges from about $50,000 to $2 million and higher.

25.2.3 Selection of Machining Centers

Machining centers generally require significant capital expenditure; to be cost effective, they may have to be operated for more than one shift per day. Consequently, there must be a sufficient and continued demand for parts to justify their purchase. Because of their inherent versatility, however, machining centers can be used to produce a wide range of products, particularly for *mass customization* or *just-in-time manufacturing* (Section 39.6).

The selection of the type and size of machining centers depends on several factors, especially the following:

- Type of products, their size, and shape complexity
- Type of machining operations to be performed and the type and number of cutting tools required
- Dimensional accuracy specified
- Production rate required.

Case Study 25.1 Machining Outer Bearing Races on a Turning Center

Outer bearing races (Fig. 25.9) are machined on a turning center. The starting material is a hot-rolled 52100 steel tube, with 91 mm OD and 75.5 mm ID. The cutting speed is 95 m/min for all operations. All tools are carbide, including the cutoff tool (used in the last operation shown), which is 3.18 mm, instead of 4.76 mm for the high-speed steel cutoff tool that formerly was used.

The amount of material saved by this change is significant, because the race width is small. The turning center was able to machine these races with repeatable tolerances of ± 0.025 mm and at high speeds (see also Example 23.2).

Source: Courtesy of McGill Manufacturing Company.

25.2.4 Reconfigurable Machines and Systems

The need for the flexibility of manufacturing operations has led to the concept of *reconfigurable machines*, consisting of various modules. The term reconfigurable stems from the fact that, by using advanced computer hardware and reconfigurable controllers, and utilizing advances in information management technologies, the machine components can be arranged and rearranged into a number of configurations to meet specific production demands.

Fig. 25.10 shows an example of how the basic machine-tool structure of a three-axis machining center can be reconfigured to become a **modular machining center**. With such flexibility, the machine can perform different operations while accommodating various workpiece sizes and part geometries. Another example is given in Fig. 25.11, where a five-axis (three linear and two rotational) machine can be reconfigured by assembling different modules. Reconfigurable machines have the aim of (a) improving the productivity and efficiency of manufacturing operations, (b) reducing lead time for production, and (c) providing a cost-effective and rapid response to market demands (see also Chapter 39).

25.3 Machine-tool Structures

This section describes the materials and design aspects of machine-tool structures that are important in producing parts, with acceptable geometric features and dimensional and surface finish characteristics.

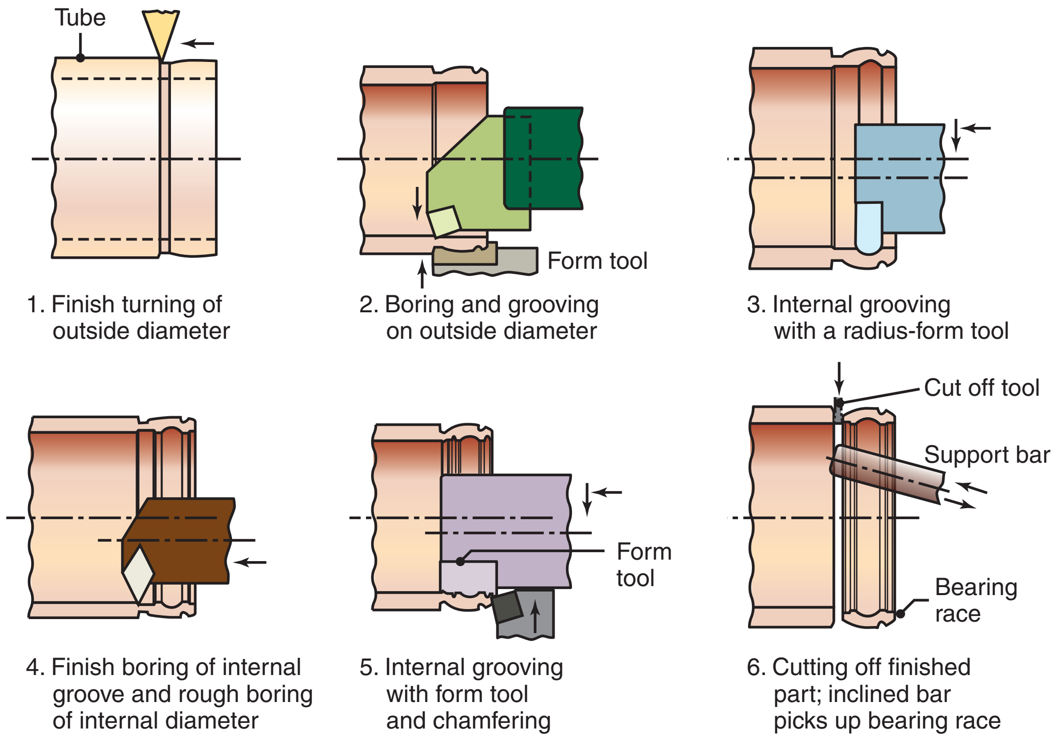


Figure 25.9: Machining of outer bearing races.

25.3.1 Materials

The following is a list of the materials that commonly have been used for machine-tool structures.

- **Gray cast iron** is the first material used in machine tool structures, and has the advantages of a good damping capacity and low cost, but has the limitation of being heavy. Most machine-tool structures are made of class 40 cast iron; some are made of class 50 (see Table 12.4).

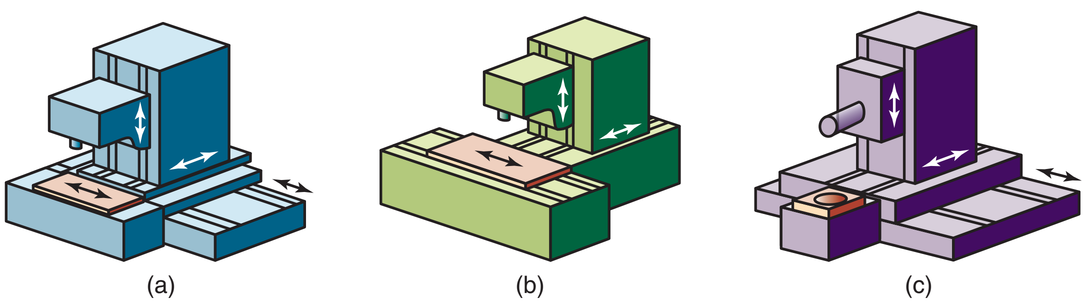


Figure 25.10: Schematic illustration of a reconfigurable modular machining center capable of accommodating workpieces of different shapes and sizes and requiring different machining operations on their various surfaces. *Source:* After Y. Koren.

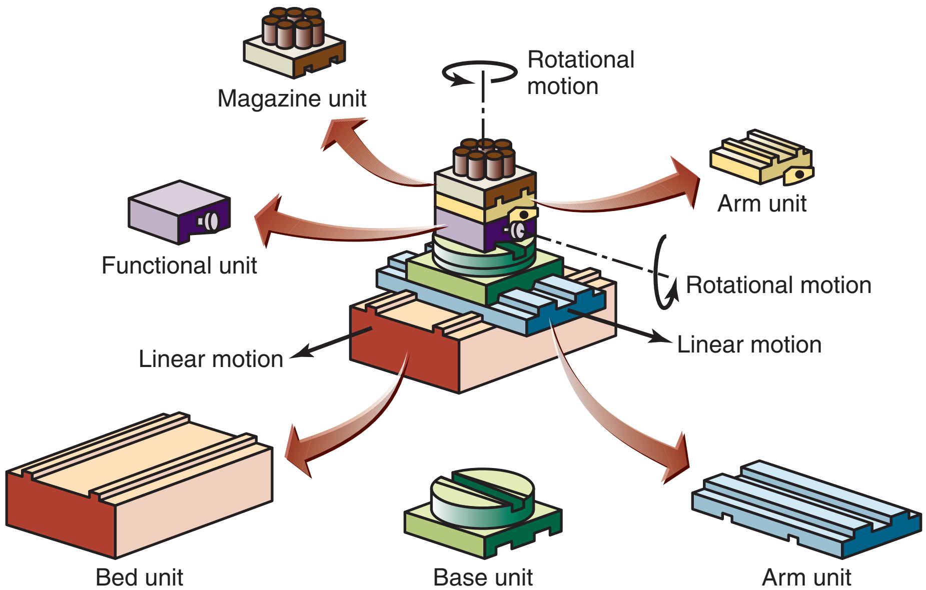


Figure 25.11: Schematic illustration of the assembly of different components of a reconfigurable machining center. *Source:* After Y. Koren.

- **Welded steel** structures (Chapters 30 and 31) are lighter than cast-iron structures. Wrought steels, typically used in these structures, (a) are available in a wide range of section sizes and shapes, such as channels, angles, and tubes, (b) have good mechanical properties such as strength and stiffness, (c) possess good formability, machinability, and weldability, and (d) have low cost. Structures made of steels can have high stiffness-to-weight ratios, by using various cross sections such as tubes and channels; in contrast, however, their damping capacity is very low.

- **Ceramic** components (Chapters 8 and 18), introduced in the 1980s, are used in advanced machine tools for their strength, stiffness, corrosion resistance, surface finish, and thermal stability. Spindles and bearings can be made of silicon nitride, which has better friction and wear characteristics than traditional metallic materials. Furthermore, their low density makes them suitable as components for high-speed machinery that undergo rapid reciprocating or rotating movements. Low inertial forces are desirable to maintain the system's stability, reduce inertial forces, and reduce the noncutting time in high-speed machining operations.

- **Composites** (Chapter 9) may consist of a polymer matrix, metal matrix, or ceramic matrix with various reinforcing materials. Their compositions can easily be tailored to provide appropriate mechanical properties in selected axes of the machine tool. Although they are presently expensive, composites are likely to be important materials for high-accuracy, high-speed machining applications.

- **Granite–epoxy composites**, with a typical composition of 93% crushed granite and 7% epoxy binder, were first used in the early 1980s in precision centerless and internal grinders (Section 26.4). These composite materials have several favorable properties: (a) good castability, thus allowing for design versatility in machine tools, (b) high stiffness-to-weight ratios, (c) thermal stability, (d) good damping capacity, and (e) resistance to environmental degradation.

- **Polymer concrete** is a mixture of crushed concrete and plastic (typically polymethylmethacrylate), and can easily be cast into desired shapes for machine bases and various components. Although it has low stiffness (about one-third that of class 40 cast iron) and poor thermal conductivity, polymer concrete has good damping capacity and can also be used for sandwich construction together with cast irons, thus combining the advantages of each type of material. Plain concrete can be poured *into* cast-iron machine-tool structures, to increase their mass and improve their damping capacity. Filling the cavities of machine bases with *loose sand* also has been demonstrated to also be an effective means of improving damping capacity.

25.3.2 Machine-tool Design Considerations

Important considerations in machine tools generally involve the following factors:

- Design, materials, and construction
- Spindle materials and type of construction
- Thermal distortion of machine components
- Error compensation and the control of moving components along slideways.

Stiffness. Stiffness, a major factor in the dimensional accuracy and vibration of a machine tool, is a function of (a) the elastic modulus of the materials used and (b) the geometry of the structural components, including the spindle, bearings, drive train, and slideways. Machine tool stiffness can be enhanced by such design improvement as using diagonally arranged interior ribs.

Damping. Damping is a critical factor in reducing or eliminating vibration and chatter in machining operations. Principally, it involves (a) the types of materials used and (b) the type and number of joints (such as bolted vs. welded) in the structure of the machine tool. For example, cast irons and polymer-matrix composites have much better damping capacity than metals or ceramics; also, the greater the number of joints in a machine structure, the more is the damping.

Thermal Distortion. An important factor in machine tools is the thermal distortion of their components, which contributes significantly to lack of precision. There are two sources of heat in machine tools:

1. *External sources*, such as from cutting fluids, nearby furnaces, heaters, other nearby machines, sunlight, and fluctuations in ambient temperature (from sources such as air-conditioning units, vents, or even someone opening or closing a door or a window).

2. *Internal sources*, such as from bearings, ballscrews, machine ways, spindle motors, pumps, and servomotors, as well as from the cutting zone during machining (Section 21.4).

These considerations are significant, particularly in **precision** and **ultraprecision machining** (Section 25.7), where dimensional tolerances and surface finish are now at nanometer range. The machine tool used for these applications are equipped with the following features:

- Various thermal and geometric real-time error-compensating features, including (a) the modeling of heating and cooling and (b) electronic compensation for accurate ballscrew positions.
- Gas or fluid hydrostatic spindle bearings, allowing tools to more easily achieve precise motions without encountering high friction or stick-slip phenomena (Section 33.4).
- New designs for traction or friction drives, for smoother linear motion.
- Extremely fine feed and position controls, using microactuators.
- Fluid-circulation channels in the machine-tool base, for maintaining thermal stability.

The structural components of the machine tool can be made of materials with high dimensional stability and low coefficient of thermal expansion, such as Super-Invar (Section 3.6), granite, ceramics, and composites. *Retrofitting* also is a viable option for enhancing the performance of older machines.

Assembly Techniques for Machine-tool Components. Traditionally, machine-tool components have been assembled using threaded fasteners and by welding (Part VI). Advanced assembly techniques now include integral casting and resin bonding. Steel guideways, with their higher stiffness, can be cast integrally over a cast-iron bed, using a hybrid casting technology. *Resin bonding* is being used to assemble machine tools, replacing mechanical fastening. Adhesives (Section 32.4) have favorable characteristics for machine-tool construction, as they do not require special preparation and are suitable for assembling both nonmetallic and metallic machine components.

Guideways. The preparation of guideways in machine tools traditionally has required significant effort. The plain cast-iron ways in machines, the most common material, require much care to achieve the required precision and service life. The movements of various components in a machine tool along its various axes typically have utilized high precision *ballscrews*, *rotating-screw drives*, and *rotary motors*. This system of mechanical and electrical components has several unavoidable design characteristics: speed limitations, length restrictions, inertial effects, gear backlash, and wear of components.

Linear Motor Drives. A *linear motor* is like a typical rotary electric motor that has been rolled out (opened) flat. This is the same principle used in some high-speed ground transportation systems in which the cars are *levitated by magnetic forces* (Maglev). The sliding surfaces in these drives are separated by an *air gap*; as a result, they have very low friction and energy loss.

Linear motor drives in machine tools have important advantages:

- Design simplicity and minimal maintenance, since there is only one moving part and no mechanical linkages
- Smooth operation, better positioning accuracy, and repeatability, at a submicron range
- A wide range of linear speeds, from 1 μm/s to 5 m/s
- Acceleration rates of about 1 to 2 g (10 to 20 m/s^2), and as high as 4 to 10 g for smaller units
- Because there is no physical contact between the sliding surfaces of the machine, the moving components do not undergo any wear.

Machine Foundations. Foundation materials, their mass, and the manner in which they are installed in a plant are major considerations, as they help reduce vibration and do not adversely affect the performance of nearby machinery in the plant. For example, in the installation of a special grinder for high-precision grinding of 2.75-m diameter marine-propulsion gears, the concrete foundation was 6.7 m deep. Its large mass, combined with the machine base, reduced the amplitude of vibrations. Even better results can be obtained when a machine is installed on an *independent* concrete slab, isolated from the rest of the plant floor with shock-isolation devices.

25.3.3 Hexapod Machines

Developments in the design and materials for machine-tool structures and their various components continue to take place, with the purposes of (a) imparting machining flexibility to machine tools, (b) increasing their *machining envelope* (the space within which machining can be done), and (c) making them lighter.

An example of a unique machine-tool structure is a self-contained octahedral (eight-sided) machine frame. Referred to as **hexapods** (Fig. 25.12) or *parallel kinematic linked machines*, these machines have a design that is based on a mechanism called the *Stewart platform* (after D. Stewart); it was first developed in 1966

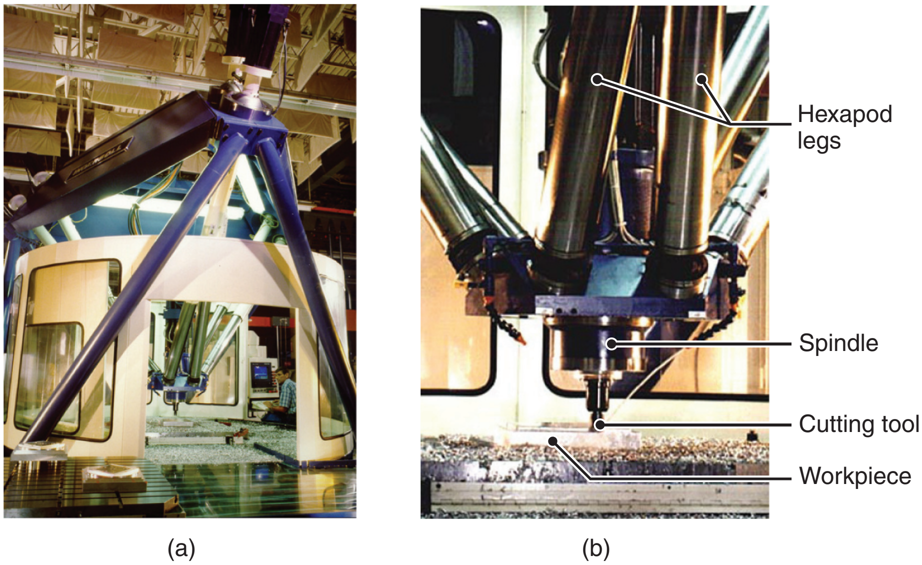


Figure 25.12: (a) A hexapod machine tool, showing its major components. (b) A detailed view of the cutting tool in a hexapod machining center. *Source:* National Institute of Standards and Technology.

and used to position aircraft cockpit simulators. The main advantage of this system is that the links in the hexapod are loaded axially, thus the bending forces and lateral deflections are minimal, resulting in a very stiff structure.

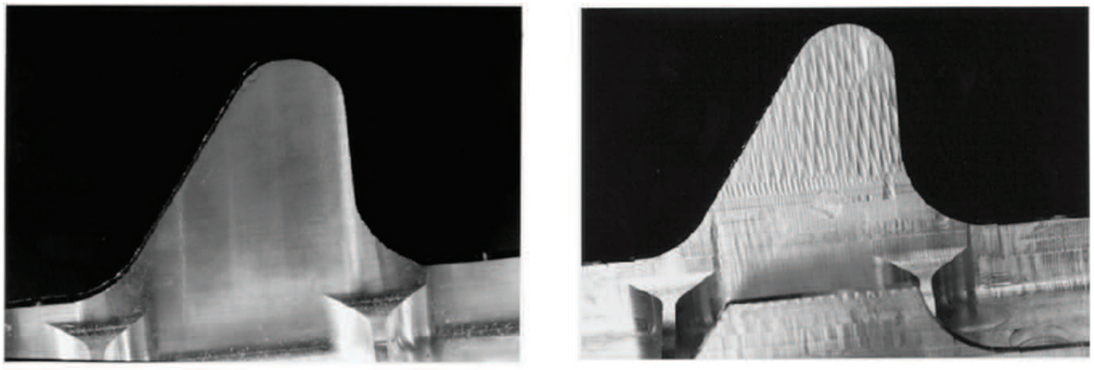
The workpiece is mounted on a stationary table. Three pairs of *telescoping tubes* (called *struts* or *legs*), each with its own motor and equipped with ballscrews, are used to maneuver a rotating cutting-tool holder. While various features and curved surfaces are being machined, the controller automatically shortens some tubes and extends others, so that the cutter can follow a specified path around the workpiece. Six sets of coordinates are involved in these machines (hence the term *hexapod*, meaning six legged): three linear sets and three rotational sets. Every motion of the cutter, even a simple linear motion, is translated into six coordinated leg lengths moving in real time. The motions of the legs are rapid; consequently, high accelerations and decelerations are involved, resulting in high inertial forces.

The machines (a) have high stiffness; (b) are not as massive as machining centers; (c) have about one-third fewer parts than machining centers; (d) have a large machining envelope (thus greater access to the work zone); (e) are capable of maintaining the cutting tool perpendicular to the surface being machined; and (f) with six degrees of freedom, they have high flexibility in the production of parts with various geometries and sizes, without the need for refixturing the work in progress. Unlike most machine tools, they are basically portable. With *hexapod attachments*, a conventional machining center can easily be converted into a hexapod machine.

A limited number of hexapod machines have been built. In view of their potential as efficient machine tools, their performance is being evaluated continually regarding stiffness, thermal distortion, friction within the struts, dimensional accuracy, speed of operation, repeatability, and reliability.

25.4 Vibration and Chatter in Machining Operations

In describing machining processes and machine tools, it was noted on several occasions that *machine stiffness* is one of the most important parameters in machining. Low stiffness can cause *vibration* and *chatter*, and thus have adverse effects on product quality. Uncontrolled vibration and chatter can result in:



(a) (b)

Figure 25.13: Chatter marks in machining. (a) smooth surface, created using the stability lobe approach to avoid chatter; (b) milling chatter marks on the surface of an Al 7075 component. Source: Courtesy Scott Smith, University of North Carolina at Charlotte.

- Poor surface finish, as shown in the right central region of Fig. 25.13
- Loss of dimensional accuracy of the workpiece
- Premature wear, chipping, and failure of the cutting tool, a critical consideration with brittle tool materials, such as ceramics, some carbides, and diamond
- Possible damage to the machine-tool components, from excessive vibration
- Objectionable noise, particularly if it is of high frequency, such as the squeal heard when turning brass on a lathe.

There are two basic types of vibration in machining: forced and self-excited.

Forced Vibration. Forced vibration is generally caused by a *periodic* applied force that develops in the machine tool, such as from gear drives, imbalance of the machine-tool components, misalignment, and motors and pumps. In operations such as milling or turning of a splined shaft, or a shaft with a keyway or a radial hole, forced vibrations are caused by the periodic engagement of the cutting tool with the workpiece (see, for example, Fig. 24.9).

The basic solution to forced vibration is to *isolate* or *remove* the forcing element. If, for example, the forcing frequency is at or near the *natural frequency* of a machine-tool component, one of these two frequencies may be raised or lowered. The *amplitude* of vibration can be reduced by increasing the stiffness or by damping the system.

The cutting parameters generally do not appear to greatly influence the magnitude of forced vibrations; however, changing the cutting speed and the tool geometry can be helpful. It is also recognized that the source of vibrations also can be minimized by changing the configuration of the machine-tool components, as may be done when the driving forces are close to, or act through, the *center of gravity* of a particular component. This approach will reduce the bending moment on the component, thus reducing deflections and improving dimensional accuracy.

Self-excited Vibration. Generally called **chatter**, self-excited vibration is caused by the interaction of the machining process with the structure of the machine tool. The vibrations usually have very high amplitude, and are audible. Chatter typically begins with a disturbance in the cutting zone, such as by (a) the type of chips produced, (b) inhomogeneities in the workpiece material or its surface condition, and (c) variations in the frictional conditions at the tool–chip interface, as influenced by cutting fluids and their effectiveness.

The most important type of self-excited vibration is **regenerative chatter**, which is caused when a tool is cutting a surface that has a roughness or geometric disturbances developed from the *previous* cut (see Figs. 21.2 and 21.23). The depth of cut varies periodically, and the resulting variations in the cutting force subject the tool to vibrations. The process continues repeatedly, hence the term *regenerative*. This type of vibration can easily be observed while driving a car over a rough road, the so-called *washboard effect*.

Self-excited vibrations generally can be controlled by:

- Increasing the *stiffness*, especially the *dynamic stiffness*, of the system. The system includes not only the tool, tool holder, machine frame, etc., but also the *workpiece* and how it is supported on the machine.
- *Damping* the system.

Dynamic stiffness is defined as the ratio of the applied-force amplitude to the vibration amplitude. For example, recall that in a trepanning operation (Fig. 23.24b), there are four machine components involved in the deflections that would cause vibrations: (a) spindle, (b) supporting arm for the cutting tool, (c) drill, and (d) cutting tool. Analysis and experience would suggest that, unless all of these machine components are sufficiently stiff, the trepanning operation will likely lead to chatter, beginning with the torsional vibration around the spindle axis and the twisting of the arm. Two similar examples are (a) long and slender drills, which may undergo torsional vibrations, and (b) cutting tools that are long or are not well supported, such as that shown schematically in Fig. 23.3.

Factors Influencing Chatter. It has been observed that the tendency for chatter during machining is proportional to the cutting forces and the depth and width of the cut. Because the forces increase with strength (hence with hardness of the workpiece material), the tendency to chatter generally increases as hardness increases. Aluminum and magnesium alloys, for example, have a lower tendency to chatter than do martensitic and precipitation-hardening stainless steels, nickel alloys, and high-temperature and refractory alloys.

Another important factor in chatter is the type of chip produced during machining. Continuous chips involve fairly steady cutting forces; such chips generally do not cause chatter. On the other hand, discontinuous chips and serrated chips (Fig. 21.5) may do so. These type of chips are produced periodically, and the resulting force variations during machining can thus cause chatter. Other factors that may contribute to chatter are using dull tools or cutters, lack of cutting fluids, and worn machine-tool ways and components.

Damping. *Damping* is defined as the rate at which vibrations decay. This effect can be demonstrated on an automobile's shock absorbers, by pushing down on the car's front or rear end and observing how rapidly the vertical motion stops. Damping is a major factor in controlling machine-tool vibration and chatter; it consists of internal and external damping.

1. **Internal damping** results from the *energy loss* in materials during vibration. For example, composite materials have a higher damping capacity than gray cast iron (Fig. 25.14). The difference in the damping capacity of materials can easily be observed by striking them with a gavel and listening to the sound emitted. Try striking first a brass cymbal, then a piece of concrete, and then a piece of wood, and listen to the distinct variations in their sound.

2. *Bolted joints* in the structure of a machine tool also are a source of damping, their effectiveness depending on size, position, and the number of joints. Because friction dissipates energy, small relative movements along dry (unlubricated) joints increase damping. Because machine tools consist of a number of large and small components, assembled by various means, this type of damping is *cumulative*. Note in Fig. 25.15, for example, how overall damping increases as the number of components on a lathe and their contact areas increase. However, the overall stiffness of the machine tool will decrease as the number of joints increases. As described and illustrated in Fig. 23.18b, damping also can be accomplished by mechanical means, whereby energy is dissipated by the *frictional* resistance of the components within the structure of the boring bar.

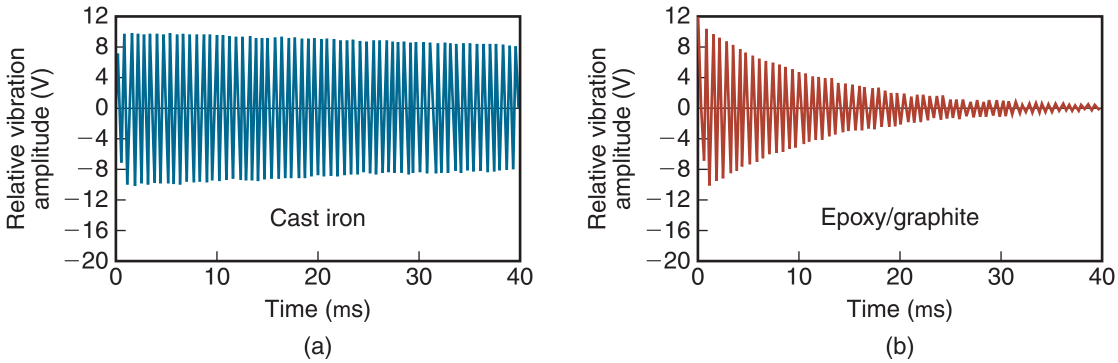


Figure 25.14: The relative damping capacity of (a) gray cast iron and (b) an epoxy–granite composite material. The vertical scale is the amplitude of vibration and the horizontal scale is time.

3. **External damping** is accomplished with external dampers, similar to shock absorbers on automobiles or machinery. Special vibration absorbers have been developed and installed on machine tools for this purpose. Also, the machines can be installed on specially prepared floors and foundations to isolate forced vibrations, such as those from nearby machinery on the same floor.

Stability Lobes. Production demands now require the selection of process parameters that result in high material-removal rates without the risk of chatter. In recent years, *stability lobes* have been studied, an example of which is shown in Fig. 25.16. It has been noted that, for a given spindle in milling, chatter occurs at certain combinations of speed and tooth depth of cut. Avoiding these combinations results in chatter-free milling.

Consider the *tooth passing frequency*, defines as the product of spindle speed and the number of teeth. The peak stable axial depths (green lines in Fig. 25.16, marked as *stability lobes*) occur at spindle speeds where the natural frequency of the machine is an integer multiple of the tooth passing frequency. The machine natural frequency can be either measured audibly or with sensors on the machine tool. Applying this concept allows the selection of processing parameters with very high speeds and material-removal rate, without a risk of chatter.

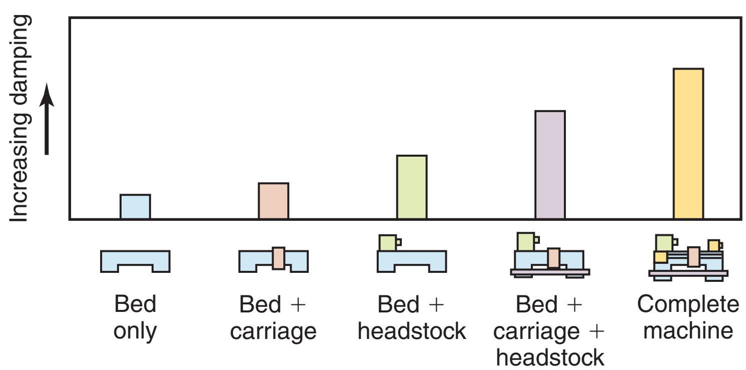


Figure 25.15: The damping of vibrations as a function of the number of components on a lathe. Joints dissipate energy; the greater the number of joints, the higher is the damping capacity of the machine. *Source:* After J. Peters.

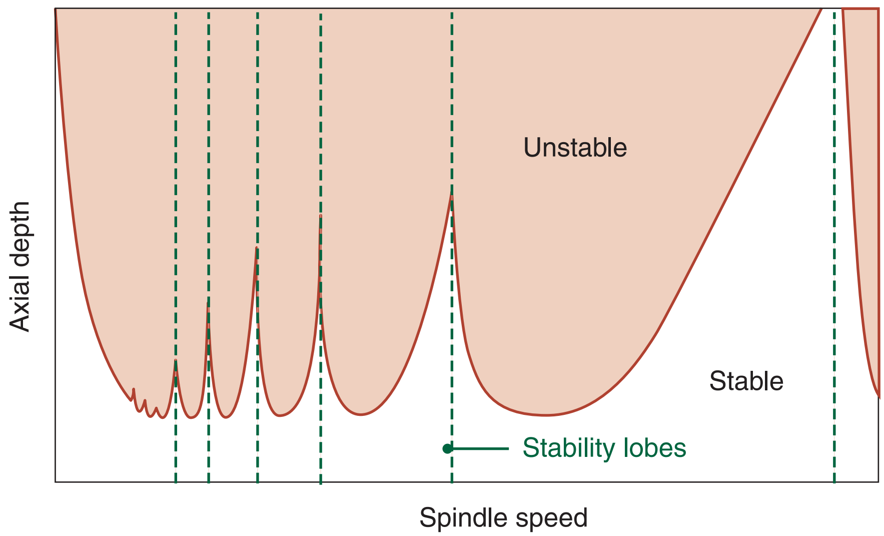


Figure 25.16: The milling stability lobe diagram, or stability map, showing the spindle speed against the limiting axial depth of cut to avoid chatter. Stable and unstable combinations are identified by the white and shaded regions. *Source:* After T. Schmitz, University of North Carolina at Charlotte.

Guidelines for Reducing Vibration and Chatter. Chatter can be eliminated by maintaining operations in a stability lobe. To make lobes larger and allow greater material removal rates, a balance must be achieved between the increased stiffness of a machine tool and the desirability of increased damping, particularly in the construction of high-precision machine tools.

Several guidelines from Chapters 23 and 24 that were given for reducing vibration and chatter in machining operations can be applied. These basic guidelines may be summarized as :

- Minimize tool overhang
- Improve the stiffness of work-holding devices and support workpieces rigidly
- Modify tool and cutter geometry to minimize forces; specifically, changing the number of flutes is often advisable
- Change process parameters, such as feed, depth of cut, and cutting fluids
- Select a cutting speed so that a stable lobe is utilized
- Increase the stiffness of the machine tool and its components, by improving their design, using larger cross sections and materials with a higher elastic modulus
- Adjust the machining operation, to shift the cutting forces into the stiffer portions of the machine tool
- Improve the damping capacity of the machine tool.

These can be seen as adjusting either the critical frequency of the machine or the tooth passing frequency, thereby ensuring machining takes place within a stable lobe.

25.5 High-speed Machining

With continuing demands for higher productivity and lower production costs, a continuing desire is increased cutting speed and material-removal rate in machining, particularly in the aerospace and automotive industries.

The term high in *high-speed machining* (HSM) is somewhat relative; as a general guide, however, an approximate range of cutting speeds may be defined as follows:

1. **High speed:** 600 to 1800 m/min
2. **Very high speed:** 1800 to 18,000 m/min
3. **Ultrahigh speed:** Higher than 18,000 m/min.

Spindle rotational speeds in machine tools now range up to 50,000 rpm, although the automotive industry generally has limited them to 15,000 rpm for better reliability and less downtime in case a failure occurs. The *spindle power* required in high-speed machining is generally on the order of 0.004 W/rpm, much less than in traditional machining (0.2 to 0.4 W/rpm). Feed rates in high-speed machining are now up to 1 m/s.

Spindles for high speeds require *high stiffness* and *accuracy*, generally involving an integral electric motor. The armature is built onto the shaft, and the stator is placed in the wall of the spindle housing. The bearings may be rolling element or hydrostatic bearings; the latter is more desirable because it requires less space. Because of *inertia* during the acceleration and deceleration of machine components, the use of lightweight materials, including ceramics and composite materials, is an important consideration.

The selection of appropriate cutting-tool materials is always a major consideration. On the basis of the discussions of tools and their selection in Chapter 22, and especially by reviewing Table 22.2, it is apparent that, depending on the workpiece material, multiphase coated carbides, ceramics, cubic-boron nitride, and diamond are all candidate tool materials for high-speed operations.

It also is important to note that high-speed machining should be considered primarily for operations in which **cutting time** is a significant portion of the total time in the overall machining operation. As described in Chapter 40, **noncutting time** and various other factors are important considerations in the overall assessment of the benefits of high-speed machining.

Studies have indicated that high-speed machining is economical for many applications. As successful examples, it has been implemented in machining (a) aluminum structural components for aircraft; (b) submarine propellers 6 m in diameter, made of a nickel–aluminum–bronze alloy, weighing 55,000 kg (45 metric tons); and (c) automotive engines, with 5 to 10 times the productivity of traditional machining. High-speed machining of complex three- and five-axis contours has been made possible by advances in CNC control technology, as described in this chapter and in Chapter 37.

Another major factor in the adoption of high-speed machining has been the requirement to further improve dimensional tolerances. Note in Fig. 21.14 that as the cutting speed increases, a large percentage of the heat generated is removed by the chip. The tool and the workpiece remain closer to ambient temperature; this is beneficial because there is no significant thermal expansion, and thus warping, of the workpiece.

Important considerations in high-speed machining are:

1. Spindle design, for stiffness, accuracy, and balance at very high rotational speeds
2. Fast feed drives
3. Inertia of the components of the machine tool
4. Selection of appropriate cutting tools
5. Processing parameters and their computer control
6. Work-holding devices that can withstand high centrifugal forces
7. Chip-removal systems that are effective at very high rates of material removal.

25.6 Hard Machining

It has been noted that as the hardness of the workpiece increases, its machinability decreases, and tool wear and fracture, surface finish, and surface integrity can become significant problems. However, it is still possible to machine hard metals and alloys by selecting an appropriate hard-tool material and using machine tools with high stiffness, power, and precision.

An example is the finish machining of heat-treated steel (45 to 65 HRC) shafts, gears, pinions, and various automotive components, using polycrystalline cubic-boron nitride (PcBN), cermet, or ceramic cutting tools. Called *hard machining* or *hard turning*, it produces machined parts with good dimensional accuracy, surface finish (25 μm), and surface integrity. The important factors are the (a) available power, (b) static and dynamic stiffness of the machine tool and its spindle, and (c) workholding devices and fixturing.

As described in Section 25.3, trends in the design and construction of modern machine tools, especially for hard machining, include the use of hydrostatic bearings for the spindles and slideways. The headstock and the slanted bed in the machines (see Fig. 23.11a) can be made of *granite-epoxy composite materials*, with unique properties, such as high stiffness-to-weight ratio, thermal stability, and good damping capacity. Cutting-tool selection and edge preparation also are important to avoid premature failure in hard machining.

From technical, economic, and ecological considerations, hard turning has been found to compete successfully with *grinding* (Chapter 26). For instance, in some specific cases, hard turning has been shown to be three times faster than grinding, requiring fewer operations to finish the part, and utilizing five times less energy. A detailed comparative case of hard turning versus grinding is presented in Example 26.4.

25.7 Ultraprecision Machining

Beginning in the 1960s, increasing demands have been made concerning precision manufacturing of components for computer, electronic, nuclear, and defense applications. Specific examples include optical mirrors and lenses, fiber optic connection components, computer memory disks, metrology equipment, and drums for photocopying machines. Surface-finish requirements are in the nanometer (10^{-9} m) range, and dimensional tolerances and shape accuracies are in the micrometer (μm) and submicrometer range.

The trend toward ultraprecision manufacturing continues to grow. Modern **ultraprecision machine tools**, with advanced computer controls, can now position a cutting tool within an accuracy approaching 1 nm, as can be seen from Fig. 25.17. Note in this figure that higher precision is now being achieved by processes such as abrasive machining, ion-beam machining, and molecular manipulation.

The cutting tool for ultraprecision machining applications is almost exclusively a single-crystal diamond, where the process is called **diamond turning**. The tool has a polished cutting edge, with a radius as small as a few nm. Wear of the diamond can be a significant problem; more recent advances include **cryogenic diamond turning**, in which the tooling system is cooled by liquid nitrogen, to a temperature of about –120°C.

The materials for ultraprecision machining include copper and aluminum alloys, silver, gold, electroless nickel, infrared materials, and plastics (acrylics). With depths of cut in the nm range, hard and brittle materials produce continuous chips, in a process called **ductile-regime cutting** (Section 26.3.4). Deeper cuts in brittle materials produce discontinuous chips.

The machine tools for ultraprecision machining are built with very high precision and high stiffness of the machine, the spindle, and the workholding devices. The machines have components that are made of structural materials with low thermal expansion and good dimensional stability (see Section 25.3). They are located in a dust-free environment (*clean rooms*; Section 28.2), where the temperature is controlled to within a fraction of one degree.

Vibrations from internal machine sources, as well as from external sources such as nearby machinery on the same floor, must be avoided. Laser metrology (Section 35.5) is used for feed and position control,

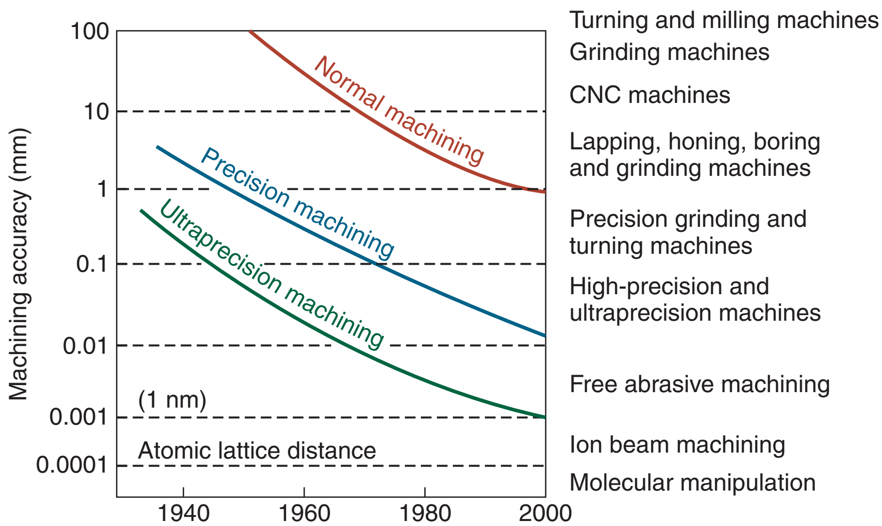


Figure 25.17: Improvements in machining accuracy over the years, using ultraprecision machining technologies. *Source:* After C.J. McKeown, N. Taniguchi, G. Byrne, D. Dornfeld, and B. Denkena.

and the machines are equipped with highly advanced computer-control systems and with thermal and geometric error-compensating features.

General Considerations for Precision Machining. There are several important factors in precision and ultraprecision machining and the machine tools, somewhat similar to those in high-speed machining:

1. Machine-tool design, construction, and assembly, including the spindle, must provide high stiffness, damping, and geometric accuracy
2. Motion control of the machine components, both linear and rotational
3. Thermal expansion of the machine tool, compensation for thermal expansion, control of the machine-tool environment
4. Real-time performance and control of the machine tool and implementation of a tool-condition monitoring system.

25.8 Machining Economics

Material and processing parameters relevant to efficient machining operations have been described in the preceding three chapters. In analyzing the *economics* of machining, several other factors also have to be considered. These factors include the costs involved in (a) machine tools, workholding devices and fixtures, and cutting tools; (b) labor and overhead associated with indirect costs; (c) the time required in setting up the machine for a particular operation; (d) material handling and movement, such as loading the blank and unloading the machined part; (e) gaging for dimensional accuracy and for surface finish; and (f) cutting and noncutting times.

Actual machining time is an important consideration; recall also the discussion in Section 25.5 regarding the role of noncutting time in high-speed machining. Unless noncutting time is a significant portion of the floor-to-floor time, high-speed machining should not be considered.

Economic analysis is based on the ability to achieve a desired outcome, such as tolerance and surface finish; as such, a machining process must be robust and under good control (see Section 36.5.1).

For example, if a milling cutter is mounted such that the exposed spindle length varies randomly with every tool change, then this alone could result in higher tolerances. The same analysis for different machine tools, where dynamic stiffness and damping may differ (see Section 25.3), the use of cutters with different numbers of inserts, or loss of ambient temperature control, all can result in variations that can significantly affect machining precision. *Full-factorial design of experiments* can characterize the machine-tool/workpiece/operator system, but this approach is complex and has its own limitations.

The section below assumes that a process has been carefully designed to be robust, so that variations in these contributing factors can be ignored, and the effect of cutting speed on economics and productivity can be explored.

Minimizing Machining Cost per Piece. As in all manufacturing processes and operations, the relevant parameters in machining can be selected and specified in such a manner that the *machining cost per piece*, as well as *machining time per piece*, is minimized. Several methods and approaches have been developed over the years to accomplish this goal, a task that has now become easier with increasing use of computers and user-friendly software. In order for the results of the methods used to be reliable, however, it is essential that input data be accurate and up to date. Described next is one of the simpler and more commonly used methods of analyzing machining costs in a *turning* operation.

In machining a part by turning, the total machining cost per piece, C_p is given by

$$C_p = C_m + C_s + C_l + C_t, \tag{25.1}$$

where

- $C_m =$ Machining cost
- $C_s =$ Cost of setting up for machining, including mounting the cutter, setting up fixtures, and preparing the machine tool for the operation
- $C_l =$ Cost of loading, unloading, and machine handling
- $C_t =$ Tooling cost, often only about 5% of the total machining operation; consequently, using the least expensive tool is not necessarily the proper way for reducing machining costs.

The **machining cost** is given by

$$C_m = T_m(L_m + B_m), \tag{25.2}$$

where T_m is the machining time per piece, L_m is the labor cost of production personnel per hour, and B_m is the *burden rate*, or *overhead charge*, of the machine, including depreciation, maintenance, and indirect labor.

The **setup cost** is a fixed figure in dollars per piece. The **loading, unloading,** and **machine-handling cost** is

$$C_l = T_l(L_m + B_m), \tag{25.3}$$

where T_l is the time involved in loading and unloading the part, in changing speeds and feed rates, and making any other adjustments before machining. The **tooling cost** is

$$C_t = \frac{1}{N_i}\left[T_c\left(L_m + B_m\right) + D_i\right] + \frac{1}{N_f}\left[T_i\left(L_m + B_m\right)\right], \tag{25.4}$$

where N_i is the number of parts machined per cutting tool insert, N_f is the number of parts that can be produced per insert edge, T_c is the time required to change the insert, T_i is the time required to index the insert, and D_i is the depreciation of the insert, in dollars.

The *time* required to machine one part is

$$T_p = T_l + T_m + \frac{T_c}{N_i} + \frac{T_i}{N_f}, \tag{25.5}$$

where T_m has to be calculated for each particular operation on the part. For example, in a turning operation the machining time (see Section 23.2) is given by

$$T_m = \frac{L}{fN} = \frac{\pi LD}{fV}, \tag{25.6}$$

where L is the length of cut, f is the feed, N is the angular speed (rpm) of the workpiece, D is the workpiece diameter, and V is the cutting speed. (Note that appropriate units must be used in all these equations.)

From Eq. (21.25) for tool life,

$$T = \left(\frac{C}{V}\right)^{1/n}, \tag{25.7}$$

where T is the time, in minutes, required to reach a flank wear of certain dimension, after which the tool has to be reground or changed. Note that the tool may have to be replaced due to other reasons as well, such as crater wear, built-up edge, or nose wear.

This analysis is restricted to *flank wear* as the important tool-failure criterion, but could also be made more elaborate to include other variables. The number of pieces machined per insert edge follows from Eq. (25.7), as

$$N_f = \frac{T}{T_m}, \tag{25.8}$$

and the number of pieces per insert is given by

$$N_i = mN_f = \frac{mT}{T_m}. \tag{25.9}$$

Combining Eqs. (25.6) through (25.9) yields

$$N_i = \frac{mfC^{1/n}}{\pi LDV^{(1/n)-1}}. \tag{25.10}$$

The cost per piece, C_p in Eq. (25.1), can now be defined in terms of several variables. To find the optimum cutting speed and the optimum tool life for **minimum cost**, C_p must be differentiated with respect to V and set to zero. Thus,

$$\frac{\partial C_p}{\partial V} = 0. \tag{25.11}$$

The *optimum cutting speed*, V_o, is

$$V_o = \frac{C\left(L_m + B_m\right)^n}{\left(\frac{1}{n} - 1\right)^n \left\{\frac{1}{m}\left[T_c\left(L_m + B_m\right) + D_i\right] + T_i\left(L_m + B_m\right)\right\}^n} \tag{25.12}$$

and the *optimum tool life*, T_o, is

$$T_o = \left(\frac{1}{n} - 1\right) \frac{\frac{1}{m}\left[T_c\left(L_m + B_m\right) + D_i\right] + T_i\left(L_m + B_m\right)}{L_m + B_m}. \tag{25.13}$$

To determine the optimum cutting speed and the optimum tool life for **maximum production**, T_p must be differentiated with respect to V and set to zero. Thus,

$$\frac{\partial T_p}{\partial V} = 0. \tag{25.14}$$

The **optimum cutting speed** then is

$$V_o = \frac{C}{\left[\left(\frac{1}{n}-1\right)\left(\frac{T_c}{m}+T_i\right)\right]^n}, \tag{25.15}$$

and the **optimum tool life** is

$$T_o = \left(\frac{1}{n}-1\right)\left(\frac{T_c}{m}+T_i\right). \tag{25.16}$$

Qualitative plots of *minimum cost per piece* and *minimum time per piece* (hence the *maximum production rate*) are given in Fig. 25.18a and b. It should be noted that the cost of machining a part also depends on the surface finish required. The additional cost increases rapidly with finer surface finish, as shown in Fig. 26.35.

The analysis above indicates the importance of (a) identifying all relevant parameters in a machining operation, (b) determining various cost factors, (c) obtaining relevant tool-life curves for the particular operation, and (d) properly measuring the various time intervals involved in the overall operation. The

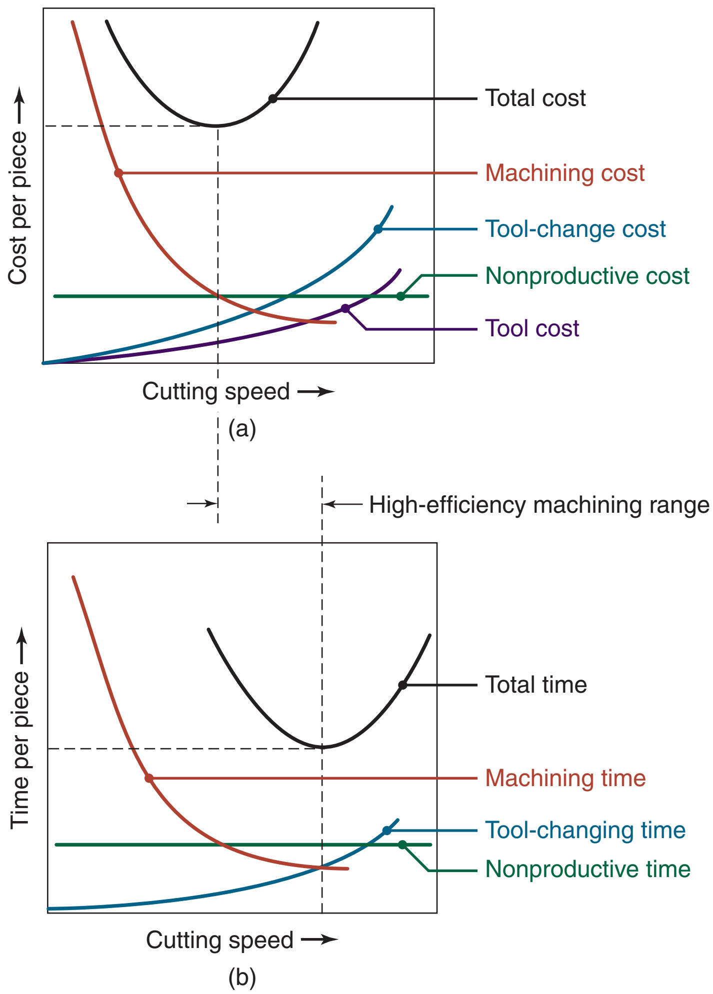


Figure 25.18: Graphs showing (a) cost per piece and (b) time per piece in machining. Note the optimum speeds for both cost and time. The range between the two is known as the *high-efficiency machining range*.

importance of obtaining accurate data is shown in Fig. 25.18. Note that small changes in cutting speed can have a significant effect on the minimum cost or minimum time per piece. The speeds and feeds recommended in Tables 23.4 and 24.2 generally lie in the *high-efficiency machining range*, which is between the speeds that yield the highest economy and highest production rate.

For many applications, such as finish machining of surfaces on soft metal castings, the machining cost per piece is fairly insensitive to cutting speed within this range; that is, the curve in Fig. 25.18 is fairly flat. With difficult-to-machine materials, however, as are routinely encountered in the medical products and aerospace industries, the cost per piece is very sensitive to cutting speed. Consequently, greater care have to be taken to ensure that machining takes place at or near the desired speed. Moreover, it should also be recognized that the data given in Tables 23.4 and 24.2 are a summary for various tool and material grades; specific data is often available for machining particular alloys.

Such an economic analysis can typically done for all manufacturing operations, and it can be a valuable tool for guiding process selection. For example, the cost per part in a sand-casting process to produce blanks, and in a machining operation to achieve final dimensional tolerances, can be calculated from an equation similar to Eq. (25.1), including also costs associated with sand casting, such as the cost of mold production and pattern depreciation. A similar calculation can be made on a processing approach that uses powder metallurgy (PM). Die and machinery costs will increase, but less machining is required because of PM's ability to produce net-shape parts and with tighter tolerances, thereby reducing machining costs. A comparison of cost estimates can then help determine a processing strategy, as described in greater detail in Section 40.10.

Summary

- Because they are versatile and capable of performing a wide variety of machining operations on small and large workpieces of various shapes, machining centers have become among the most important machine tools. Their selection depends on such factors as part complexity, the number and type of machining operations to be performed, the number of cutting tools required, and the dimensional accuracy and production rate specified.
- Vibration and chatter in machining operations are important considerations for workpiece dimensional accuracy, surface finish, and tool life. Stiffness and damping capacity of machine tools are major factors in controlling vibration and chatter. With proper selection of machining variables, operations can take place inside of stability lobes to avoid chatter.
- The economics of machining operations depends on such factors as nonproductive costs, machining costs, tool-change costs, and tool costs. Optimum cutting speeds can be determined for both minimum machining time per piece and minimum machining cost per piece.

Key Terms

Automatic pallet changer
Automatic tool changer
Chatter
Chip collection
Damping
Dynamic stiffness
Forced vibration
Hard machining
Hexapods
High-efficiency machining range
High-speed machining
Machining center
Modular machining center
Pallet
Reconfigurable machines
Regenerative chatter

Self-excited vibration
Stability lobes
Stiffness
Tool-exchange arm
Tool- and part-checking station
Touch probes
Turning center
Ultraprecision machining
Universal machining center
Work envelope

Bibliography

Cheng, K., **Machining Dynamics: Fundamentals, Applications and Practices**, Springer, 2010.
Dashchenko, A.I. (ed.), **Reconfigurable Manufacturing Systems and Transformable Factories**, Springer, 2006.
Davim, J.P., **Machining of Hard Materials**, Springer, 2011.
—, **Machining of Complex Sculptured Surfaces**, Springer, 2012.
Erdel, B., **High-speed Machining**, Society of Manufacturing Engineers, 2003.
Gegg, B.C., Suh, C.S., and Luo, C.J., **Machine Tool Vibrations and Cutting Dynamics**, Springer, 2011.
Ito, Y., **Modular Design for Machine Tools**, McGraw-Hill, 2008.
—, **Thermal Deformation in Machine Tools**, McGraw-Hill, 2010.
Joshi, P.H., **Machine Tools Handbook**, McGraw-Hill, 2008.
Lopez de Lacalle, N., and Lamikiz, A. (eds.), **Machine Tools for High Performance Machining**, Springer, 2009.
Mickelson, D., **Hard Milling & High Speed Machining: Tools of Change**, Industrial Press, 2006.
Rivin, E.I., **Stiffness and Damping in Mechanical Design**, ASME Press, 2010.
Schmitz, T.L., and Smith, K.S., **Machining Dynamics: Frequency Response to Improved Productivity**, Springer, 2008.
Suh, C.S., and Liu, M.-K., **Control of Cutting Vibration and Machining Instability**, Wiley, 2013.
Zhang, D., **Parallel Robotic Machine Tools**, Springer, 2010.

Review Questions

25.1. Describe the distinctive features of machining centers and explain why these machines are so versatile.

25.2. Explain how the tooling system in a machining center functions. What are the typical tool-changing times?

25.3. Explain the trends in materials used for machine-tool structures.

25.4. Is there any difference between chatter and vibration? Explain.

25.5. What are the differences between forced and self-excited vibration?

25.6. Explain the importance of foundations for installing machine tools.

25.7. Explain why automated pallet changers and automatic tool changers are important parts of machining centers.

25.8. What types of materials are machine-tool bases typically made from? Why?

25.9. What is meant by the "modular" construction of machine tools?

25.10. What is a hexapod? What are its advantages?

25.11. What factors contribute to costs in machining operations?

25.12. What is the high-efficiency machining range? Why is it so called?

25.13. List the reasons that temperature is important in machining operations.

Qualitative Problems

25.14. Explain the technical and economic factors that led to the development of machining centers.

25.15. Spindle speeds in machining centers vary over a wide range. Explain why this is so, giving specific applications.

25.16. Explain the importance of stiffness and damping of machine tools. Describe how they are implemented.

25.17. Are there machining operations described in Chapters 23 and 24 that cannot be performed in machining and turning centers? Explain, with specific examples.

25.18. How important is the control of cutting-fluid temperature in operations performed in machining centers? Explain.

25.19. Review Fig. 25.10 on modular machining centers, and describe some workpieces and operations that would be suitable on such machines.

25.20. Review Fig. 25.15 and estimate the amount of damping you would expect in a hexapod. Is vibration a serious concern with hexapods? Explain.

25.21. Describe the adverse effects of vibration and chatter in machining operations.

25.22. Describe some specific situations in which thermal distortion of machine-tool components would be important.

25.23. Explain the differences in the functions of a turret and of a spindle in turning centers.

25.24. Explain how the pallet arrangements shown in Fig. 25.4a and b would be operated in using these machines on a shop floor.

25.25. Review the tool changer shown in Fig. 25.5. Are there any constraints on making their operations faster in order to reduce the tool changing time? Explain.

25.26. List the parameters that influence the temperature in metal cutting, and explain why and how they do so.

25.27. List and explain factors that contribute to poor surface finish in machining operations.

25.28. Can high-speed machining be performed without the use of cutting fluids? Explain.

25.29. In addition to the number of joints in a machine tool (see Fig. 25.15), what other factors influence the rate at which damping increases? Explain.

25.30. Describe types and sizes of workpieces that would not be suitable for machining on a machining center. Give specific examples.

25.31. Other than the fact that they each have a minimum, are the overall shapes and slopes of the total-cost and total-time curves in Fig. 25.18 important? Explain.

25.32. Explain the advantages and disadvantages of machine-tool frames made of gray-iron castings.

25.33. What are the advantages and disadvantages of (a) welded-steel frames, (b) bolted steel frames, and (c) adhesively bonded components of machine tools? Explain.

25.34. What would be the advantages and limitations of using concrete or polymer–concrete in machine tools?

25.35. Explain how you would go about reducing each of the cost factors in machining operations. What difficulties would you encounter in doing so?

25.36. Describe workpieces that would not be suitable for machining on a machining center. Give specific examples.

25.37. Give examples of forced vibration or self-excited vibration in general engineering practice.

Quantitative Problems

25.38. A machining-center spindle and tool extend 200 mm from their machine-tool frame. Calculate the temperature change that can be tolerated in order to maintain a tolerance of 0.025 mm in machining. Assume that the spindle is made of steel.

25.39. Using the data given in the example, estimate the time required to manufacture the parts in Example 25.1 with conventional machining and with high-speed machining.

25.40. A machining-center spindle and tool extend 500 mm from its machine-tool frame. What temperature change can be tolerated to maintain a tolerance of 0.0025 mm in machining? A tolerance of 0.025 mm? Assume that the spindle is made of steel.

25.41. In the production of a machined valve, the labor rate is \$30.00 per hour, and the general overhead rate is \$25.00 per hour. The tool is a ceramic insert with four faces and costs \$30.00, takes six minutes to change and one minute to index. Estimate the optimum cutting speed from a cost perspective. Use $C = 100$ for V_o in $\mathrm{m/min}$.

25.42. Estimate the optimum cutting speed in Problem 25.41 for maximum production.

25.43. Develop an equation for optimum cutting speed in face milling using a cutter with inserts.

25.44. Develop an equation for optimum cutting speed in turning where the tool is a high speed steel tool that can be reground periodically.

Synthesis, Design, and Projects

25.45. If you were the chief engineer in charge of the design of advanced machining and turning centers, what changes and improvements would you recommend on existing models? Explain.

25.46. Review the technical literature and outline the trends in the design of modern machine tools. Explain why there are those trends.

25.47. Make a list of components of machine tools that could be made of ceramics, and explain why ceramics would be suitable.

25.48. Survey the company literature from various machine-tool manufacturers, and prepare a comprehensive table indicating the capabilities, sizes, power, and costs of machining and turning centers. Comment on your observations.

25.49. The cost of machining and turning centers is considerably higher than for traditional machine tools. Since many operations performed by machining centers also can be done on conventional machines, how would you go about justifying the high cost of these centers? Explain with appropriate examples.

25.50. In your experience using tools or other devices, you may have come across situations in which where you experienced vibration and chatter. Describe your experience and explain how you would go about minimizing the vibration and chatter.

25.51. Describe your thoughts on whether or not it is feasible to include grinding operations (see Chapter 26) in machining centers. Explain the nature of any difficulties that may be encountered.

25.52. Is the accuracy and surface finish that can be achieved in a machining center a function of the number of inserts on a cutter? Explain.

25.53. The following experiment is designed to better demonstrate the effect of tool overhang on vibration and chatter: With a sharp tool, scrape the surface of a piece of soft metal by holding the tool with your arm fully outstretched. Repeat the experiment, this time holding the tool as close to the workpiece as possible. Describe your observations regarding the tendency for the tool to vibrate. Repeat the experiment with different types of metallic and nonmetallic materials.

25.54. Review the part in Fig. 25.1a and list the machining operations and machine tools you would recommend to produce this part.

Chapter 26

Abrasive Machining and Finishing Operations

- Abrasive machining is among the major finishing operations in the production of parts, and is important because of its capability to impart high dimensional accuracy and surface finish.
- This chapter opens with a description of the grinding process, the mechanics of material removal, and the types of abrasives and bonds used in grinding wheels.

- Some abrasive machining operations, including polishing, buffing, honing, and sanding, require a bonded or coated abrasive; others, such as ultrasonic machining, lapping, abrasive flow machining, and electrochemical machining and grinding, have loose abrasives.
- The fundamentals of all abrasive processes are described in detail, including their principles, applications, and design considerations.
- The chapter concludes with a discussion of economic considerations for finishing operations.

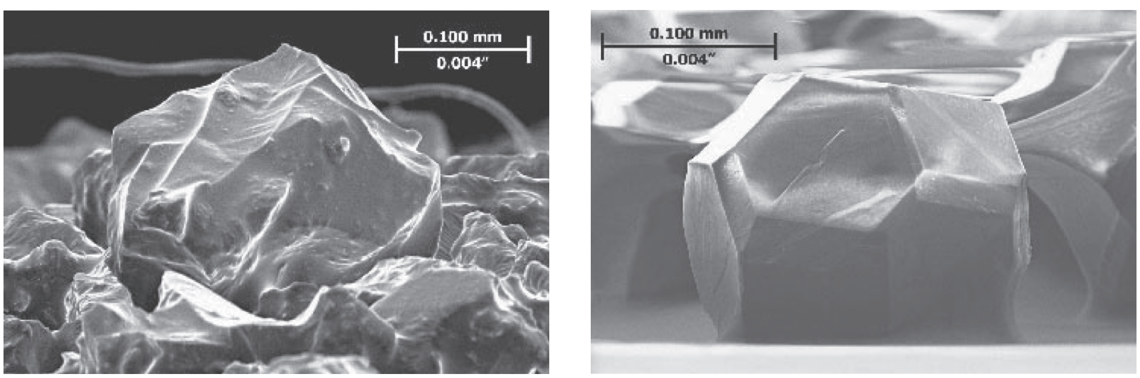
Typical parts made: Any part requiring high dimensional accuracy and surface finish, such as ball and roller bearings, piston rings, valves, cams, gears, and tools and dies.

Alternative processes: Precision machining, electrical-discharge machining, electrochemical machining and grinding, and abrasive-jet machining.

26.1 Introduction

There are numerous situations in manufacturing where the processes described thus far cannot produce the required dimensional accuracy or surface finish, or the workpiece material is too hard or too brittle to process. For example, consider the dimensional accuracy and fine surface finish required on ball bearings, pistons, valves, cylinders, cams, gears, molds and dies, and the precision components in instrumentation. One of the most common and economical methods for producing such demanding characteristics on parts is *abrasive machining*.

An **abrasive** is a very small, hard particle having sharp edges and an irregular shape (Fig. 26.1). The simplest example is *sand*, which is capable of removing small amounts of material from a surface by scratching it, producing tiny chips. Familiar applications of abrasives are *sandpaper* or *emery cloth*, used to smoothen



(a) (b)

Figure 26.1: Typical abrasive grains; note the angular shape with sharp edges. (a) A single, 80-mesh Al_2O_3 grit in a freshly dressed grinding wheel; (b) an 80/100 mesh diamond grit. Diamond and cubic boron nitride grains can be manufactured in various geometries, including the "blocky" shape shown. *Source:* Courtesy of J. Badger.

Figure 26.2: A variety of bonded abrasives used in abrasive-machining processes. *Source:* Shutterstock/ Praethip Docekalova.

surfaces and remove sharp corners, and *grinding wheels*, as shown in Figs. 26.2 and 26.3, to sharpen knives, tools, or to impart good dimensional accuracy and surface finish. Abrasives also are used to hone, lap, buff, and polish workpieces.

With the use of computer-controlled machines, abrasive processes and equipment are now capable of producing a wide variety of part shapes, as can be seen in Fig. 26.3, and very fine dimensional accuracy and surface finishes, as shown in Figs. 23.14 and 33.5. For example, dimensional tolerances on parts can now be less than 1 μm and surface roughness can be as fine as 0.025 μm.

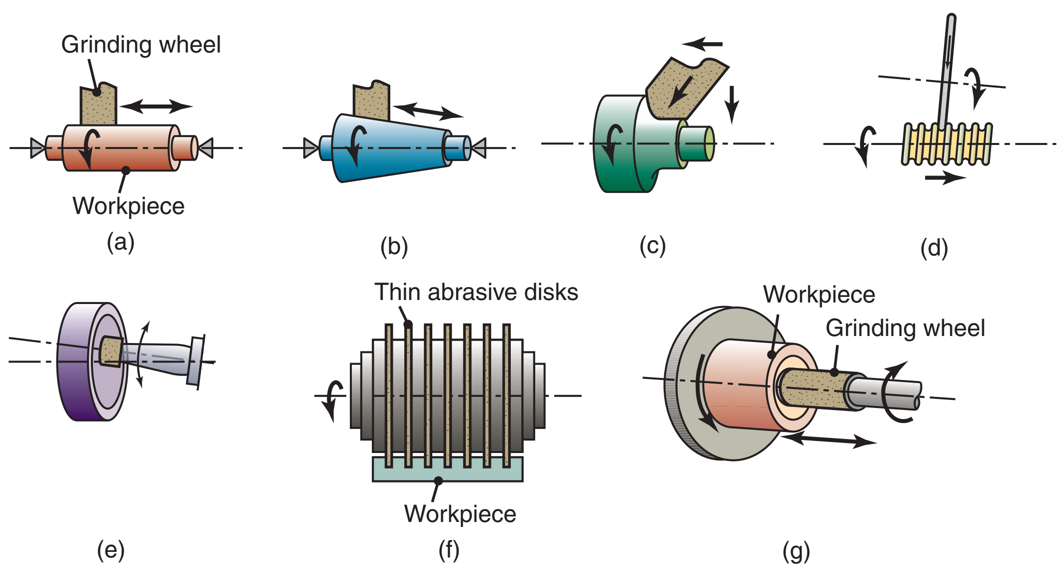


Figure 26.3: The types of workpieces and operations typical of grinding: (a) cylindrical surfaces, (b) conical surfaces, (c) fillets on a shaft, (d) helical profiles, (e) concave shape, (f) cutting off or slotting with thin wheels, and (g) internal grinding.

Because they are hard, abrasives also are used in *finishing processes* for heat-treated metals and alloys, and for very hard parts, in such applications as (a) finishing of ceramics and glasses, (b) cutting off lengths of bars, structural shapes, masonry, and concrete, (c) removing unwanted weld beads and spatter, (d) creating a very smooth and flat surface on silicon wafers to make integrated circuits, (e) polishing bearings and races, and (f) cleaning surfaces with jets of air or water containing abrasive particles.

This chapter begins with a description of abrasive characteristics, along with their use in various abrasive material-removal processes. The mechanics of abrasive operations is first described. This knowledge is essential in establishing the interrelationships between the (a) workpiece material and process variables and (b) dimensional accuracy, surface finish, and surface integrity of the parts produced by abrasive machining.

26.2 Abrasives and Bonded Abrasives

Abrasives that are used most commonly in abrasive-machining operations are:

Conventional abrasives

- *Aluminum oxide* (Al_2O_3)
- *Silicon carbide* (SiC)

Superabrasives

- *Cubic boron nitride* (cBN)
- *Diamond*

As described in Chapter 8, these abrasives are much harder than conventional cutting-tool materials, as may be seen by comparing Tables 22.1 and 26.1 (see also Fig. 2.15). Cubic boron nitride and diamond are listed as superabrasives because they are the two hardest materials known.

In addition to hardness, an important characteristic of abrasives is **friability**, defined as the ability of abrasive grains to fracture into smaller pieces. This property gives abrasives their *self-sharpening* characteristics, essential in maintaining their sharpness during use. High friability indicates low strength or low fracture resistance of the abrasive. Thus, a highly friable abrasive grain fragments more rapidly under grinding forces than one with low friability. Aluminum oxide, for example, has lower friability than silicon carbide and, correspondingly, lower tendency to fragment.

The *shape* and *size* of the abrasive grain affect its friability. For example, *blocky* grains, which are analogous to a negative rake angle in single-point cutting tools (Fig. 21.3), are less friable than less blocky or platelike grains. Moreover, because the probability of defects decreases as grain size decreases, smaller grains are stronger and less friable than larger ones (a phenomenon known as *size effect*).

Table 26.1: Ranges of Knoop Hardness for Various Materials and Abrasives.

Common glass	350–500
Flint, quartz	800–1100
Zirconium oxide	1000
Hardened steels	700–1300
Tungsten carbide	1800–2400
Aluminum oxide	2000–3000
Titanium nitride	2000
Titanium carbide	1800–3200
Silicon carbide	2100–3000
Boron carbide	2800
Cubic boron nitride	4000–5000
Diamond	7000–8000

Abrasive Types. The abrasives commonly found in nature are *emery*, *corundum* (alumina), *quartz*, *garnet*, and *diamond*. Because in their natural state abrasives generally contain impurities and possess nonuniform properties, their performance as an abrasive can be inconsistent and unreliable. Consequently, abrasives have been made *synthetically* for many years.

- **Aluminum oxide** was first made in 1893, and is produced by fusing bauxite, iron filings, and coke. Fused aluminum oxides are categorized either as *dark* (less friable), *white* (very friable), or *single crystal*.

- **Seeded gel** is the purest form of *unfused aluminum oxide*; it was first introduced in 1987. Also known as *ceramic aluminum oxide*, it has a grain size on the order of 0.2 μm (coarse human hair is about 200 μm), which is much smaller than other types of commonly used abrasive grains. The grains are *sintered* (heating without melting; Section 17.4) to become larger in size. Because they are harder than fused alumina and have relatively high friability, seeded gels maintain their sharpness and are especially effective for difficult-to-grind materials.

- **Silicon carbide** was first discovered in 1891, and is made with silica sand and petroleum coke. Silicon carbides are classified as *black* (less friable) or *green* (more friable). They generally have higher friability than aluminum oxide; hence, they display greater tendency to fracture and thus remain sharp.

- **Cubic boron nitride** was first developed in the 1970s; its properties and characteristics are described in Sections 8.2.3 and 22.7.

- **Diamond**, also known as *synthetic* or *industrial diamond*, was first used as an abrasive in 1955; its properties and characteristics are described in Sections 8.7 and 22.9.

Abrasive Grain Size. As used in manufacturing operations, abrasives generally are very small when compared to the size of cutting tools and inserts (Chapters 21 and 22). They have sharp edges, allowing removal of very small quantities of material from a workpiece surface, resulting in very fine surface finish and dimensional accuracy.

The size of an abrasive grain is identified by a **grit number**, which is a function of sieve size; the smaller the grain size, the larger is the grit number. For example, grit number 10 is typically regarded as very coarse, 100 as fine, and 500 as very fine. Sandpaper and emery cloth also are identified in this manner, as can readily be observed by noting the grit number printed on the backs of abrasive papers or cloth.

Compatibility of Abrasive and Workpiece Material. As in selecting cutting-tool materials for machining, the *affinity* of an abrasive grain to the workpiece material is an important consideration. The less the reactivity of the two materials, the less the wear and dulling of the grains during grinding, making the operation more efficient and causing less damage to the workpiece surface (see Section 26.3.1 for details). As an example of chemical affinity, diamond (a form of carbon, Section 8.7) cannot be used for grinding steels, since diamond dissolves in iron at the high temperatures encountered in grinding. Generally, the following recommendations are made with regard to selecting abrasives:

- **Aluminum oxide:** Carbon steels, ferrous alloys, and alloy steels.

- **Silicon carbide:** Nonferrous metals, cast irons, carbides, ceramics, glass, and marble.

- **Cubic boron nitride:** Steels and cast irons above 50 HRC hardness and high-temperature alloys.

- **Diamond:** Ceramics, carbides, and some hardened steels where the hardness of diamond is more significant than its reactivity with the carbon in steel.

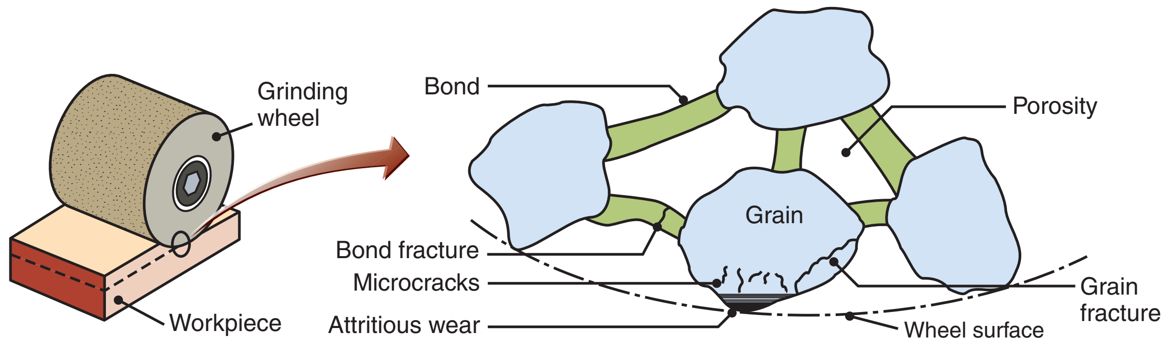


Figure 26.4: Schematic illustration of a physical model of a grinding wheel, showing its structure and its wear and fracture patterns.

26.2.1 Grinding Wheels

Each abrasive grain typically removes only a tiny amount of material at a time; consequently, high material-removal rates can only be achieved if a very large number of these grains remove material. This is often done by using **bonded abrasives**, typically in the form of a grinding wheel, in which the abrasive grains are distributed and oriented randomly.

As shown schematically in Fig. 26.4, the abrasive grains in a grinding wheel are held together by a **bonding material** (Section 26.2.2), which acts as supporting posts or braces between the grains. In bonded abrasives, *porosity* is essential to provide clearance for the chips being produced; otherwise, the chips would severely interfere with the grinding operation. Porosity can be observed by looking at the surface of a grinding wheel with a magnifying glass or under a microscope.

A very wide variety of types and sizes of abrasive wheels is now available. Some of the more commonly used types of grinding wheels made of conventional abrasives are shown in Fig. 26.5; superabrasive wheels are shown in Fig. 26.6. Note that, due to their high cost, only a small volume of superabrasive material is used on the periphery of these wheels; also, it is not necessary to have more abrasives because their wear is extremely small.

Bonded abrasives are marked with a standardized system of letters and numbers, indicating the type of abrasive, grain size, grade, structure, and bond type. Figure 26.7 shows the marking system for aluminum-oxide and silicon-carbide bonded abrasives. The marking system for diamond and cubic boron nitride bonded abrasives is shown in Fig. 26.8.

The cost of grinding wheels depends on the type and size of the wheel. Small wheels [up to about 25 mm in diameter] cost approximately $2 to $15 for conventional abrasives, $30 to $100 for diamond, and $50 to $300 for cubic boron nitride wheels. For a large wheel of about 500 mm in diameter and 250 mm in width, the costs are $500 for conventional abrasives, $5000 to $8000 for diamond, and as high as $20,000 for cubic boron nitride.

26.2.2 Bond Types

The common types of bonds used in bonded abrasives are:

Vitrified. Also called *ceramic* bond, *vitrified bonds* (from the Latin *vitrum* for glass; Section 8.4) are the most common and widely used material. The raw materials consist of feldspar (a crystalline mineral) and clays. They are mixed with the abrasives, moistened, and molded under pressure into the shape of grinding wheels. These *green* wheels are then fired slowly, up to a temperature of about 1250°C, to fuse the glass and develop structural strength. The wheels are then cooled slowly (to avoid temperature gradients within the

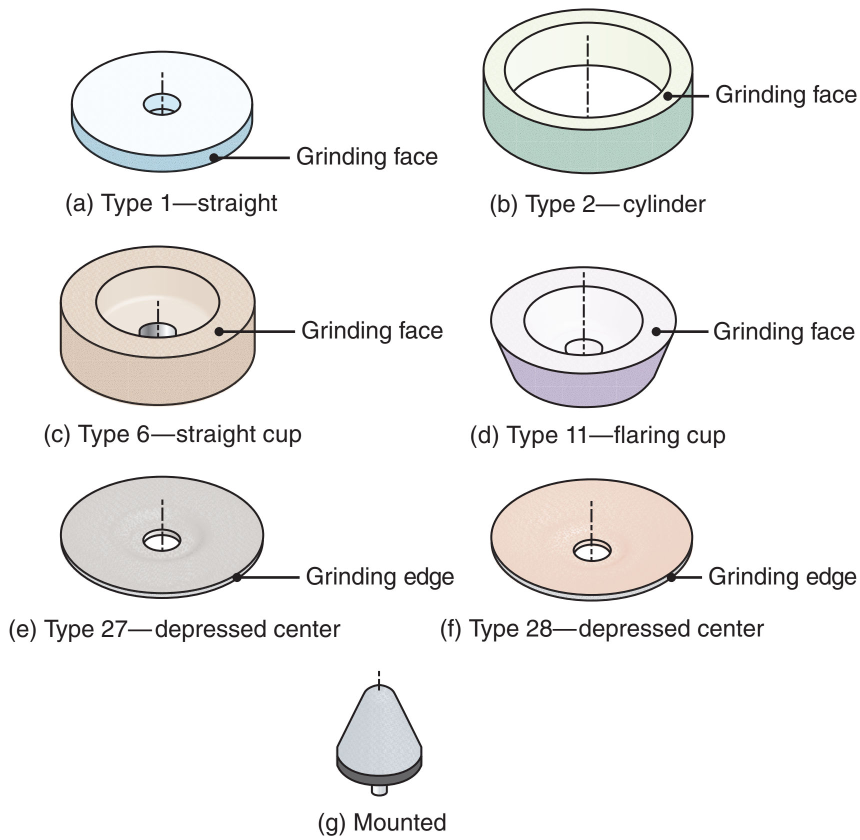


Figure 26.5: Common types of grinding wheels made with conventional abrasives. Note that each wheel has a specific *grinding face*; grinding on other surfaces is improper and unsafe.

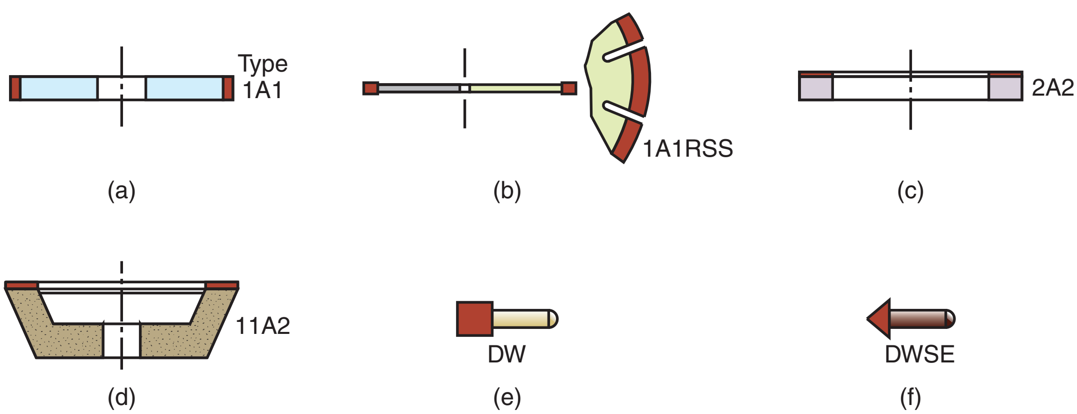


Figure 26.6: Examples of superabrasive wheel configurations. The annular regions (rims) are grinding surfaces; the wheel itself (core) generally is made of metal or composites. The bonding materials for superabrasives are (a), (d), and (e) resinoid, metal, or vitrified; (b) metal; (c) vitrified; and (f) resinoid.

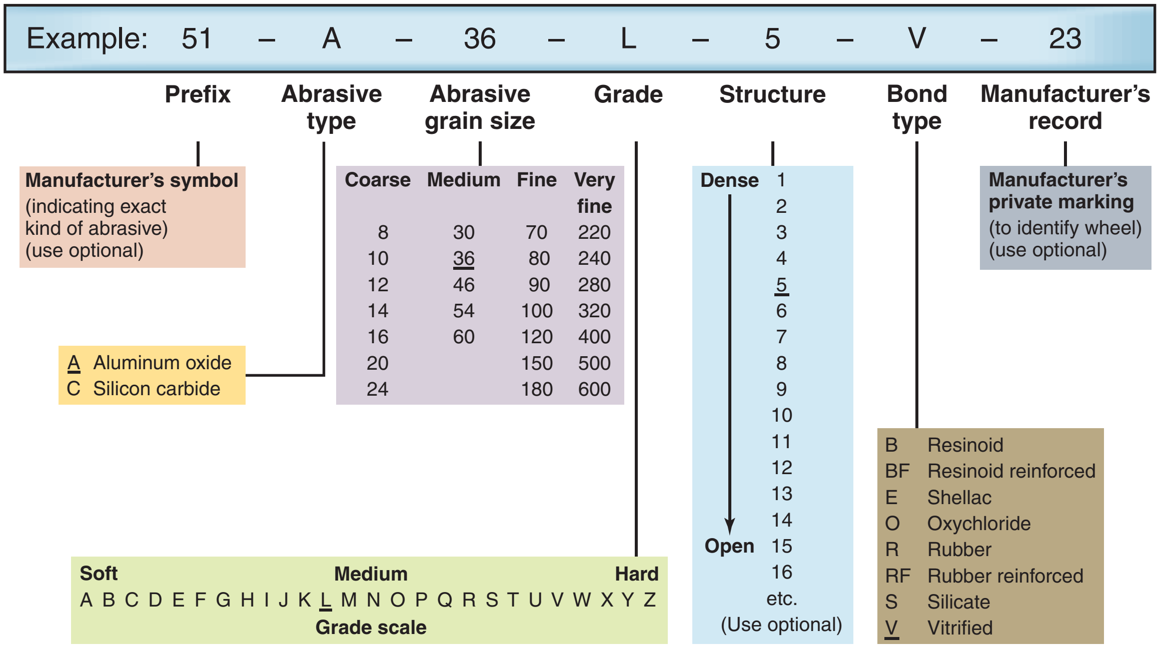


Figure 26.7: Standard marking system for aluminum-oxide and silicon-carbide bonded abrasives.

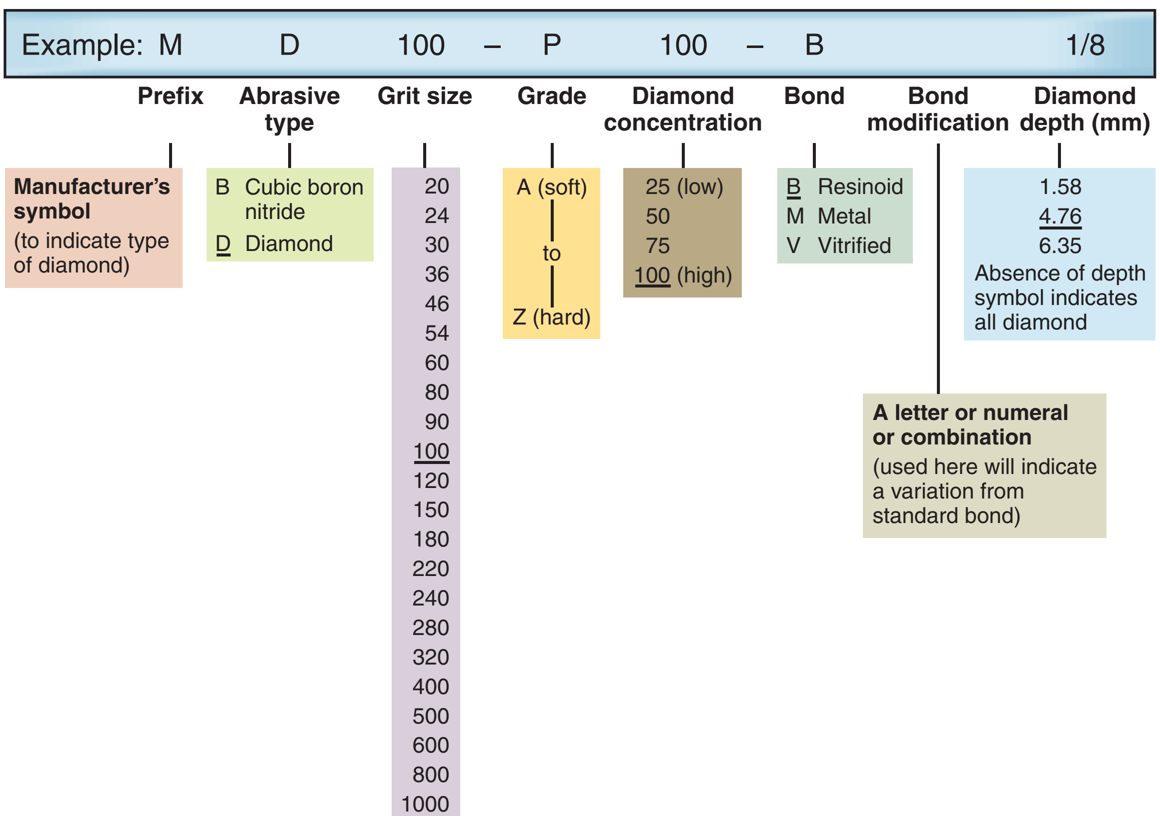


Figure 26.8: Standard marking system for cubic boron nitride and diamond bonded abrasives.

wheels and associated thermal cracking), finished to size, inspected for quality and dimensional accuracy, and tested for any defects.

Wheels with vitrified bonds are strong, stiff, and resistant to oils, acids, and water; however, they are brittle and thus lack resistance to mechanical and thermal shock. To improve strength during their use, vitrified wheels also are made with steel-backing plates or cups, for better structural support of the bonded abrasives. The color of a grinding wheel can be modified by adding various elements during its manufacture, so that wheels can be color coded for use with specific workpiece materials, such as ferrous, nonferrous, and ceramic.

Resinoid. Resinoid bonding materials are *thermosetting resins*, and are available in a wide range of compositions and properties (Sections 7.4 and 7.7). Because the bond is an organic compound, wheels with *resinoid bonds* also are called **organic wheels**. The process for making them consists basically of (a) mixing the abrasive with liquid or powdered phenolic resins and additives, (b) pressing the mixture into the shape of a grinding wheel, and (c) curing it at temperatures of about 175°C to set the bond. In addition to pressing, *injection molding* also is used to make grinding wheels (see Sections 17.3 and 19.3).

Because the elastic modulus of thermosetting resins is lower than that of glasses (see Table 2.2), resinoid wheels are more flexible than vitrified wheels. As a bonding material, *polyimide* (Section 7.7) also is used as a substitute for the phenolic resin; it is tougher and more resistant to higher temperatures.

Reinforced Wheels. These wheels typically consist of one or more layers of *fiberglass mats* of various mesh sizes. The fiberglass in this laminate structure retards the disintegration of the wheel should the wheel break for some reason during its use. Large-diameter resinoid wheels can be further supported by using one or more internal rings, made of round steel bars inserted during molding of the wheel.

Thermoplastic. In addition to thermosetting resins, thermoplastic plastics (Section 7.3) also are used in grinding wheels. Wheels are available with sol-gel abrasives bonded with thermoplastics.

Rubber. The most flexible matrix used in abrasive wheels is rubber (Section 7.9). The manufacturing process consists of (a) mixing crude rubber, sulfur, and the abrasive grains together, (b) rolling the mixture into sheets, (c) cutting out disks of various diameters, and (d) heating the disks under pressure to vulcanize the rubber. Thin wheels (called *cutoff blades*) can be made in this manner, and are used like circular saws for cutting-off operations.

Metal. Using powder-metallurgy techniques, the abrasive grains, usually diamond or cubic boron nitride, are bonded to the periphery of a metal wheel to depths of 6 mm or less (Fig. 26.6). Metal bonding is carried out under high pressure and temperature. The wheel itself (the core) may be made of aluminum, bronze, steel, ceramics, or composite materials, depending on such requirements as strength, stiffness, and dimensional stability. Superabrasive wheels may be *layered*, so that a single abrasive layer is plated or brazed to a metal wheel. Layered wheels are lower in cost, and are used for small production batches.

26.2.3 Wheel Grade and Structure

The **grade** of a bonded abrasive is a measure of its bond strength, including both the type and the amount of bonding material in the wheel. Because strength and hardness are directly related (Section 2.6.2), the grade is also referred to as the **hardness** of a bonded abrasive. Thus, a hard wheel has a stronger bond and/or a larger amount of bonding material between the grains than does a soft wheel.

The **structure** of a bonded abrasive is a measure of its *porosity* (the spacing between the grains, as shown in Fig. 26.4). The structure ranges from *dense* to *open*, as shown in Fig. 26.7. Recall that some porosity is essential to provide clearance for the chips, as otherwise they would interfere with the grinding operation.

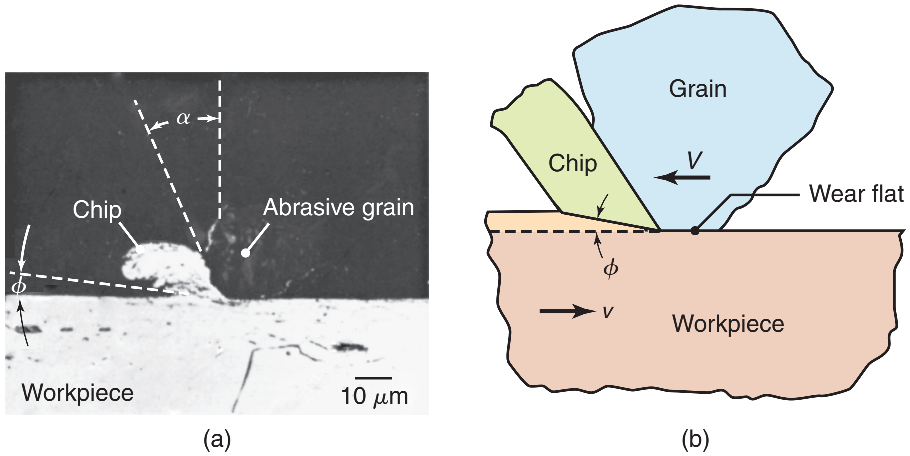


Figure 26.9: (a) Grinding chip being produced by a single abrasive grain; note the large negative rake angle of the grain. (b) Schematic illustration of chip formation by an abrasive grain with a wear flat; note the negative rake angle of the grain and the small shear angle. *Source:* (a) After M.E. Merchant.

26.3 The Grinding Process

Grinding is a chip-removal process that uses an individual abrasive grain as the cutting tool (Fig. 26.9a). The major differences between the action of an abrasive grain and that of a single-point cutting tool can be summarized as:

- Individual abrasive grains have *irregular shapes* (Fig. 26.1), and are spaced randomly along the periphery of the wheel (Fig. 26.10).

- The average rake angle of the grains is highly negative, typically -60° or even less; consequently, grinding chips undergo much larger plastic deformation than they do in other machining processes (see Section 21.2).

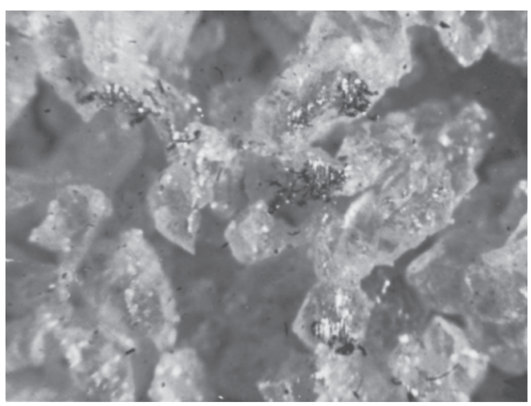


Figure 26.10: The surface of a grinding wheel (A46-J8V), showing abrasive grains, wheel porosity, wear flats on grains, and metal chips from the workpiece adhering to the grains. Note the random distribution and shape of the abrasive grains. Magnification: $50\times$. *Source:* S. Kalpakjian.

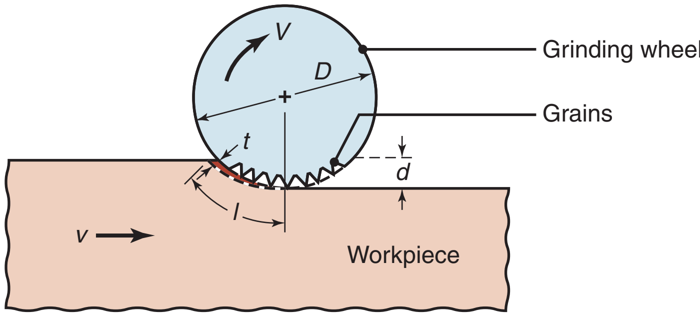


Figure 26.11: Schematic illustration of the surface-grinding process, showing various process variables. The figure depicts conventional (up) grinding.

- The radial positions of the grains over the peripheral surface of a wheel vary, and not all grains are active during grinding.
- Surface speeds of grinding wheels (equivalent to cutting speeds) are very high, typically 20 to 30 m/s, and can be as high as 150 m/s in high-speed grinding, using specially designed wheels.

The grinding process and its parameters can best be observed in a *surface-grinding* operation, shown schematically in Fig. 26.11. A straight grinding wheel (Fig. 26.5a), with a diameter of D, removes a layer of metal at a depth d (called **wheel depth of cut**). An individual grain on the periphery of the wheel moves at a tangential velocity of V, while the workpiece moves at a velocity of v. Each abrasive grain produces a small chip, which has an *undeformed thickness* (**grain depth of cut**), t, and an *undeformed length*, l.

Typical chips from grinding operations are shown in Fig. 26.12; note that the chips, just as in machining, are thin and long. From geometric relationships, it can be shown that the undeformed chip length in surface grinding (Fig. 26.11) is approximated by the equation

$$l = \sqrt{Dd} \tag{26.1}$$

and the undeformed chip thickness, t, by

$$t = \sqrt{\left(\frac{4v}{VCr}\right)\sqrt{\left(\frac{d}{D}\right)}}, \tag{26.2}$$

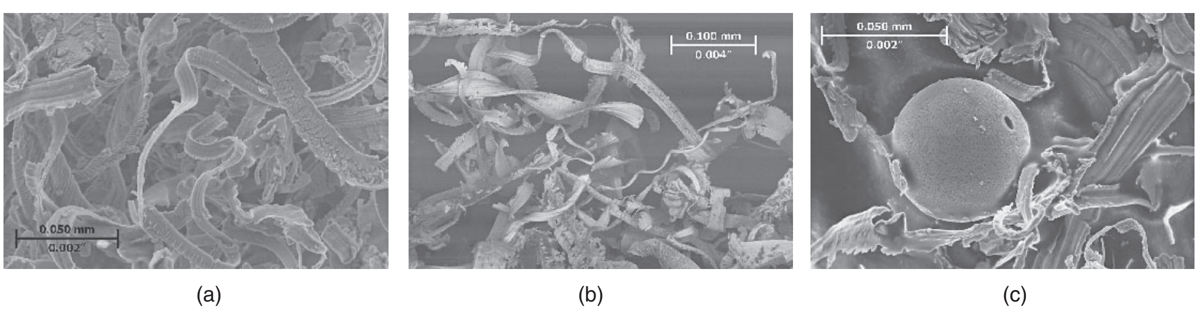


Figure 26.12: Typical chips, or swarf, from grinding operations. (a) Swarf from grinding a conventional HSS drill bit; (b) swarf of nickel-alloy workpiece using an Al_2O_3 wheel; (c) swarf of M2 high-speed steel using an Al_2O_3 wheel, showing a melted globule among the chips. *Source:* Courtesy of J. Badger, *The Grinding Doc*.

where C is the number of cutting points per unit area of the wheel periphery. Generally, C is in the range from 0.1 to 10 per mm^2. The quantity r is the ratio of chip width to average undeformed chip thickness, and is estimated between 10 and 20.

As an example, l and t can be taken as functions of process parameters. Consider the case where $D = 200$ mm, $d = 0.05$ mm, $v = 30$ m/min, and $V = 1800$ m/min. Using the preceding formulas gives

$$l = \sqrt{(200)(0.05)} = 3.2 \text{ mm}$$

Assuming that $C = 2$ per mm^2 and that $r = 15$ gives

$$t = \sqrt{\frac{(4)(30)}{(1800)(2)(15)}\sqrt{\frac{0.05}{200}}} = 0.006 \text{ mm}$$

Because of plastic deformation during chip formation, the actual chip will be shorter and thicker than the values calculated (see Figs. 26.9 and 26.12). Note from this example that grinding chip dimensions typically are much smaller than those in metal-cutting operations.

Grinding Forces. A knowledge of grinding forces is essential for

- Estimating power requirements.
- Designing grinding machines and workholding devices and fixtures.
- Determining the deflections that the workpiece, as well as the grinding machine and its components, may undergo. Deflections adversely affect dimensional accuracy, and are especially critical in precision and ultraprecision grinding.

Assuming that the cutting force on the grain is proportional to the cross-sectional area of the undeformed chip, it can be shown that the **grain force** (acting tangential to the wheel) is a function of process variables and is given as:

$$\text{Grain force} \propto \left(\frac{v}{V}\sqrt{\frac{d}{D}}\right)(S_{\text{ut}}). \tag{26.3}$$

Because of the small dimensions involved, forces in grinding are typically much smaller than those in the machining operations described in Chapters 23 and 24. Grinding forces should be kept low, in order to avoid distortion and to maintain high dimensional accuracy of the workpiece.

Specific Energy. The *energy* dissipated in producing a grinding chip consists of the energy required for the following:

- *Plastic deformation* in chip formation.
- *Plowing*, as shown by the ridges formed in Fig. 26.13.
- *Friction*, caused by rubbing of the abrasive grain along the workpiece surface.

Note in Fig. 26.9b that, after some use, the grains along the periphery of the wheel develop a **wear flat**, a phenomenon similar to *flank wear* in cutting tools, shown in Fig. 21.15. The wear flat continuously rubs over the ground surface, dissipates energy (because of friction), and thus makes the grinding operation less efficient.

The **specific energy** in grinding is defined as the energy per unit volume of material ground from the workpiece surface (Table 26.2). Note that the energy levels are much higher than those in machining operations (Table 21.2). This difference has been attributed to such factors as the presence of a wear flat, high negative rake angles of the abrasive grains (which require more energy; Section 21.3), and a possible

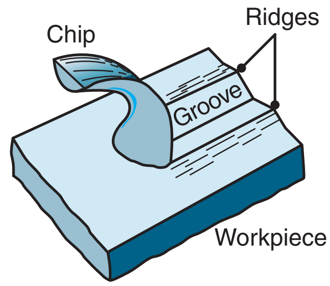


Figure 26.13: Chip formation and plowing of the workpiece surface by an abrasive grain.

Table 26.2: Approximate Specific-energy Requirements for Surface Grinding.

Workpiece material	Hardness	Specific energy W-s/mm^3
Aluminum	150 HB	7–27
Cast iron (class 40)	215 HB	12–60
Low-carbon steel (1020)	110 HB	14–68
Titanium alloy	300 HB	16–55
Tool steel (T15)	67 HRC	18–82

contribution of the size effect, that is, the smaller the chip, the higher the specific energy required to produce it. Also, it has been observed that with effective lubrication, the specific energy in grinding can be reduced by a factor of four or more. Also, using engineered abrasives with controlled orientation, the specific energy can be reduced by a factor of three or more, as described in Case Study 26.1 at the end of the chapter.

Example 26.1 Forces in Surface Grinding

Given: Assume that a surface-grinding operation is being carried out on low-carbon steel, with a wheel of diameter $D = 250$ mm and rotating at $N = 4000$ rpm, and a width of cut of $w = 25$ mm. The depth of cut is $d = 0.050$ mm, and the feed rate of the workpiece, v, is 25 mm/s.

Find: Calculate the *grinding force* (the force tangential to the wheel), F_c, and the *thrust force* (the force normal to the workpiece surface), F_n, using specific-energy data.

Solution: The material-removal rate (MRR) is determined as

$$\text{MRR} = dwv = (0.05)(25)(25) = 31.25 \text{ mm}^3/\text{min}.$$

The power consumed is given by

$$\text{Power} = (u)(\text{MRR}),$$

where u is the specific energy, which can be obtained from Table 26.2 (see also Section 21.3). For low-carbon steel, it is estimated to be 110 W-s/mm^3. Thus,

$$\text{Power} = (110)(31.25) = 3440 \text{ W}.$$

Since power is defined as

$$\text{Power} = T\omega,$$

where the torque $T = F_c D/2$ and ω is the rotational speed of the wheel in radians per second ($\omega = 2\pi N/60$). It then follows that

$$3440 = (F_c)\left(\frac{0.25}{2}\right)(2\pi)(4000/60),$$

so that $F_c = 65.7$ N. The thrust force, F_n, can be calculated directly; however, it also can be estimated by noting from experimental data in the technical literature that it is about 30% higher than the cutting force, F_c. Consequently,

$$F_n = (1.3)(65.7) = 85.4 \text{ N}.$$

Temperature. The temperature rise in grinding is an important consideration because of the following reasons:

- It can adversely affect the surface properties of the workpiece, including metallurgical changes.
- The temperature rise can cause residual stresses in the workpiece.
- Temperature gradients in the workpiece cause distortions due to thermal expansion and contraction of the workpiece surface, thus making it difficult to control dimensional accuracy.

The *surface-temperature rise*, ΔT, in grinding is related to process variables by the following expression:

$$\Delta T \propto D^{1/4} d^{3/4} \left(\frac{V}{v}\right)^{1/2}. \tag{26.4}$$

Thus, temperature increases with increasing depth of cut, d, wheel diameter, D, and wheel speed, V, and decreases with increasing workpiece speed, v. Note from this equation that the depth of cut has the largest exponent; hence, it has the greatest influence on temperature.

Although *peak temperatures* during grinding can reach 1600°C, the time involved in producing a chip is on the order of microseconds; thus, the chip produced may or may not melt. Because the chips carry away much of the heat generated, as do chips formed in high-speed machining processes (Section 25.5), only a small fraction of the heat generated in grinding is conducted to the workpiece. If this was not the case, it would be very difficult to grind workpieces with sufficient dimensional accuracy and without causing any possible metallurgical changes to the workpiece.

Sparks. The sparks produced when grinding metals are actually chips that glow due to the *exothermic* (heat producing) reaction of the hot chips with oxygen in the atmosphere. Sparks do not occur during grinding in an oxygen-free environment or when the workpiece material does not readily oxidize at elevated temperatures. The color, intensity, and shape of sparks depend on the composition of the metal being ground. Charts are available that, from the appearance of its sparks, help identify the type of metal being ground. If the heat generated due to exothermic reaction is sufficiently high, chips can melt, acquiring and solidifying into spherical shape because of surface tension (see Fig. 26.12c).

Tempering. Excessive temperature rise in grinding can cause *tempering* and *softening* of the workpiece surface; processing variables must therefore be selected properly to avoid excessive temperature rise. Using grinding fluids (Section 26.4) is an effective means of controlling temperature.

Burning. Excessive temperature rise during grinding may burn the workpiece surface. A *burn* is characterized by a bluish color on ground steel surfaces, an indication that high temperatures have caused oxidation of the workpiece. A burn can be detected by etching and metallurgical techniques; it may not be objectionable in itself, unless surface layers have undergone *phase transformations* (Chapter 4). For example, if martensite forms in higher carbon steels from rapid cooling (called a **metallurgical burn**), it will adversely affect the surface properties of ground parts, and reduce surface ductility and toughness.

Heat Checking. High temperatures in grinding may also develop cracks in the workpiece surface, known as *heat checking*. The cracks usually are perpendicular to the grinding direction, although under severe conditions, parallel cracks also may appear. As expected, such a surface lacks toughness and has low fatigue and corrosion resistance. Heat checking also occurs in dies during die casting (see Section 11.4.5).

Residual Stresses. Temperature gradients within the workpiece during grinding are primarily responsible for the development of *residual stresses*. Grinding fluids and their method of application, as well as processing parameters such as depth of cut and speed, significantly influence the magnitude and type of residual stresses. Because of the adverse effect of tensile residual stresses on fatigue strength, processing variables should be selected accordingly. Residual stresses usually can be reduced by lowering wheel speed and increasing workpiece speed (called **low-stress grinding** or *gentle grinding*). Softer grade wheels, known as **free-cutting** grinding wheels, also may be used to reduce residual stresses.

26.3.1 Grinding-wheel Wear

Similar to the wear on cutting tools, grinding-wheel wear is an important consideration, because it adversely affects the shape and dimensional accuracy of ground surfaces. Wear of grinding wheels is caused by three different mechanisms, as described below.

Attritious Grain Wear. In *attritious wear*, which is similar to flank wear in cutting tools (see Fig. 21.15), the cutting edges of an originally sharp grain become dull and develop a *wear flat* (Fig. 26.9b). This type of wear involves both physical and chemical reactions, and is caused by the interaction of the grain material with the workpiece material. These complex reactions involve diffusion, chemical degradation or decomposition of the grain, fracture at a microscopic scale, plastic deformation, and melting.

Attritious wear is low when the two materials are *chemically inert* with respect to each other, much like what has been observed with cutting tools (Section 22.1). The more inert the materials, the lower is the tendency for reaction and adhesion to occur between the grain and the workpiece. Thus, for example, because aluminum oxide is relatively inert with respect to iron, its rate of attritious wear when used to grind steels is much lower than that of silicon carbide and diamond. By contrast, silicon carbide can dissolve in iron, and hence it is not suitable for grinding steels. Cubic boron nitride has a higher inertness with respect to steels, and hence it is suitable as an abrasive.

Grain Fracture. Because abrasive grains are brittle, their fracture characteristics in grinding are important. If the wear flat caused by attritious wear is excessive, the grain becomes dull and grinding becomes inefficient, and produces undesirably high temperatures. Ideally, a dull grain should fracture or fragment at a moderate rate, so that new sharp edges are produced continuously during grinding. This situation is equivalent to breaking a dull piece of chalk or a stone into two or more pieces in order to expose new sharp edges (see *friability* in Section 26.2).

The selection of grain type and size for a particular application also depends on the attritious wear rate. A grain–workpiece material combination that has a high attritious wear and low grain friability dulls the grains and develops a large wear flat; grinding then becomes inefficient, and surface damage and burning are likely to occur.

Bond Fracture. The strength of the bond (*grade*) is a significant parameter in grinding. If, for example, the bond is too strong, dull grains cannot be easily dislodged, preventing other sharp grains along the circumference of the wheel from contacting the workpiece. Conversely, if the bond is too weak, the grains are dislodged easily, and the wear rate of the wheel increases; maintaining dimensional accuracy then becomes difficult.

In general, softer bonds are recommended for harder materials, to reduce residual stresses and thermal damage to the workpiece. Hard-grade wheels are used for softer materials, for removing large amounts of material at high rates.

26.3.2 Grinding Ratio

The *grinding ratio*, G, correlates the grinding wheel wear with the amount of workpiece material removed, and is defined as

$$G = \frac{\text{Volume of material removed}}{\text{Volume of wheel wear}}. \tag{26.5}$$

In practice, G varies widely, ranging from 2 to 200, and it may even be higher, depending on the type of wheel, workpiece material, grinding fluid, and processing parameters, such as the depth of cut and the speeds of the wheel and the workpiece. It has also been shown that effective grinding fluids can increase G by a factor of 10 or more, greatly improving wheel life.

During grinding, a particular wheel may **act soft** (thus exhibiting high wear rate) or **act hard** (low wear rate), regardless of the wheel grade. Note, for example, that an ordinary pencil acts soft when writing on rough paper, but it acts hard when writing on soft paper, even though it is the same pencil. Acting hard or soft is a function of the force on the individual grain on the periphery of the wheel. The higher the force, the higher the tendency for the grains to fracture or to be dislodged from the wheel surface, and the higher the wheel wear and the lower the grinding ratio.

Note from Eq. (26.3), that the grain force (a) increases with the strength of the workpiece material, work speed, and depth of cut and (b) decreases with increasing wheel speed and wheel diameter. Note also that attempting to obtain a high grinding ratio in practice (to extend wheel life) isn't always desirable, because high ratios may indicate grain dulling and possible surface damage to the workpiece. A lower ratio may be acceptable when an overall technical and economic analysis justifies it.

Example 26.2 Action of a Grinding Wheel

Given: A surface-grinding operation is being carried out with the wheel running at a constant spindle speed. Assume that the depth of cut, d, remains constant and the wheel is dressed periodically (see Section 26.3.3).

Find: Will the wheel act soft or hard as the wheel wears down over time?

Solution: Referring to Eq. (26.3), note that the parameters that change over time in this operation are the wheel diameter, D, and the surface speed, V. As D becomes smaller, the relative grain force increases, thus the wheel acts softer. To accommodate the changes due to the wheel diameter reduction over time or to make provisions for using wheels of different diameters, some grinding machines are equipped with variable-speed spindle motors.

26.3.3 Dressing, Truing, and Shaping of Grinding Wheels

Dressing is the process of (a) *conditioning*, that is, producing *sharp new edges* on worn grains on the grinding surface of a wheel and (b) *truing*, producing a *true circle* on a wheel that, for whatever reason, has become out of round. Dressing is necessary when excessive attritious wear dulls the wheel, called **glazing** (because of the shiny appearance of the wheel surface), or when the wheel becomes *loaded* (see below). For softer wheels, truing and dressing are done separately, but for harder wheels, such as cBN, both are done in one operation.

Loading of a grinding wheel occurs when the porosities on the wheel surfaces (Fig. 26.10) become filled or clogged with chips from the workpiece. Loading can occur while grinding soft materials or from improper selection of wheels or processing parameters. A loaded wheel grinds inefficiently and generates much frictional heat, resulting in surface damage and loss of dimensional accuracy of the workpiece.

The techniques used to dress grinding wheels are the following:

- A specially shaped *diamond-point tool* or *diamond cluster* is moved across the width of the grinding face of a rotating wheel, removing a very thin layer from the wheel surface with each pass. This method can be performed either dry or wet, depending on whether the wheel is to be used dry or wet (using grinding fluids), respectively.
- A set of *star-shaped steel disks* is pressed against the wheel. Material is removed from the wheel surface by crushing the grains. This method produces a coarse surface on the wheel and is used only for rough grinding operations on bench or pedestal grinders.
- *Abrasive sticks* are used to dress grinding wheels, particularly softer wheels; however, this technique is not appropriate for precision grinding operations.
- Dressing techniques for metal-bonded diamond wheels involve the use of *electrical-discharge* and *electrochemical machining* techniques (Chapter 27). These processes erode very thin layers of the metal bond, exposing new diamond cutting edges.
- Dressing for form grinding involves *crush dressing* or *crush forming*, consisting of pressing a metal roll on the surface of the grinding wheel, which typically is a vitrified wheel. The roll, usually made of high-speed steel, tungsten carbide, or boron carbide, has a machined or ground profile on its periphery. Thus, it reproduces a replica of this profile on the surface of the grinding wheel being dressed (see Section 26.3.3).

Dressing techniques and their frequency are important for quality control, because they affect grinding forces and workpiece surface finish. Computer-controlled grinders are equipped with automatic dressing features, which dress the wheel simultaneously as grinding progresses. The first contact of the dressing tool with the grinding wheel is very important, as it determines the nature of the new surface produced. This action is monitored precisely, by using piezoelectric or acoustic-emission sensors (Section 37.7). Vibration sensors, power monitors, and strain gages also are used in the dressing setup of high-precision grinding machines.

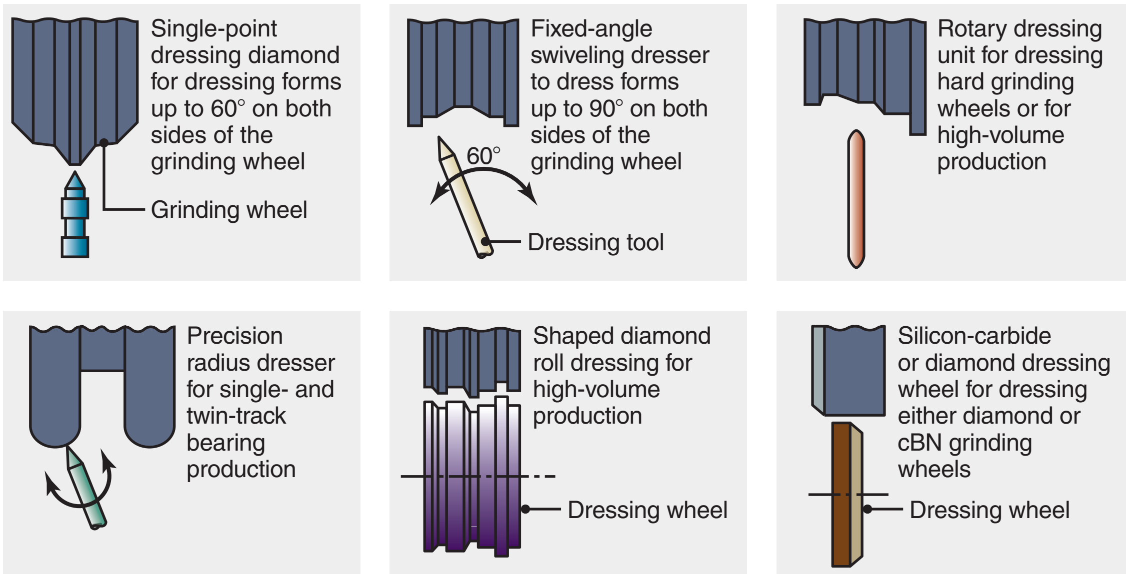
For a typical aluminum-oxide wheel, the depth removed during dressing is on the order of 5 to 15 μm, but for a cBN wheel, it could be 2 to 10 μm. Modern dressing systems have a resolution as low as 0.25 to 1 μm.

Grinding wheels can be *shaped* to the form to be ground on the workpiece (Section 26.4). The grinding face on the Type 1 straight wheel shown in Fig. 26.5a is cylindrical; thus, it produces a flat ground surface. The wheel surface also can be shaped into various forms by dressing it (Fig. 26.14a). Modern grinders are equipped with computer-controlled shaping features. Unless it already has the desired form, the diamond dressing tool traverses the wheel face automatically along a certain prescribed path (Fig. 26.14b), producing very accurate surfaces. Note in Fig. 26.14b that the axis of the diamond dressing tool remains normal to the grinding-wheel face at the point of contact.

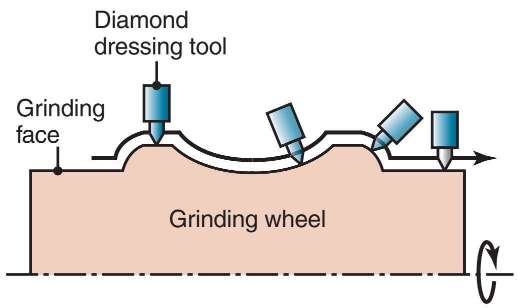
26.3.4 Grindability of Materials and Wheel Selection

The term *grindability* of materials, as in terms like *machinability* (Section 21.7) or *forgeability* (Section 14.5), is difficult to define precisely. It is a general indicator of how easy it is to grind a material, and includes such considerations as the quality of the surface produced, surface finish, surface integrity, wheel wear, grinding cycle time, and overall economics of the operation. Grindability of a material can be greatly enhanced by proper selection of processing parameters (Table 26.3), grinding wheels, grinding fluids, and by using the appropriate machine characteristics, fixturing methods, and workholding devices.

Grinding practices are well established for a wide variety of metallic and nonmetallic materials, including newly developed composites. Specific recommendations for selecting wheels and appropriate process



(a)



(b)

Figure 26.14: (a) Types of grinding-wheel dressing. (b) Shaping the grinding face of a wheel by dressing it by computer control. Note that the diamond dressing tool is normal to the surface at the point of contact with the wheel. *Source:* Courtesy of Okuma Machinery Works, Ltd.

parameters for metals can be found in various handbooks, manufacturers' literature, and the references in the Bibliography of this chapter.

Ductile-regime Grinding. It has been shown that with light passes and machine tools with high stiffness and damping capacity, it is possible to produce *continuous* chips and good surface finish in grinding of brittle materials, such as ceramics (Fig. 26.13), a process known as *ductile-regime grinding*. This regime produces fewer surface cracks and leads to better performance in fatigue and bearing applications. Ceramic chips,

Table 26.3: Typical Ranges of Speeds and Feeds for Abrasive Processes.

Process variable	Grinding, conventional	Grinding, creep-feed	Polishing	Buffing
Wheel speed (m/min)	1500–3000	1500–3000	1500–2400	1800–3500
Work speed (m/min)	10–60	0.1–1	—	—
Feed (mm/pass)	0.01–0.05	1–6	—	—

typically 1 to 10 μm in size, are more difficult to remove from grinding fluids than metal chips, requiring the use of fine filters and special techniques.

26.4 Grinding Operations and Machines

The selection of a grinding process and a machine tool for a particular application depends on the workpiece shape and features, size, ease of fixturing, and production rate required (Table 26.4). Modern grinding machines are computer controlled, and have such features as automatic workpiece loading and unloading, part clamping, and automatic dressing and wheel shaping. Grinders can be equipped with probes and gages, for determining the relative position of the wheel and workpiece surfaces (see also Fig. 25.6), as well as with tactile sensing features, whereby diamond dressing-tool breakage, for example, can be monitored readily during the dressing cycle.

Surface Grinding. *Surface grinding* (Fig. 26.15) generally involves grinding flat surfaces. In this operation, a straight wheel is mounted on the horizontal spindle of the surface grinder. In *traverse grinding*, the table reciprocates longitudinally and is fed laterally (in the direction of the spindle axis) after each stroke. The workpiece is held on a *magnetic chuck*, attached to the worktable of the grinder (Fig. 26.16); nonmagnetic materials are held by vises, vacuum chucks, or other fixtures.

The movement of the grinding wheel may be along the surface of the workpiece (*traverse grinding*, *through-feed* grinding, or *cross-feeding*), or the wheel may move radially into the workpiece (*plunge* grinding), as is the case when grinding a groove (Fig. 26.15b). Surface grinders make up the largest percentage of grinders used in industry, followed by bench grinders (typically with two wheels at each end of the spindle), cylindrical grinders, tool and cutter grinders, and internal grinders, as described below.

In addition to the surface grinder shown in Fig. 26.16, other types include *vertical spindle* and *rotary table* (referred to as the *Blanchard* type, Fig. 26.15c). These configurations allow several pieces to be ground in one setup. Steel balls for ball bearings, for example, are ground in special setups and at high production rates (Fig. 26.17).

Table 26.4: General Characteristics of Abrasive Machining Processes and Machines.

Process	Characteristics	Typical maximum dimensions, length and diameter (m)*
Surface grinding	Flat surfaces on most materials; production rate depends on table size and level of automation; labor skill depends on part complexity; production rate is high on vertical-spindle rotary-table machines	Reciprocating table L: 6 Rotary table D: 3
Cylindrical grinding	Round workpieces with stepped diameters; low production rate unless automated; low to medium labor skill	Workpiece D: 0.8, roll grinders D: 1.8, universal grinders D: 2.5
Centerless	Round and slender workpieces; high production rate; low to medium labor skill	Workpiece D: 0.8
Internal	Holes in workpiece; low production rate; low to medium labor skill	Hole D: 2
Honing	Holes in workpiece; low production rate; low labor skill	Spindle D: 1.2
Lapping	Flat, cylindrical, or curved workpieces; high production rate; low labor skill	Table D: 3.7
Chemical mechanical polishing	Flat surfaces, generally used for semiconductors for microelectronics or MEMS applications; moderate production rate; high labor skill.	D: 0.3
Abrasive flow machining	Used for debarring and finishing of complex geometries; low production rate; low labor skill	D: 0.3
Ultrasonic machining	Holes and cavities with various shapes; suitable for hard and brittle materials; medium labor skill	—

*Larger capacities are available for special applications.

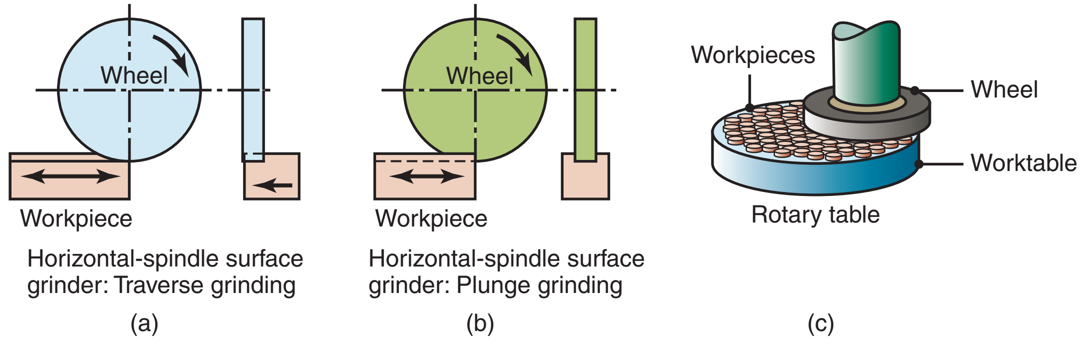


Figure 26.15: Schematic illustrations of various surface-grinding operations. (a) Traverse grinding with a horizontal-spindle surface grinder. (b) Plunge grinding with a horizontal-spindle surface grinder, producing a groove in the workpiece. (c) A vertical-spindle rotary-table grinder (also known as the *Blanchard* type).

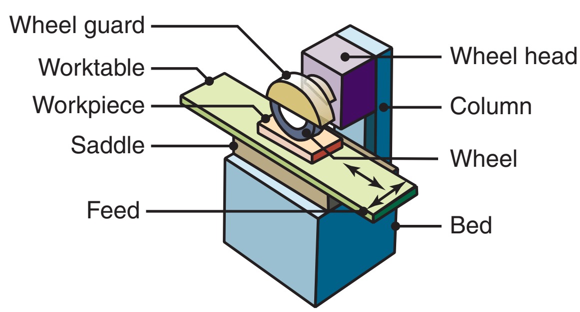


Figure 26.16: Schematic illustration of a horizontal-spindle surface grinder.

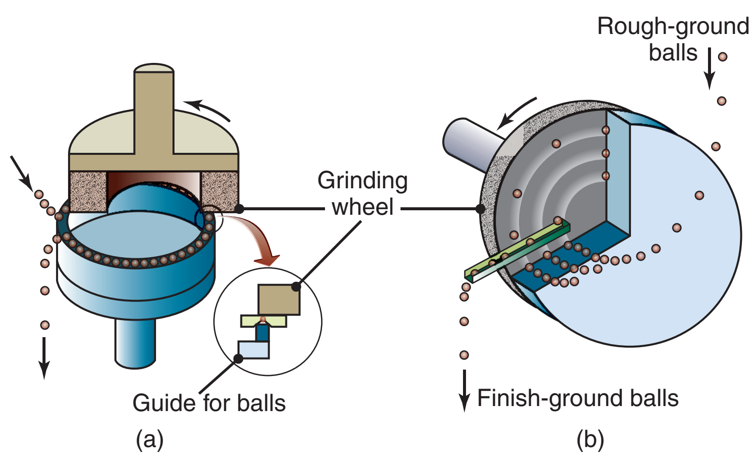


Figure 26.17: (a) Rough grinding of steel balls on a vertical-spindle grinder. The balls are guided by a special rotary fixture. (b) Finish grinding of balls in a multiple-groove fixture. The balls are ground to within 0.013 mm of their final size.

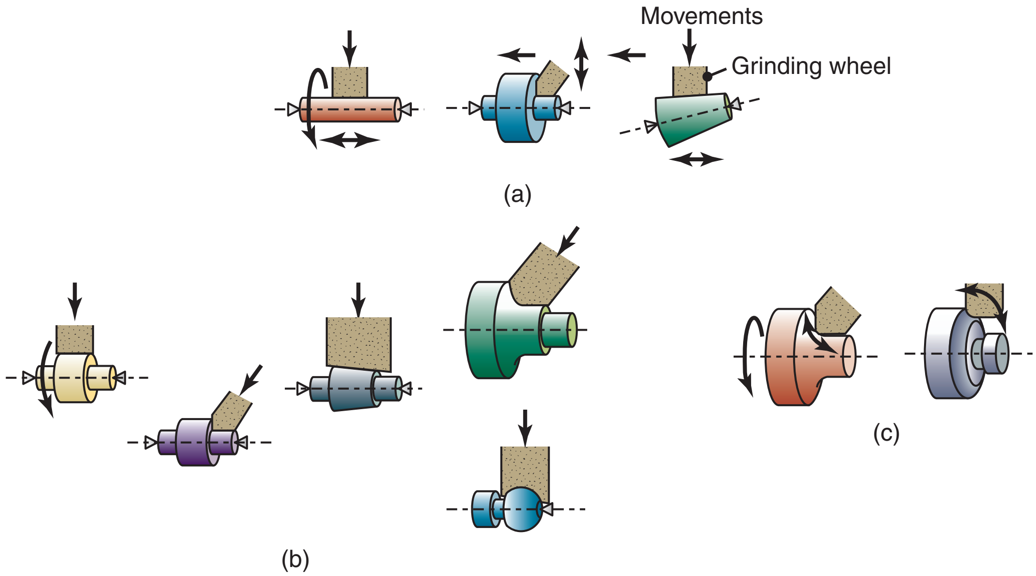


Figure 26.18: Examples of various cylindrical-grinding operations: (a) traverse grinding, (b) plunge grinding, and (c) profile grinding. *Source:* Courtesy of Okuma Machinery Works, Ltd.

Cylindrical Grinding. In *cylindrical grinding*, also called *center-type grinding* (Fig. 26.18; see also Fig. 26.3), the external cylindrical surfaces and shoulders of workpieces, such as crankshaft bearings, spindles, pins, and bearing rings, are ground. The rotating cylindrical workpiece reciprocates laterally along its axis, to cover the whole width to be ground. In *roll grinders*, used for large and long workpieces such as rolls for rolling mills (see Fig. 13.1), the grinding wheel reciprocates. These machine tools are capable of grinding rolls as large as 1.8 m in diameter.

The workpiece in cylindrical grinding is held between centers or in a chuck, or it is mounted on a *faceplate* in the headstock of the grinder. For straight cylindrical surfaces, the axes of rotation of the wheel and the workpiece are parallel, and each is driven by a separate motor and at different speeds. Long pieces with two or more diameters also can be ground on cylindrical grinders. As with *form grinding* and *plunge grinding*, the operation also can produce shapes in which the wheel is dressed to the workpiece form to be ground (Fig. 26.19).

Cylindrical grinders are identified by the maximum diameter and length of the workpiece that can be ground. In *universal grinders*, both the workpiece and the wheel axes can be moved and swiveled around a horizontal plane, thus permitting grinding tapers and various shapes.

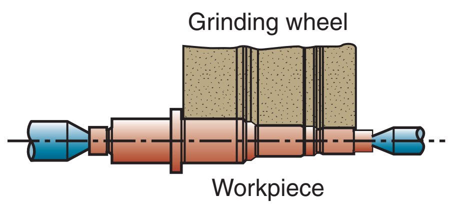


Figure 26.19: Plunge grinding of a workpiece on a cylindrical grinder with the wheel dressed to a stepped shape.

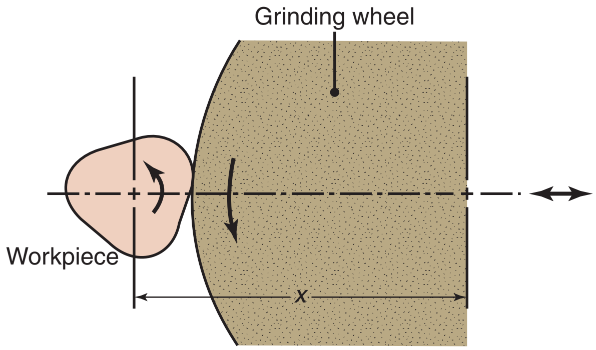


Figure 26.20: Schematic illustration of grinding a noncylindrical part on a cylindrical grinder with computer controls to produce the shape. The part rotation and the distance x between centers are varied and synchronized to grind the particular workpiece shape.

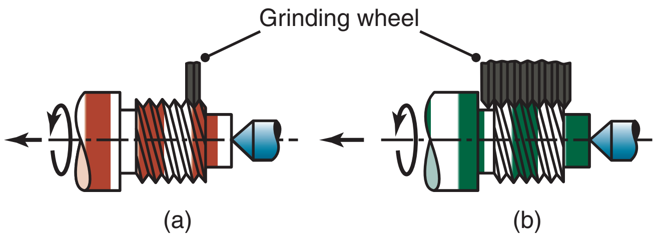


Figure 26.21: Thread grinding by (a) traverse and (b) plunge grinding.

With computer control, *noncylindrical* parts such as cams also can be ground. As illustrated in Fig. 26.20, the workpiece spindle speed is synchronized, such that the radial distance, x, between the workpiece and the wheel axes is continuously varied to grind a particular shape, such as the one shown in the figure.

Thread grinding is done on cylindrical grinders, using specially dressed wheels matching the shape of the threads (Fig. 26.21, see also *centerless grinding*). Although expensive, threads produced by grinding are the most accurate of any manufacturing process, and have very fine surface finish. Typical applications requiring such threads include ballscrew mechanisms, used for precise movement of various machine components. The workpiece and wheel movements are synchronized to produce the pitch of the thread, usually in about six passes.

Example 26.3 Cycle Patterns in Cylindrical Grinding

As in most grinding operations, the grinding wheel typically makes several passes along a path, in order to produce the final geometry on the workpiece. Figure 26.22 illustrates the cycle patterns for producing various shapes on a multifunctional, computer-controlled precision grinder. The downward arrowheads with numbers in the figures indicate the beginning of the grinding cycle.

Determination of the optimum and most economical pattern for minimum cycle time depends on the volume of material to be removed, the shape of the part, and the process parameters. All the patterns shown are automatically generated by the software in the computer controls of the grinder.

Source: Courtesy of Toyoda Machinery.

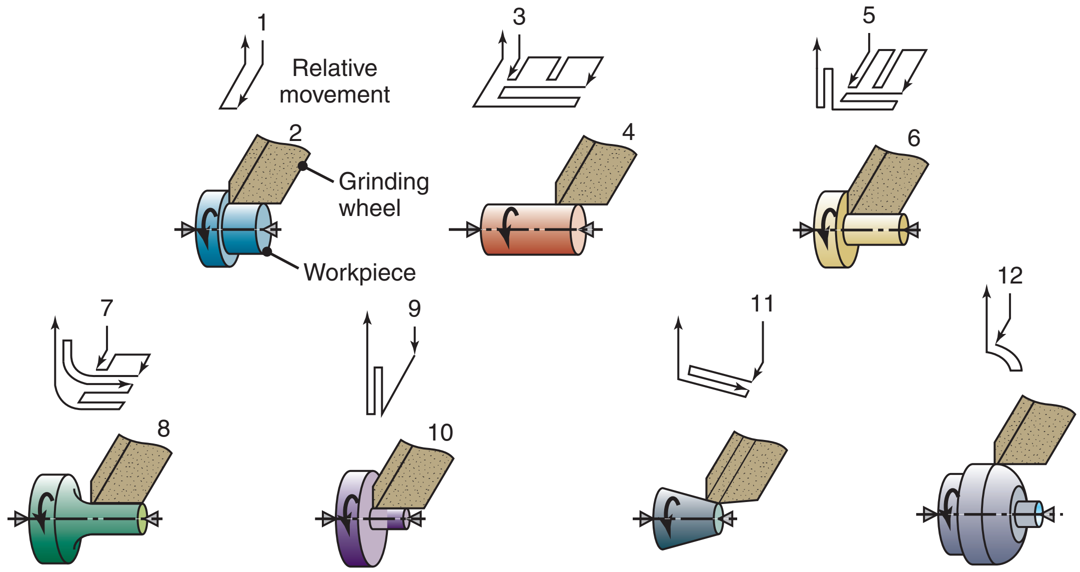


Figure 26.22: Cycle patterns for a CNC precision grinder.

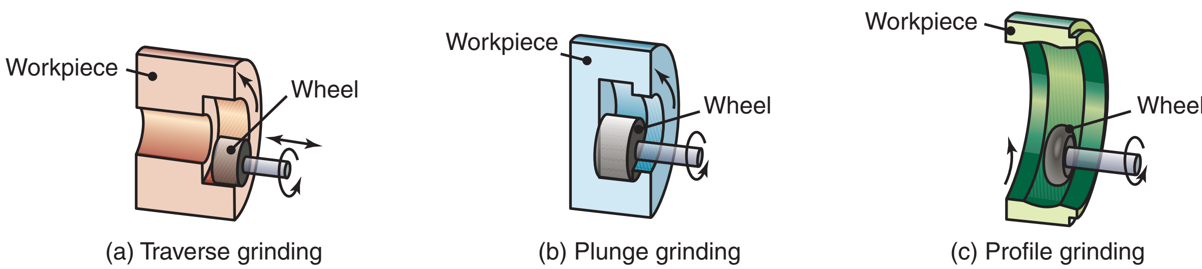


Figure 26.23: Schematic illustrations of internal grinding operations: (a) traverse grinding, (b) plunge grinding, and (c) profile grinding.

Internal Grinding. In *internal grinding* (Fig. 26.23), a small wheel is used to grind the inside diameter of the part, such as in bushings and bearing races. The workpiece is held in a rotating chuck; the wheel rotates at 30,000 rpm or higher. Internal profiles also can be ground with *profile-dressed* wheels, that move radially into the workpiece. The headstock of internal grinders can be swiveled on a horizontal plane for grinding tapered holes.

Centerless Grinding. *Centerless grinding* is a high-production process for grinding cylindrical surfaces. The workpiece is supported not by centers (hence the term centerless) or chucks, but by a *blade*, as shown in Fig. 26.24a and b. Typical parts ground are roller bearings, piston pins, engine valves, and camshafts; parts with diameters as small as 0.1 mm can be ground. Centerless grinders are capable of wheel surface speeds on the order of 10,000 m/min, typically using cubic boron nitride wheels.

In *through-feed grinding*, the workpiece is supported on a work-rest blade and is ground continuously (hence the term through-feed) between two wheels (Fig. 26.24a). Grinding is done by the larger wheel, while the smaller wheel regulates the axial movement of the workpiece. The rubber-bonded *regulating wheel* is tilted and runs at a much slower surface speed of about one-twentieth of the grinding-wheel speed.

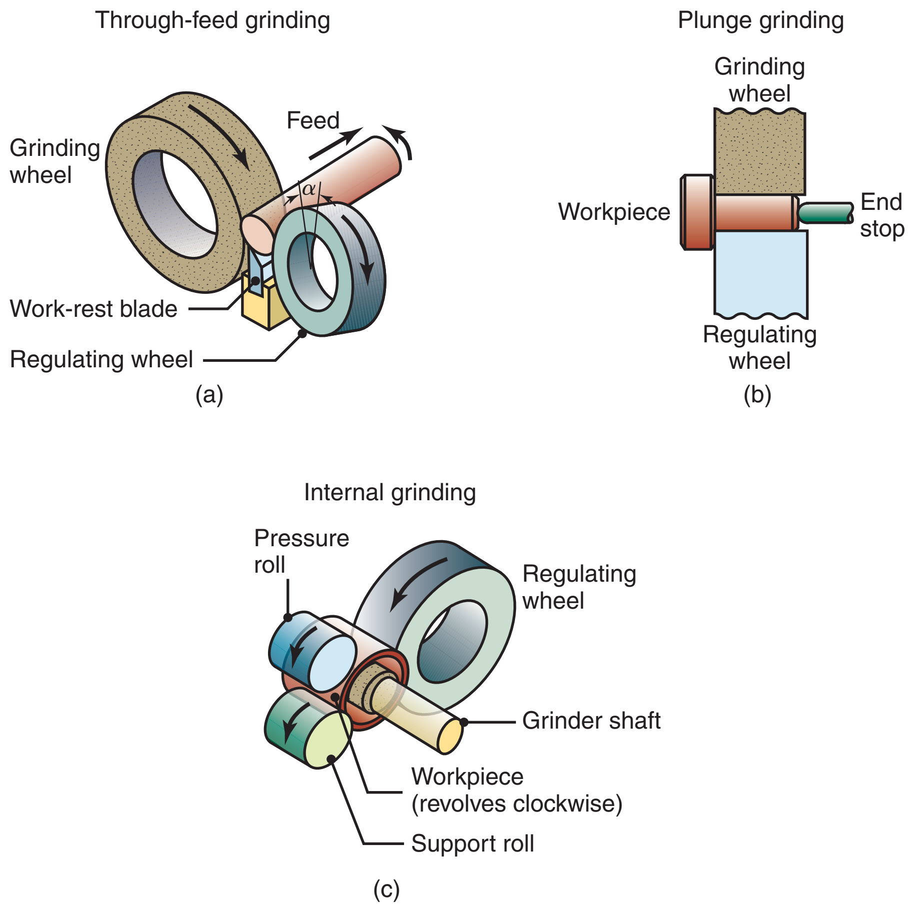


Figure 26.24: Schematic illustrations of centerless-grinding operations: (a) through-feed grinding, (b) plunge grinding, (c) and internal grinding.

Parts with variable diameters, such as bolts, valve tappets, and multiple-diameter shafts, can be ground by a process called *infeed* or *plunge grinding* (Fig. 26.24b), an operation similar to plunge or form grinding on cylindrical grinders. Tapered pieces are centerless ground by *end-feed grinding*. *Thread grinding* can be done at high-production rates with centerless grinders, using specially dressed wheels. In *internal centerless grinding*, the workpiece is supported between three rolls and is ground internally; typical applications are sleeve-shaped parts and rings (Fig. 26.24c).

Creep-feed Grinding. Although grinding traditionally has been associated with small rates of material removal (Table 26.3) and fine surface finishing operations, it can also be used for large-scale metal-removal operations. In *creep-feed grinding*, the wheel depth of cut, d, is as much as 6 mm and the workpiece speed is low (Fig. 26.25). The wheels are softer grade resin bonded and have an open structure (see Fig. 26.7), in order to keep workpiece temperatures low and to improve surface finish.

The machine tools for creep-feed grinding have special features, such as power up to 225 kW, high stiffness (because of the high forces due to the large depth of material removed), high damping capacity, variable spindle and worktable speeds, and ample capacity for the grinding fluids required. They can continuously dress the wheel, using a diamond roll as the dressing tool.

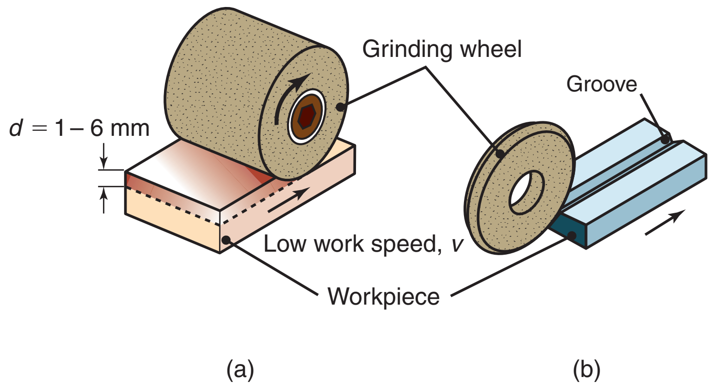


Figure 26.25: (a) Schematic illustration of the creep-feed grinding process; note the large wheel depth of cut, d. (b) A shaped groove produced on a flat surface by creep-feed grinding in one pass. Groove depth is typically on the order of a few mm. This operation also can be performed by some of the processes described in Chapter 27.

Creep-feed grinding can be competitive with other machining processes, such as milling, broaching, and planing. It is economical for such specific applications as shaped punches, key seats, twist-drill flutes, roots of turbine blades, and various complex superalloy parts. Because the wheel is dressed to the shape to be produced, the workpiece does not have to be shaped previously by milling, shaping, or broaching; near-net-shape castings and forgings are therefore suitable for creep-feed grinding. Although a single grinding pass generally is sufficient, a second pass may be necessary for improved surface finish.

Heavy Stock Removal by Grinding. Grinding can also be used for heavy stock removal by increasing process parameters, such as wheel depth of cut. This operation can be economical in certain specific applications and it can compete favorably with machining processes, particularly milling, turning, and broaching. In this operation, (a) surface finish is of secondary importance, (b) the dimensional tolerances are on the same order as those obtained by most machining processes, and (c) the grinding wheel or the belt can be utilized to its fullest capabilities, while minimizing grinding cost per piece. Heavy stock removal by grinding is also performed on welds, castings, and forgings to finish weld beads and remove flash.

Example 26.4 Grinding versus Hard Turning

In some specific applications, grinding and hard turning (Section 25.6) can be competitive. Hard turning continues to be increasingly competitive with grinding, and dimensional tolerances and surface finish are approaching those obtained by grinding. Consider the case of *machining* of heat-treated steels, with hardness above 45 HRC, using a single-point polycrystalline cubic boron nitride tool, versus *grinding* these steels.

In comparing Tables 21.2 and 26.2, it will be noted that (a) turning requires much less energy than grinding; (b) thermal and other types of damage to the workpiece surface are less likely to occur in machining; (c) cutting fluids may not be necessary; and (d) lathes are less expensive than grinders. Moreover, finishing operations, including finish grinding, can be performed on the turned part while it is still chucked in the lathe.

On the other hand, workholding devices for large and especially slender workpieces during hard turning can present significant problems, because cutting forces are higher than grinding forces. Furthermore, tool wear and its control can be a significant challenge as compared with the automatic dressing of grinding wheels. It is evident that the competitive positions of hard turning versus grinding must be evaluated individually for each application, in terms of product surface finish, integrity, quality, and overall economics.

A number of grinders are used for various operations:

- **Universal tool and cutter grinders** are used for grinding single-point or multipoint tools and cutters, including drills. They are equipped with special workholding devices for accurate positioning of the tools to be ground. A variety of CNC tool grinders is available, making the operation simple and fast and with consistent results.

- **Tool-post grinders** are self-contained units, usually attached to the tool post of a lathe (see Fig. 23.2). The tool is mounted on the headstock and is ground by moving the tool post. These grinders are versatile, but it is essential for the lathe components to be protected from the abrasive debris.

- **Swing-frame grinders** are typically used in foundries for grinding large castings. Rough grinding of castings is called **snagging**, and is usually done on *floorstand grinders*, using wheels as large as 0.9 m in diameter.

- **Portable grinders** are used for such operations as grinding off weld beads and *cutting off*, using thin abrasive disks. They are driven either pneumatically, electrically, or with a flexible shaft connected to an electric motor or a gasoline engine.

- **Bench** and **pedestal grinders** are used for routine grinding of tools and small parts. They usually are equipped with two grinding wheels, mounted on the two ends of the shaft of an electric motor; generally, one wheel is coarse for rough grinding and the other is fine for finish grinding.

Grinding Fluids. The functions of grinding fluids are similar to those of cutting fluids (Section 22.12). Although grinding and other abrasive removal processes can be performed dry, the use of a fluid is important because it:

- Reduces temperature rise in the workpiece

- Improves part surface finish and dimensional accuracy

- Improves the efficiency of the operation, by reducing wheel wear, reducing loading of the wheel, and lowering power consumption.

Grinding fluids typically are *water-based emulsions*, for general grinding, and *oils*, for thread grinding, (Table 26.5). They may be applied as a stream (flood) or as mist (a mixture of fluid and air). Because of the high surface speeds involved, an airstream (*air blanket*) around the periphery of the grinding wheel may prevent the fluid from reaching the wheel–workpiece interface. Special *nozzles* that conform to the shape of the cutting surface of the grinding wheel have been designed whereby the grinding fluid is supplied under high pressure.

There can be a significant rise in the temperature of water-based grinding fluids as they remove heat from the grinding zone, causing the workpiece to expand, thus making it difficult to control its dimensional accuracy. A common method to maintain a low workpiece temperature is to use refrigerating systems (*chillers*), through which the grinding fluid is circulated continuously and is maintained at about a constant

Table 26.5: General Recommendations for Grinding Fluids.

Material	Grinding fluid
Aluminum	E, EP
Copper	CSN, E, MO + FO
Magnesium	D, MO
Nickel	CSN, EP
Refractory metals	EP
Steels	CSN, E
Titanium	CSN, E

D = dry; E = emulsion; EP = extreme pressure; CSN = chemicals and synthetics; MO = mineral oil; FO = fatty oil (see also Section 33.7).

temperature. As described in Section 22.12, the *biological* and *ecological* aspects of disposal, treatment, and recycling of metalworking fluids are important considerations in their selection and use. The practices employed must comply with federal, state, and local laws and regulations.

Grinding Chatter. *Chatter* is particularly important in grinding as a finishing operation, because it can adversely affect surface finish and wheel performance. Studying **chatter marks** on ground surfaces often can help identify their source, which may include: (a) bearings and spindles of the grinding machine, (b) nonuniformities in the grinding wheel, as manufactured, (c) uneven wheel wear, (d) improper dressing techniques, (e) grinding wheels that are not balanced properly, and (f) external sources, such as nearby machinery. The grinding operation itself can cause *regenerative chatter*, as it does in machining (Section 25.4).

The important factors in controlling chatter in grinding are the stiffness of the machine tool, the stiffness of work-holding devices, and damping of the system. General guidelines include (a) using soft-grade grinding wheels, (b) dressing the wheel frequently, (c) changing dressing techniques, when necessary, (d) reducing the material-removal rate, and (e) supporting the workpiece rigidly.

Safety in Grinding Operations. Because grinding wheels are brittle and rotate at high speeds, they can fracture. Certain procedures must be followed in their handling, storage, and use. Failure to follow these procedures, and the instructions and warnings printed on individual wheel labels, may result in serious injury or fatality. Grinding wheels should be stored properly and protected from environmental extremes, such as temperature or humidity. They should be inspected visually for cracks and damage prior to installing them. Vitrified wheels should be tested prior to their use by *ringing* them (supporting them at the hole, tapping them gently, and listening to the sound). A damaged wheel will have a flat ring to it, similar to that of a cracked dinner plate.

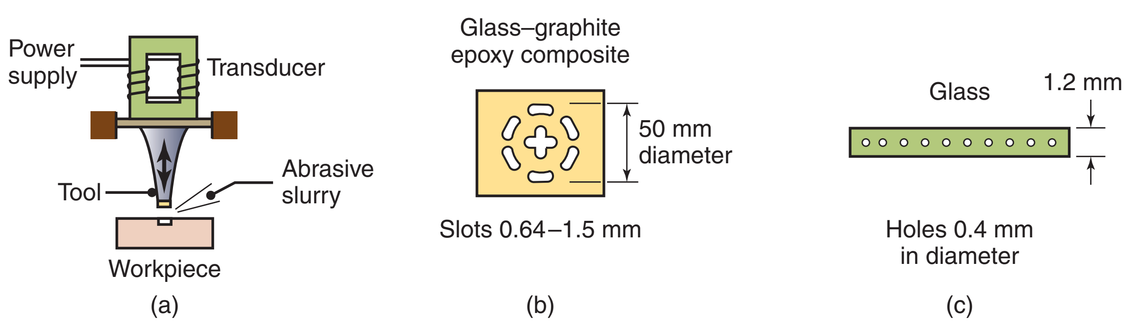


Figure 26.26: (a) Schematic illustration of the ultrasonic machining process. (b) and (c) Types of parts made by this process. Note the small size of the holes produced.

Damage to a grinding wheel can severely reduce its **bursting speed**. Defined as the surface speed at which a freely rotating wheel bursts (*explodes*), the bursting speed (expressed in rpm) depends on the type of wheel, such as its bond, grade, and structure. In diamond and cBN wheels (Fig. 26.6), which are operated at high surface speeds, the type of the core material used in the wheel affects the bursting speed. Metal cores, for example, have the highest bursting speed, typically on the order of about 250 m/s.

26.5 Design Considerations for Grinding

Design considerations for grinding are similar to those for machining, as described in various sections in Chapters 23 and 24. In addition, specific attention should be given in grinding to the following:

- Parts should be designed so that they can be mounted securely, either in chucks, magnetic tables, or suitable fixtures and workholding devices. Thin, straight, or tubular workpieces may distort during grinding, requiring special attention.
- If high dimensional accuracy is required, interrupted surfaces, such as holes and keyways, should be avoided, as they can cause vibrations and chatter.
- Parts for cylindrical grinding should be balanced; long and slender designs should be avoided to minimize deflections. Fillets and corner radii should be as large as possible, or relief should be provided for them during previous machining in these regions.
- In centerless grinding, short pieces may be difficult to grind accurately, because the blade may not support them sufficiently. In through-feed grinding, only the largest diameter on the parts can be ground.
- The design of parts requiring accurate form grinding should be kept as simple as possible, to avoid frequent form dressing of the wheel.
- Deep and small holes, and blind holes requiring internal grinding, should be avoided or they should include a relief.

In general, part designs should have a minimum amount of material to be removed by grinding, except for creep-feed grinding. Moreover, in order to maintain good dimensional accuracy, designs preferably should allow for all grinding to be done without having to reposition the workpiece.

26.6 Ultrasonic Machining

In *ultrasonic machining* (UM), material is removed from a surface by *microchipping* and *erosion*, with fine and loose abrasive grains in a water slurry (Fig. 26.26a). The tip of the tool (called a **sonotrode**) vibrates at a frequency of 20 kHz and an amplitude of 0.0125 to 0.075 mm. Vibration imparts high velocity to abrasive grains between the tool and the workpiece. The stress produced by the abrasive particles impacting the workpiece surface is high, because (a) the time of contact between the particle and the surface is on the order of only 10 to 100 μs and (b) the area of contact is very small. In brittle materials, impact stresses are sufficiently high to remove material from the workpiece surface.

The abrasive grains are typically boron carbide, although aluminum oxide or silicon carbide grains are also used, with sizes ranging from grit number 100 for roughing to grit number 1000 (see Fig. 26.7) for finishing operations. The grains are carried in water slurry, with concentrations of 20 to 60% by volume; the slurry also carries the debris away from the cutting zone.

Ultrasonic machining is best suited for materials that are hard and brittle, such as ceramics, carbides, precious stones, and hardened steels; two examples are shown in Fig. 26.26b. A special tool is required for each shape to be produced; thus it is also called a *form tool*. The tip of the tool, which is attached to a transducer through the toolholder, is usually made of mild steel.

Rotary Ultrasonic Machining. In this process, the abrasive slurry is replaced by a tool with metal-bonded diamond abrasives, either impregnated or electroplated on the tool surface. The tool is vibrated ultrasonically and rotated at the same time while being pressed against the workpiece surface at a constant pressure. The process is similar to face-milling (Fig. 24.5), but with the inserts being replaced with abrasives. The chips produced are washed away by a coolant, pumped through the core of the rotating tool. Rotary ultrasonic machining (RUM) is particularly effective in producing deep holes in brittle materials and at high material-removal rates.

Design Considerations for Ultrasonic Machining. The basic design guidelines for UM include the following:

- Avoid demanding profiles, sharp corners, and radii, because they can be eroded by the flow of the abrasive slurry.
- Holes produced will have some taper.
- Because of the tendency of brittle materials to chip at the exit end of holes, the bottom of the parts should have a backup plate.

26.7 Finishing Operations

Several other processes utilize fine abrasive grains and are used as a final finishing operation. Because these operations can significantly affect production time and product cost, they should be specified only after due consideration to their costs and benefits.

Coated Abrasives. Common examples of *coated abrasives* are sandpaper and emery cloth; the majority is made of aluminum oxide, with silicon carbide and zirconia alumina making up the rest. Coated abrasives usually have a much more open structure than grinding wheels, and their grains are more pointed and aggressive. The grains are deposited electrostatically on flexible backing materials, such as paper or cloth.

As shown in Fig. 26.27, the bonding material (matrix) typically is resin. It first is applied to the backing (called *make coat*); then the grains are bonded with a second layer (*size coat*). The grains have their long axes aligned perpendicular to the plane of the backing, thus improving their cutting action. Coated abrasives are available as sheets, belts, and disks. They are used extensively to finish flat or curved surfaces of metallic and nonmetallic parts, metallographic specimens, and also in woodworking.

Belt Grinding. Coated abrasives also are used as *belts* for high-rate material removal with good surface finish. *Belt grinding* is an important production process, and in some cases competes well with, and is preferred to, conventional grinding operations. Belts with grit numbers ranging from 16 to 1500 (see Figs. 22.7 and 22.8) are available. Belt speeds are in the range of 700 to 1800 m/min. Machines for abrasive-belt operations require proper belt support and rigid construction to minimize vibrations.

Conventional coated abrasives have randomly placed abrasives on their surface, and may consist of single or multiple layers of abrasives. An alternative surface is produced by **microreplication**, in which

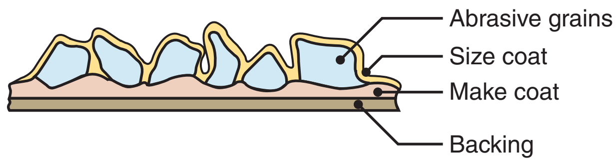


Figure 26.27: Schematic illustration of the structure of a coated abrasive. Sandpaper (developed in the 16th century) and emery cloth are common examples of coated abrasives.

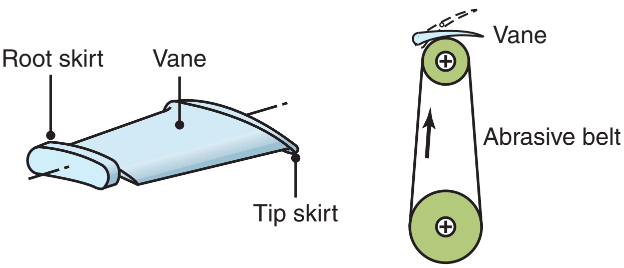


Figure 26.28: Turbine nozzle vane considered in Example 26.5.

abrasives in the shape of tiny aluminum-oxide pyramids are placed in a predetermined orderly arrangement on the belt surface. When used on stainless steels and superalloys, their performance is more consistent than conventional coated abrasives, and the temperature rise is lower. Typical applications include belt grinding of golf clubs, firearms, turbine blades, surgical implants, and various medical and dental instruments.

Example 26.5 Belt Grinding of Turbine Nozzle Vanes

The turbine nozzle vane shown in Fig. 26.28 was investment cast (Section 11.3.2) from a cobalt-based superalloy. To remove a thin diffusion layer from the root skirt and tip skirt sections of the vane, it was ground on a cloth-backed abrasive belt (60-grit aluminum oxide). The vanes were mounted on a fixture, and ground dry at a belt surface speed of 1800 m/min. The production rate was 93 seconds per piece. Each vane weighed 21.65 g before and 20.25 g after belt grinding, a reduction in weight of about 6.5%.

Source: Courtesy of ASM International.

Wire Brushing. In this process, also called *power brushing*, the workpiece is held against a circular wire brush that rotates at speeds from 1750 rpm for large wheels to 3500 rpm for small wheels. The tips of the wires produce longitudinal scratches on the workpiece surface. Performed under the proper conditions, wire brushing also may be considered as a very light material-removal process. Wire brushing is used to produce a fine or controlled surface texture. In addition to metal wires, polymeric wires (such as nylon; Section 7.6) embedded with abrasives can be used effectively (see also *diamond wire saws*; Section 24.5).

Honing. *Honing* is an operation used primarily to improve the surface finish of holes made by such processes as boring, drilling, and internal grinding. The honing tool consists of a set of aluminum oxide or silicon-carbide bonded abrasive sticks, called *stones* (Fig. 26.29). They are mounted on a mandrel that rotates

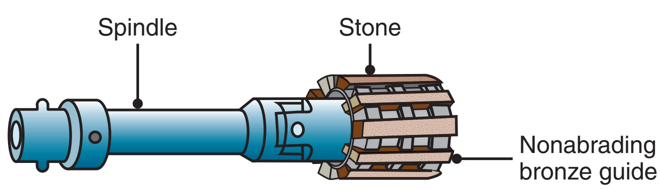


Figure 26.29: Schematic illustration of a honing tool used to improve the surface finish of bored or ground holes.

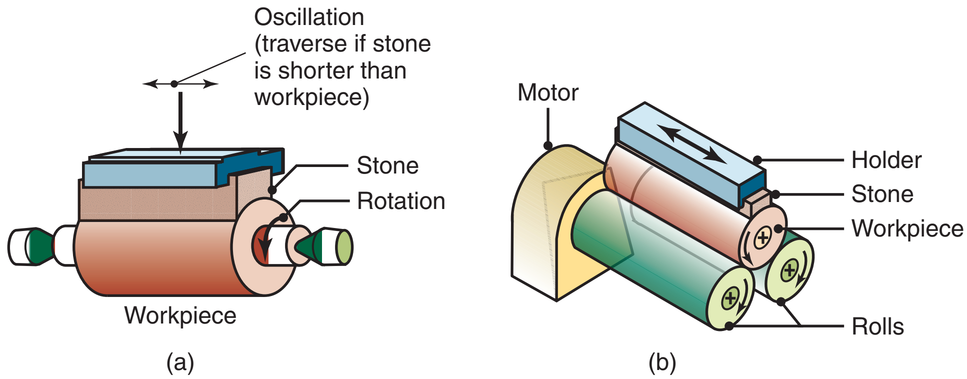


Figure 26.30: Schematic illustrations of the superfinishing process for a cylindrical part. (a) Cylindrical microhoning. (b) Centerless microhoning.

at surface speeds of 45 to 90 m/min in the hole, applying a radial outward force on the hole surface. They can be adjusted radially for different hole sizes.

The tool has a reciprocating axial motion, producing a crosshatched pattern on the hole surface. Oil- or water-based honing fluids are used to flush away the debris and keep temperatures low. Honing is also done on external or flat surfaces, and to manually remove sharp edges on cutting tools and inserts.

The quality of the surface finish produced by honing can be controlled by the type and size of the abrasive used, the pressure applied, and rotational speed. If not performed properly, honing can produce holes that are neither straight nor cylindrical, but rather in shapes that are bell mouthed, wavy, barrel shaped, or tapered.

Superfinishing. In this process, the pressure applied is very light and the motion of the honing stone has short strokes. The motion is controlled so that the grains do not travel along the same path on the workpiece surface. Examples of external superfinishing of a round part are shown in Fig. 26.30.

Lapping. This is an operation for finishing flat, cylindrical, or curved surfaces. Generally, the *lap* (Fig. 26.31a) is relatively soft and porous, and is made of such materials as cast iron, copper, leather, or cloth. The abrasive

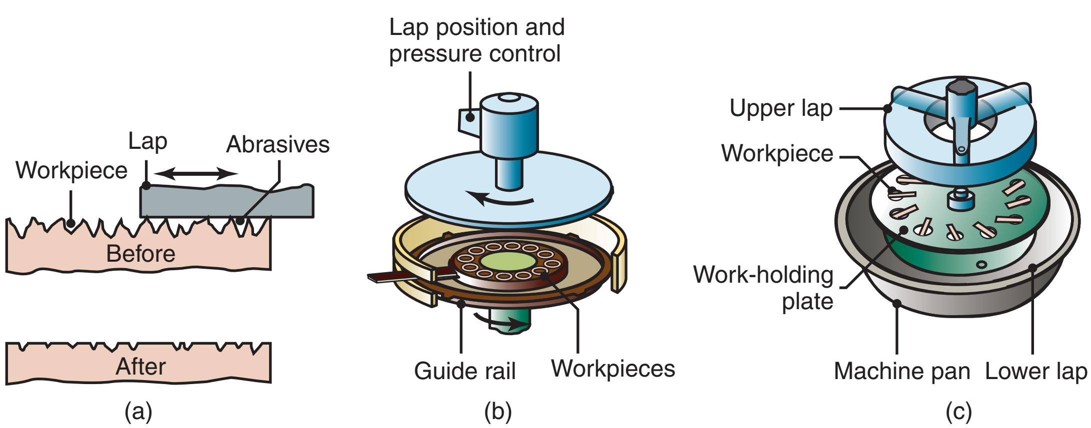


Figure 26.31: (a) Schematic illustration of the lapping process. (b) Production lapping on flat surfaces. (c) Production lapping on cylindrical surfaces.

particles either are embedded in the lap or may be carried in a slurry. Lapping of spherical objects and glass lenses is done with specially shaped laps. *Running-in* of mating gears can be done by lapping, as on hypoid gears for rear axles of automobiles. Depending on the type and hardness of the workpiece, lapping pressures range from 7 to 140 kPa.

Dimensional tolerances on the order of ± 0.0004 mm can be obtained in lapping by using fine abrasives (up to grit size 900), and the surface finish can be as smooth as 0.025 to 0.1 μm. Production lapping on flat or cylindrical parts is done on machines similar to those shown in Fig. 26.31b and c.

Polishing. *Polishing* is a process that produces a smooth, lustrous surface. The basic mechanism involved in the polishing process is the softening and smearing of surface layers by frictional heating developed during polishing, as well as by some very fine-scale abrasive removal from a workpiece. The shiny appearance commonly observed on polished surfaces results from a smearing action.

Polishing is done with disks or belts made of fabric, leather, or felt that typically are coated with fine powders of aluminum oxide or diamond. In *double-sided polishing*, pairs of pads are attached to the faces of platens that rotate in opposite directions. Parts with irregular shapes, sharp corners, deep recesses, and sharp projections can be difficult to polish.

Chemical–mechanical Polishing. *Chemical–mechanical polishing* (CMP) is extremely important in semiconductor manufacturing (Chapter 28). This process, shown in Fig. 26.32, uses a suspension of abrasive particles in a water-base solution, with a chemistry selected to cause controlled corrosion. Workpiece surface changes are through combined actions of abrasion and corrosion; the result is an exceptionally fine finish, as well as a very flat part. For this reason, the process is often referred to as **chemical–mechanical planarization** (Section 28.4).

A major application of this process is polishing of silicon wafers (Section 28.4), in which the primary function of CMP is to polish at the micrometer level without any *lay* (see Section 33.3). To remove material evenly and across the whole wafer surface, the wafer is held face down on a rotating carrier, and is pressed against a polishing pad attached to a rotating disk (Fig. 26.32). The angular velocities of the carrier and the pad are selected such that wear is uniform across the entire wafer surface. The velocities are adjusted such that there is a constant relative velocity between the carrier and the pad on the axis connecting their centers. The pad has grooves intended to uniformly supply slurry to all wafers; pad rotation ensures that a linear lay does not develop (see Section 33.3).

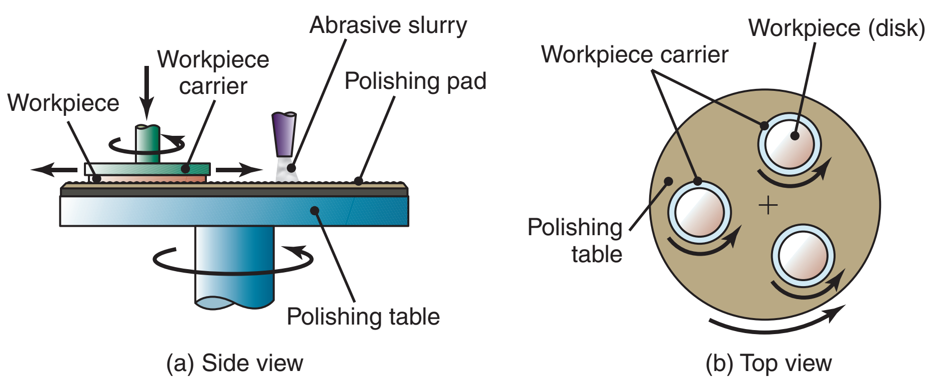


Figure 26.32: Schematic illustration of the chemical–mechanical polishing process. This process is used widely in making silicon wafers and integrated circuits and also is known as *chemical–mechanical planarization*. For other materials and applications, more carriers and more disks per carrier are possible.

Specific abrasive- and solution-chemistry combinations have been developed for polishing copper, silicon, silicon dioxide, aluminum, tungsten, and other metals. For silicon dioxide or silicon polishing, for example, an alkaline slurry of colloidal silica (SiO_2 particles in a KOH solution or in NH_4OH) is fed continuously to the pad-wafer interface.

Electropolishing. Mirrorlike finishes can be obtained on metal surfaces by electropolishing, a process that is the reverse of electroplating (Section 34.9). Because there is no mechanical contact with the workpiece, this process is particularly suitable for polishing irregular shapes as well. The electrolyte preferentially attacks the projections and the peaks on the workpiece, producing a smooth surface. Electropolishing is also used for deburring operations (Section 26.8).

Polishing in Magnetic Fields. In this technique, abrasive slurries are supported with magnetic fields. There are two basic methods:

1. In the **magnetic-float polishing** of ceramic balls, illustrated schematically in Fig. 26.33a, a magnetic fluid (containing abrasive grains and extremely fine ferromagnetic particles in a carrier fluid such as water or kerosene) is filled in the chamber within a guide ring. The ceramic balls are located between a driveshaft and a float. The abrasive grains, the ceramic balls, and the float (made of a nonmagnetic material) are all suspended by magnetic forces. The balls are pressed against the rotating driveshaft and are polished by the abrasive action. The forces applied by the abrasive particles on the balls are extremely small and are controllable, and hence the polishing action is very fine. Because polishing times are much lower than those involved in other polishing methods, this process is highly economical and the surfaces produced have few, if any, significant defects.

2. In the magnetic-field-assisted polishing of ceramic rollers (Fig. 26.33), a ceramic or steel roller (as the workpiece) is clamped and rotated on a spindle. The magnetic poles are then oscillated, introducing a vibratory motion to the magnetic–abrasive conglomerate, an action that polishes the cylindrical roller surface. Bearing steels, with a hardness of 63 HRC, have been mirror finished in 30 s with this process.

Buffing. This process is similar to polishing, with the exception that an even finer surface finish is obtained using very fine abrasives on soft disks, typically made of cloth.

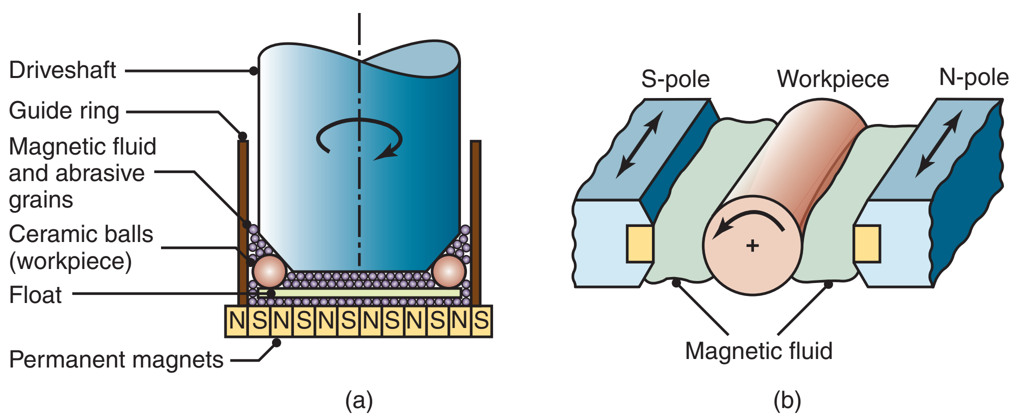


Figure 26.33: Schematic illustration of polishing of balls and rollers by magnetic fields. (a) Magnetic-float polishing of ceramic balls. (b) Magnetic-field-assisted polishing of rollers. *Source:* After R. Komanduri, M. Doc, and M. Fox.

26.8 Deburring Operations

Burrs are thin ridges, usually triangular in shape, that develop along the edges of a workpiece from such operations as machining, shearing sheet metals (see Figs. 16.2 and 16.3), and trimming of forgings and castings. Burrs can be detected by such simple means as by direct touch or with a toothpick or cotton swab. Visual inspection of burrs includes the use of magnifiers and microscopes.

Burrs can be detrimental: (a) They may interfere with the assembly of parts, and can cause jamming, misalignment of parts, and can cause short circuiting in electrical components. (b) Because they are usually sharp, they can be a safety hazard to personnel in handling parts. (c) Burrs may reduce the fatigue life of components. (d) Sheet metal may have lower bendability if the burr is on the tensile side (see Section 16.2). On the other hand, burrs can be useful on drilled or tapped thin components, such as tiny parts in mechanical watches and mechanisms, by providing additional thickness and thus improve the holding torque of screws.

Several **deburring processes** are available. Their cost-effectiveness depends on such factors as the extent of deburring required, part complexity, burr location, the number of parts to be deburred, floor space available, labor costs, and safety and environmental considerations. Deburring operations include:

1. Manual deburring, using files and scrapers; however, it is estimated that manual deburring can contribute up to 10% of the cost of manufacturing a part
2. Mechanical deburring by various means
3. Wire brushing or using rotary nylon brushes consisting of filaments embedded with abrasives
4. Abrasive belts
5. Ultrasonic machining
6. Electropolishing
7. Electrochemical machining
8. Magnetic–abrasive finishing
9. Vibratory finishing
10. Shot blasting or abrasive blasting
11. Abrasive-flow machining, such as extruding a semisolid abrasive slurry over the edges of the part
12. Thermal energy, using lasers or plasma.

The last four processes are described next; other processes are covered elsewhere in this book.

Vibratory and Barrel Finishing. These processes are used to remove burrs from large numbers of relatively small parts. This is a batch-type operation, in which specially shaped *abrasive pellets* of nonmetallic or metallic media (stones or balls) are placed in a container, along with the parts to be deburred. The container is then *vibrated* or *tumbled* by various means. The impact of individual abrasives and metal particles removes the burrs and sharp edges from the parts. Depending on the application, this can be a *dry* or a *wet* process. Liquid compounds may be added for such purposes as adding corrosion resistance to the parts being deburred. When chemically active fluids and abrasives are used, this process becomes a form of chemical–mechanical polishing (Section 26.7).

Shot Blasting. Also called **grit blasting**, this process involves abrasive particles (usually sand or specially engineered abrasives) propelled by a high-velocity jet of air, or by a rotating wheel, onto the surface of the part. Shot blasting is particularly useful in deburring metallic or nonmetallic materials, and in stripping, cleaning, and removing surface oxides. The surfaces have a matte finish, although surface damage can result if the process parameters are not controlled properly. **Microabrasive blasting** consists of small-scale polishing and etching, using very fine abrasives, on bench-type units.

Abrasive-flow Machining. This process involves using abrasive grains, such as silicon carbide or diamond, mixed in a puttylike matrix, and then forced back and forth through the openings and passageways in the part. The movement of the abrasive matrix under pressure erodes away burrs and sharp corners, polishing the part. Abrasive-flow machining (AFM) is particularly suitable for workpieces with internal cavities, such as those produced by casting, that are inaccessible by other means. The pressure applied ranges from 0.7 to 22 MPa. External surfaces also can be deburred with this method, by containing the workpiece within a fixture that directs the abrasive media to the edges and the areas of the part to be deburred. The deburring of a turbine impeller by this process is illustrated in Fig. 26.34.

In **microabrasive-flow machining**, the process mechanics are similar to those in abrasive-flow machining, but with much smaller abrasive media and using less viscous carriers. This technique allows the media to flow through very small holes, ranging from 50 μm to 750 μm diameter. Micro-AFM has been applied to the production of high-quality diesel-fuel injectors and other small nozzles, where a burr or rough surface finish could otherwise adversely affect the flow quality.

Thermal Energy Deburring. This process consists of placing the part in a chamber, which is then injected with a mixture of natural gas and oxygen. When the mixture is ignited, a burst of heat is produced, at a temperature of about 3300°C. The burrs are instantly heated and they melt while the temperature of the part only reaches about 150°C. There are, however, drawbacks to this process: (a) Larger burrs tend to form beads after melting, (b) thin and slender parts may distort, and (c) the process does not polish or buff the surfaces, as occurs in other deburring processes.

Robotic Deburring. Deburring and flash removal from finished products are being performed increasingly by *programmable robots* (Section 37.6). Using a force-feedback system for controlling the path and rate of burr removal, this method eliminates tedious and expensive manual labor, and it results in more consistent and repeatable deburring. In another application, the manual deburring of a double-helical gear for a helicopter gearbox was deburred in 150 min, whereas robotic deburring required 15 min.

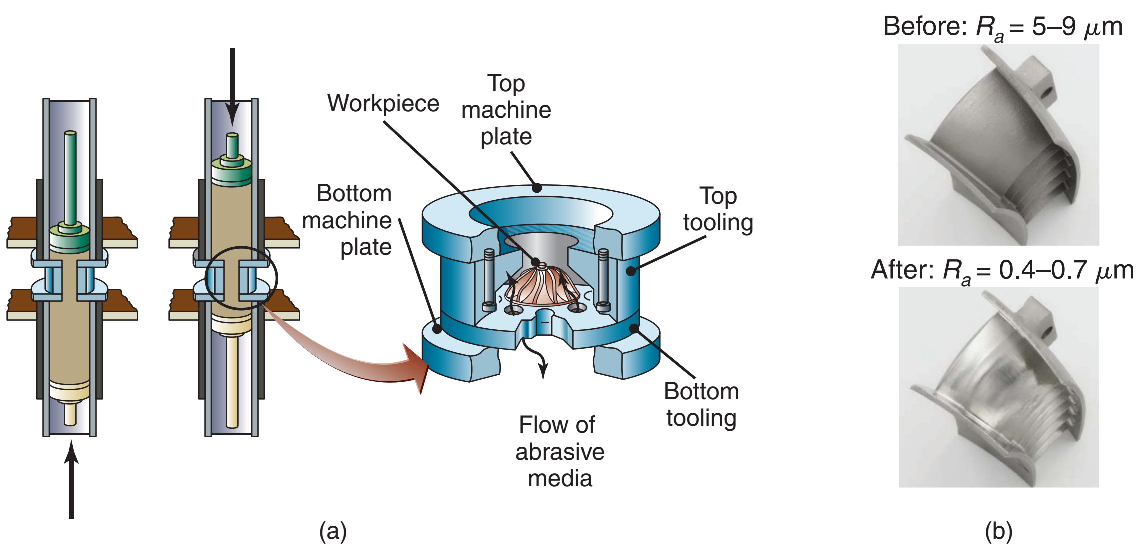


Figure 26.34: (a) Schematic illustration of abrasive-flow machining to deburr a turbine impeller. The arrows indicate movement of the abrasive media. Note the special fixture, which is usually different for each part design. (b) Valve fittings subjected to abrasive-flow machining to eliminate burrs and improve surface quality. *Source:* Courtesy of Kennametal Extrude Hone.

26.9 Economics of Abrasive Machining and Finishing Operations

Abrasive machining and finishing operations often are necessary, because forming, shaping, and machining processes alone do not achieve sufficiently high dimensional accuracy or surface finish. Abrasive processes may be used both as a finishing and as large-scale material-removal operation. For example, creep-feed grinding is an economical alternative to machining operations, such as milling or broaching, even though wheel wear is high.

Much progress has been made in automating the equipment involved in these operations, including the use of computer controls, sensors, process optimization, and robotic handling of parts. Labor costs and production times have been reduced, even though such machinery generally requires major capital investment.

Because they are additional operations, the processes described in this chapter can significantly affect product cost, especially since many of these processes are relatively slow. Moreover, as surface-finish requirements increase, more operations may be necessary, further increasing production costs, as clearly seen in Fig. 26.35. Note how rapidly the cost increases as surface finish is improved, by such additional processes as grinding and honing.

The total cost of abrasive operations depends on several factors, such as part size, shape, surface finish, and dimensional accuracy required, as well as machine tools, tooling, fixturing, and labor involved. Whereas machinery costs can be high for grinding, the costs for machinery for finishing processes are rather low. Grinding-wheel costs are generally low, although they can go up to hundreds or even thousands of dollars, depending on their composition and size. The costs of finishing tools, such as those for honing and lapping, vary widely, and labor costs and operator skill depend greatly on how well the equipment is automated.

If finishing is likely to be an important factor in manufacturing a part, the conceptual and original design stages should involve an analysis of the level of surface finish and the dimensional accuracy

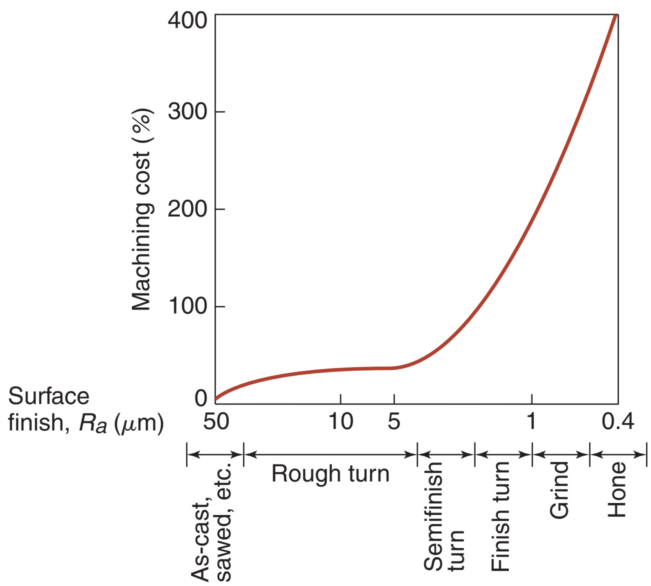


Figure 26.35: Increase in the cost of machining and finishing a part as a function of the surface finish required. This is the main reason that the surface finish specified on parts should not be any finer than is necessary for the part to function properly.

required, and whether they can be relaxed. Furthermore, all processes that precede finishing operations should be analyzed for their capability to produce more acceptable surface characteristics. This can be accomplished through proper selection of tools, selection of processing parameters, metalworking fluids, and the characteristics of the machine tools, their level of automation, computer controls, and of the workholding devices involved.

Case Study 26.1 Gear Grinding with Engineered Abrasives

Grinding in the gear industry has usually focused on finishing steel gear profiles, diameters, and bores; the gears are generally made of hardened steel. Gear manufacturers have to consider grinding in context, weighing the performance and economic factors of grinding as compared to hard machining processes, such as hard hobbing and power skiving (Section 24.2.4), or to various honing or lapping technologies (Section 26.7).

Grinding offers improved (a) surface quality, (b) dimensional accuracy, and (c) process reliability, with improvements in, for example, noise reduction and allowable contact stress. However, these are at the expense of reduced metal-removal rates, additional cost, and a somewhat higher environmental impact due mainly to the use of cutting fluid.

In general, little attention has been paid to the potential of grinding the rough gear form prior to heat treatment; conventional gear cutting is preferred due to the inability of traditional abrasive technologies to provide the required metal-removal rates and power efficiencies to be competitive.

Consider hobbing, where removal rates are on the order of 50–100 mm^3/sec; specific power requirements are around 4 J/mm^3. Hobbing is fast and energy efficient, due to the relatively small numbers of large chips produced.

The disadvantages of a process like hobbing lie in the flexibility and reliability of the tooling. The Indiana Tool and Manufacturing Company (ITAMCO) product line involves unique gears, both in terms of size and geometry. For these applications, the cutting tools, although normally long-lived are costly, are custom to a given gear, and they require long lead times to make. Machining performance and quality also change as the tool wears.

Grinding with conventional abrasives has historically been limited by low grinding ratios (Section 26.4), metal-removal rates, and high specific grinding energy requirements. Bonded cubic boron nitride superabrasive wheels also have been unable to achieve the high rates of material removal required, even at high wheel speeds; the wheels are also expensive. Moreover, like hobs, they are specific to a given gear tooth profile.

Grinding soft steel also has been a problem due to wheel loading, unless accompanied by continuous dressing. CDCF (continuous dress creep feed grinding) was developed in the 1970s using high porosity wheels to take deep form cuts. It was found that by continuously dressing with a formed diamond roll dresser, to keep the abrasive grains clean and sharp, the specific grinding energy was reduced significantly. In combination with good coolant access through high wheel porosity, CDCF allowed an order of magnitude increase in stock removal rates. However, although this was very effective on tough-to-grind metals, such as Inconel, the level of wheel wear from continuous dressing (typically 1μm/rev at 1000 rpm) resulted in uneconomic wear rates as compared to machining. Combined with the need for a specifically shaped diamond roll, again made CDCF impractical for rough grinding of gears.

Recent advances in wheel technology, especially those related to dressable vitrified bonds utilizing *engineered ceramic grains*, such as 3M's Cubitron II, make grinding feasible for these operations.

Conventional wheel grains have random shapes, typically with roughly equiaxed grains that offers a random cutting edge. The Cubitron II has specially shaped ceramic grains that are more aggressive and are uniformly oriented on a grinding wheel or disk. Typical grains are shown in Fig. 26.36. Compared to conventional abrasives (Fig. 26.1), the oriented grains grind more efficiently, allowing the grinding wheel to operate at lower temperature and higher efficiency, and removing up to three times the removal rate of conventional wheels.

As one of the largest open gear manufacturers in the world, ITAMCO has developed technologies that can achieve specific grinding energies and metal-removal rates that approach those of hobbing and shaping, but with a grinding ratio that allows deep form grinding without continuous dressing. This approach offers the possibility of grinding, in the soft state, on standard finish gear grinders. Dressing is done using a standard CNC contour diamond dress roll, instead of a specialty tool.

As an example of processing capabilities, the large gear (Fig. 26.37) has a 25 mm diametrical pitch, 600 mm face width, 3 m diameter with a 9 degree helix angle, and weighing over 23,587 kg. ITAMCO was able to finish the gear to size in 120 hours in one workholding, using only two 3M Cubitron II grinding wheels, thus saving time and material.

The immediate benefits of the approach described above are fast turnaround times, since there is no need to wait for specialty hob or shaper manufacture. Moreover, power requirements are significantly reduced, and more flexibility is achieved by using grinders that normally would be underpowered for demanding applications.

Source: Courtesy of J. Neidig, ITAMCO.

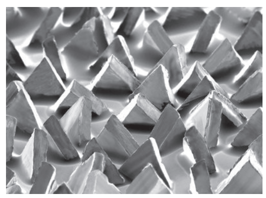


Figure 26.36: Cubitron II abrasives. The shaped particles are much more aggressive than conventional abrasives (Fig. 26.1). *Source:* Courtesy of J. Neidig, ITAMCO.

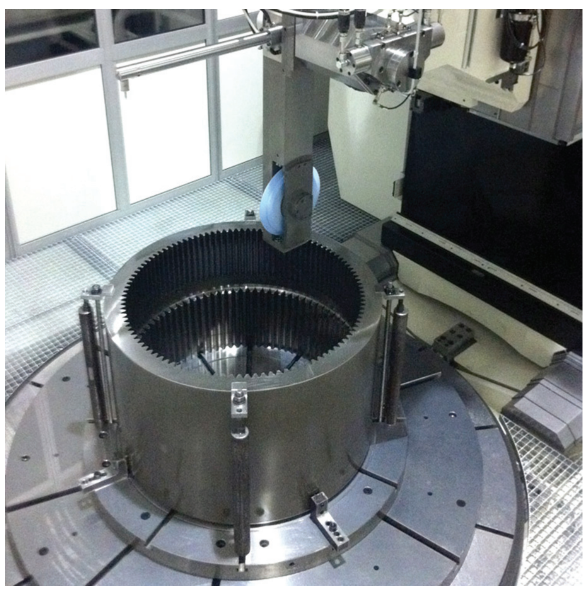


Figure 26.37: Grinding of a 3-m diameter gear.

Summary

- Abrasive machining often is necessary and economical when workpiece hardness and strength are high, the materials are brittle, and surface finish and dimensional tolerance requirements are demanding.
- Conventional abrasives consist of aluminum oxide and silicon carbide; superabrasives consist of cubic boron nitride and diamond. The friability of abrasive grains is an important factor in their performance, as are the shape and size of the grains.
- Grinding wheels, also known as bonded abrasives (in contrast to loose abrasives), consist of a combination of abrasive grains and bonding agents. Important characteristics of wheels are type of abrasive grain and bond, grade, and hardness. Wheels may be reinforced to maintain their integrity, if and when a crack develops during their normal use.
- Grinding wheel wear is an important consideration in the surface quality and integrity of the ground part. Dressing and truing of wheels are necessary operations, and are done by various techniques.
- A variety of abrasive-machining processes and machinery is available for surface, external, and internal grinding. The process is also used for large-scale material-removal processes, such as creep-feed grinding, making it competitive with processes such as milling and turning.
- The selection of abrasives and process variables, including grinding fluids, is important in obtaining the desired surface finish and dimensional accuracy; otherwise, damage to surfaces, such as burning, heat checking, detrimental residual stresses, and chatter may develop.
- Several finishing operations are available for improving surface finish. Because they can significantly affect product cost, the appropriate selection and implementation of these operations is important.
- Deburring may be necessary for some finished components. Commonly used methods are vibratory finishing, barrel finishing, and shot blasting, although thermal energy and other methods also are available.

Key Terms

Abrasive-flow machining
Abrasives
Aluminum oxide
Attritious wear
Barrel finishing
Belt grinding
Bonded abrasives
Buffing
Burning
Burr
Chatter marks
Chemical-mechanical polishing
Coated abrasives
Creep-feed grinding
Cubic boron nitride
Deburring
Diamond
Dressing
Ductile-regime grinding
Electropolishing
Engineered abrasive
Finishing
Free-cutting wheels
Friability
Glazing
Grade
Grain depth of cut
Grain size

Grindability
Grinding
Grinding ratio
Grit number
Hardness of wheel
Heat checking
Honing
Lapping
Loading
Low-stress grinding
Magnetic-field-assisted polishing
Magnetic-float polishing
Metallurgical burn
Microabrasive-flow machining
Microreplication
Polishing
Reinforced wheels
Resinoid bond
Robotic deburring
Rotary ultrasonic machining
Seeded gel
Shot blasting
Silicon carbide
Snagging
Sonotrode
Sparks
Specific energy
Structure of wheel
Superabrasives
Superfinishing
Tempering
Truing
Ultrasonic machining
Vibratory finishing
Vitrified bond
Wear flat
Wheel depth of cut
Wire brushing

Bibliography

Astashev, V.K., and Babitsky, V.I., **Ultrasonic Processes and Machines**, Springer, 2007.
Davim, J.P., (ed.), **Nontraditional Machining Processes**, Springer, 2013.
Doi, T., Uhlmann, E., and Marinescu, I.D., **Handbook of Ceramics Grinding and Polishing**, 2nd ed., William Andrew, 2015.
Gillespie, L.K., **Deburring and Edge Finishing Handbook**, Society of Manufacturing Engineers/American Society of Mechanical Engineers, 2000.
—, **Mass Finishing Handbook**, Industrial Press, 2006.
Grzesik, W., **Advanced Machining Processes of Metallic Materials**, 2nd. ed., Elsevier, 2016.
Hwa, L.S., **Chemical Mechanical Polishing in Silicon Processing**, Academic Press, 1999.
Jackson, M.J., and Davim, M.J., **Machining with Abrasives**, Springer, 2010.
Kuchle, A., **Manufacturing Processes 2: Grinding, Honing, Lapping**, Springer, 2009.
Malkin, S., and Guo, C., **Grinding Technology**, 2nd ed., Industrial Press, 2008.
Marinescu, I.D., Hitchiner, M., Uhlmann, E., Rowe, W.B., and Inasaki, I., **Handbook of Machining with Grinding Wheels**, 2nd ed., CRC Press, 2016.
Marinescu, I.D., and Rowe, W.B., **Handbook of Abrasive Machining Processes**, 2nd ed., William Andrew, 2012.
Oliver, M.R., **Chemical Mechanical Planarization of Semiconductor Materials**, Springer, 2004.
Rowe, W.B., **Principles of Modern Grinding Technology**, 2nd ed., William Andrew, 2013.

Review Questions

26.1. What is an abrasive? What are superabrasives?

26.2. What are the advantages of superabrasives over conventional abrasives?

26.3. How is the size of an abrasive grain related to its number?

26.4. Why are most abrasives made synthetically?

26.5. Describe the structure of a grinding wheel and its features.

26.6. Explain the characteristics of each type of bond used in bonded abrasives.

26.7. What causes grinding sparks in grinding? Is it useful to observe them? Explain.

26.8. Define *metallurgical burn*.

26.9. Define (a) friability, (b) wear flat, (c) grinding ratio, (d) truing, and (e) dressing.

26.10. What is creep-feed grinding and what are its advantages?

26.11. How is centerless grinding different from cylindrical grinding?

26.12. What are the differences between coated and bonded abrasives?

26.13. What is the purpose of the slurry in chemical mechanical polishing?

Qualitative Problems

26.14. Explain why grinding operations may be necessary for components that have previously been machined.

26.15. Why is there such a wide variety of types, shapes, and sizes of grinding wheels?

26.16. Explain the reasons for the large difference between the specific energies involved in machining (Table 21.2) and in grinding (Table 26.2).

26.17. Explain the factors involved in selecting the appropriate type of abrasive for a particular grinding operation.

26.18. Explain how the grinding ratio, G, depends on the following factors: (a) type of grinding wheel; (b) workpiece hardness; (c) wheel depth of cut; (d) wheel and workpiece speeds; and (e) type of grinding fluid.

26.19. What are the consequences of allowing the temperature to rise during grinding? Explain.

26.20. Explain why speeds are much higher in grinding than in machining operations.

26.21. It was stated that ultrasonic machining is best suited for hard and brittle materials. Explain.

26.22. Explain why parts with irregular shapes, sharp corners, deep recesses, and sharp projections can be difficult to polish.

26.23. Describe your understanding of the role of friability of abrasive grains on grinding-wheel performance.

26.24. List the finishing operations commonly used in manufacturing operations. Why are they necessary? Explain why they should be minimized.

26.25. Referring to the preceding chapters on processing of materials, list the operations in which burrs can develop on workpieces.

26.26. Explain the reasons that so many deburring operations have been developed over the years.

26.27. Outline the methods that are generally available for deburring parts. Discuss the advantages and limitations of each.

26.28. What precautions should you take when grinding with high precision? Comment on the machine, process parameters, grinding wheel, and grinding fluids.

26.29. Describe the factors involved in a grinding wheel acting "soft" or acting "hard." Can the same grinding wheel act soft or hard? Explain.

26.30. What factors could contribute to chatter in grinding? Explain.

26.31. Generally, it is recommended that, in grinding hardened steels, the grinding wheel be of a relatively soft grade. Explain.

26.32. In Fig. 26.5, the proper grinding faces are indicated for each type of wheel. Explain why the other surfaces of the wheels should not be used for grinding and what the consequences may be in doing so.

26.33. Describe the effects of a wear flat on the overall grinding operation.

26.34. What difficulties, if any, could you encounter in grinding thermoplastics? Thermosets? Ceramics?

26.35. Observe the cycle patterns shown in Fig. 26.22 and comment on why they follow those particular patterns.

26.36. Which of the processes described in this chapter are suitable particularly for workpieces made of (a) ceramics, (b) thermoplastics, (c) thermosets, (d) diamond, and (e) annealed aluminum? Why?

26.37. Grinding can produce a very fine surface finish on a workpiece. Is this finish necessarily an indication of the quality of a part? Explain.

26.38. Jewelry applications require the grinding of diamonds into desired shapes. How is this done, since diamond is the hardest material known?

26.39. Why should we be interested in the magnitude of the thrust force in grinding? Explain.

26.40. List and explain factors that contribute to poor surface finish in the processes described in this chapter.

Quantitative Problems

26.41. Calculate the chip dimensions in surface grinding for the following process variables: $D = 250$ mm, $d = 0.03$ mm, $v = 0.20$ m/s, $V = 30$ m/s, $C = 1$ per mm^2, and $r = 20$.

26.42. If the strength of the workpiece material is increased by 50%, what should be the percentage decrease in the wheel depth of cut, d, in order to maintain the same grain force, with all other variables being the same?

26.43. Assume that a surface-grinding operation is being carried out under the following conditions: $D = 300$ mm, $d = 0.15$ mm, $v = 0.6$ m/s, and $V = 60$ m/s. These conditions are then changed to the following: $D = 200$ mm, $d = 0.15$ mm, $v = 0.35$ m/s, and $V = 30$ m/s. What is the difference in the temperature rise from the initial condition?

26.44. Estimate the percent increase in the cost of the grinding operation if the specification for the surface finish of a part is changed from 6.4 to 0.8 μm.

26.45. Assume that the energy cost for grinding an aluminum part with a specific energy requirement of 8 W-s/mm^3 is \$1.50 per piece. What would be the energy cost of carrying out the same operation if the workpiece material were T15 tool steel?

26.46. In describing grinding processes, we have not given the type of equations regarding feeds, speeds, material-removal rates, total grinding time, etc., as we did in the turning and milling operations discussed in Chapters 23 and 24. Study the quantitative relationships involved and develop such equations for grinding operations.

26.47. What would be the answers to Example 26.1 if the workpiece is high-strength titanium and the width of cut is $w = 20$ mm? Give your answers in newtons.

26.48. It is known that heat checking occurs when grinding with a spindle speed of 4000 rpm, a wheel diameter of 250 mm, and a depth of cut of 0.0375 mm for a feed rate of 0.25 m/s. For this reason, the spindle speed should be kept at 3500 rpm. If a new, 200-mm-diameter wheel is used, what spindle speed can be employed before heat checking occurs? What spindle speed should be used to keep the same grinding temperatures as those encountered with the existing operating conditions?

26.49. A grinding operation is taking place with a 250-mm grinding wheel at a spindle rotational speed of 4000 rpm. The workpiece feed rate is 0.25 m/s, and the depth of cut is 0.050 mm. Contact thermometers record an approximate maximum temperature of 950°C. If the workpiece is steel, what is the temperature if the spindle speed is increased to 5000 rpm? What if it is increased to 10,000 rpm?

26.50. Derive an expression for the angular velocity of the wafer shown in Fig. 26.31b as a function of the radius and angular velocity of the pad in chemical–mechanical polishing.

26.51. It is desired to grind a hard aerospace aluminum alloy. A depth of 0.080 mm is to be removed from a cylindrical section 250-mm long and with a 80-mm diameter. If each part is to be ground in not more than one minute, what is the approximate power requirement for the grinder? What if the material is changed to a hard titanium alloy?

26.52. A 150-mm diameter tool steel ($u = 60$ W-s/mm^3) work roll for a metal rolling operation is being ground by a 250-mm diameter, 75-mm wide, Type I grinding wheel. Estimate the chip dimensions if $d = 0.04$ mm and $C = 5$ grains per mm^2. If the wheel rotates at $N = 3000$ rpm, estimate the cutting force if the work roll rotates at 1 rpm.

Synthesis, Design, and Projects

26.53. With appropriate sketches, describe the principles of various fixturing methods and devices that can be used for the processes described in this chapter.

26.54. Describe the methods you would use to determine the number of active cutting points per unit surface area on the periphery of a straight (i.e., Type 1; see Fig. 26.5a) grinding wheel. What is the significance of this number?

26.55. Make a comprehensive table of the process capabilities of abrasive-machining operations. Using several columns, describe the features of the machines involved, the type of abrasive tools used, the shapes of blanks and parts produced, typical maximum and minimum sizes, surface finish, tolerances, and production rates.

26.56. Vitrified grinding wheels (also called ceramic wheels) use a glasslike bond to hold the abrasive grains together. Given your understanding of ceramic-part manufacture (as described in Chapter 18), list methods of producing vitrified wheels.

26.57. Conduct a literature search, and explain how observing the color, brightness, and shape of sparks produced in grinding can be a useful guide to identifying the type of material being ground and its condition.

26.58. Visit a large hardware store and inspect the grinding wheels that are on display. Make a note of the markings on the wheels and, on the basis of the marking system shown in Fig. 26.7, comment on your observations, including the most common types of wheels available in the store.

26.59. Obtain a small grinding wheel or a piece of a large wheel. (a) Using a magnifier or a microscope, observe its surfaces and compare them with Fig. 26.9. (b) Rub the abrasive wheel by pressing it hard against a variety of flat metallic and nonmetallic materials. Describe your observations regarding the surfaces produced.

26.60. In reviewing the abrasive machining processes in this chapter, you will note that some use bonded abrasives while others involve loose abrasives. Make two separate lists for these processes and comment on your observations.

26.61. On the basis of the contents of this chapter, describe your thoughts on whether or not it would be possible to design and build a "grinding center" (see Chapter 25). Comment on any difficulties that may be encountered in such machines and operations.

26.62. Assume that you are an instructor covering the topics described in this chapter and you are giving a quiz on the numerical aspects to test the understanding of the students. Prepare three quantitative problems and supply the answers.

Chapter 27

Advanced Machining Processes

- It is often necessary to machine or finish products made of very hard or strong materials where conventional machining and grinding strategies described thus far become impractical or uneconomical. This chapter describes advanced machining processes that are based on nonmechanical methods of material removal.
- The chapter begins by examining chemical machining and photochemical blanking processes, in which material is removed through the corrosive action of a fluid.
- Electrochemical machining and grinding are then described, where material is removed by the action of an electrical power source and ion transfer inside an electrolytic fluid. Electrical-discharge machining removes material by melting small portions of the workpiece by a spark.

- Laser-beam and electron-beam machining processes, as well as water-jet and abrasive-jet machining operations, also are described, with examples of their unique applications.
- The chapter ends with a review of trends in hybrid machining operations and the economics of advanced machining processes.

Typical parts made: Skin panels for missiles and aircraft, turbine blades, nozzles, parts with complex cavities and small-diameter deep holes, dies, laser cutting of sheet metals, cutting of thick metallic and nonmetallic parts.

Alternative methods: Abrasive machining, ultrasonic machining, and precision machining.

27.1 Introduction

The machining processes described in the preceding chapters involve material removal by mechanical means of chip formation, abrasion, or microchipping. However, there are situations where mechanical methods are not satisfactory, economical, or even possible, for the following reasons:

- The **strength** and **hardness** of the workpiece material are very high, typically above 400 HB (Fig. 2.15).
- The material is too **brittle** to be machined without damage to the part, which typically is the case with highly heat-treated alloys, glass, ceramics, and powder-metallurgy parts.
- The workpiece is **too flexible** or **slender** to withstand the forces involved in machining or grinding, or the parts are difficult to clamp in fixtures and workholding devices.
- The part has a **complex shape** (Fig. 27.1), with such features as internal and external profiles or holes with high length-to-diameter ratios in very hard materials.
- The part has special **surface finish** and **dimensional tolerance** requirements that cannot be obtained by other processes or are uneconomical to do so.
- The **temperature rise** during processing and **residual stresses** developed in the workpiece are not acceptable.

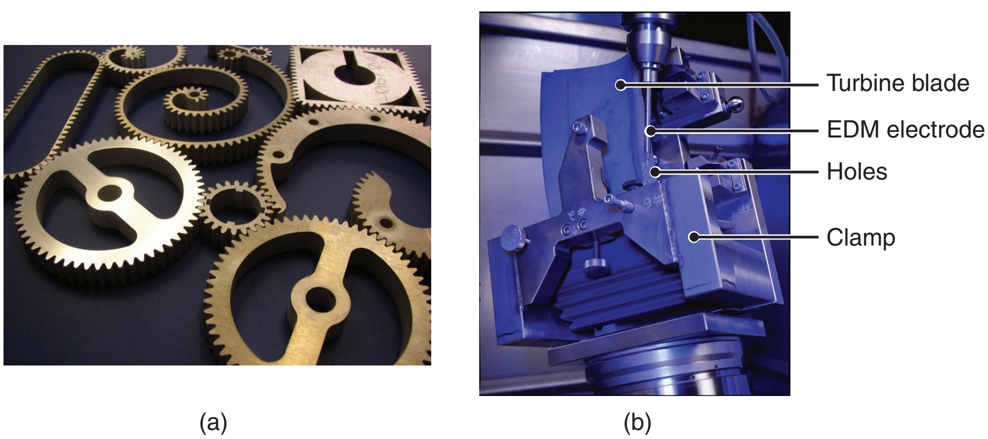


Figure 27.1: Examples of parts made by advanced machining processes. (a) Samples of parts produced by water-jet cutting. (b) Turbine blade, produced by plunge electrical-discharge machining, in a fixture to produce forced air cooling channels also by electrical-discharge machining. *Source:* (a) Courtesy of OMAX Corporation; (b) Courtesy of HI-TEK Mfg., Inc.

Table 27.1: General Characteristics of Advanced Machining Processes.

Process	Characteristics	Process parameters and typical material-removal rate or cutting speed
Chemical machining (CM)	Shallow removal on large flat or curved surfaces; blanking of thin sheets; low tooling and equipment cost; suitable for low-production runs	0.0025–0.1 mm/min.
Electrochemical machining (ECM)	Complex shapes with deep cavities; highest rate of material removal among other nontraditional processes; expensive tooling and equipment; high power consumption; medium-to-high production quantity	V: 5–25 D.C.; A: 1.5-8 A/mm^2; 2.5–12 mm/min, depending on current density
Electrochemical grinding (ECG)	Cutting off and sharpening hard materials, such as tungsten-carbide tools; also used as a honing process; higher removal rate than grinding	A: 1–3 A/mm^2; typically 25 mm^3/s per 1000 A
Electrical-discharge machining (EDM)	Shaping and cutting complex parts made of hard materials; some surface damage may result; also used as a grinding and cutting process; expensive tooling and equipment	V: 50–380; A: 0.1–500; typically 300 mm^3/min
Wire electrical-discharge machining	Contour cutting of flat or curved surfaces; expensive equipment	Varies with material and thickness
Laser-beam machining (LBM)	Cutting and hole making on thin materials; heat-affected zone; does not require a vacuum; expensive equipment; consumes much energy	0.50–7.5 m/min
Laser microjet	Water-jet guided laser uses a 25–100 μm diameter stream to mill or cut; large depth of field; little thermal damage from laser machining	Varies with material; up to 20 mm in silicon, 2 mm in stainless steel; up to 300 mm/s in 50 μm thick silicon.
Electron-beam machining (EBM)	Cutting and hole making on thin materials; very small holes and slots; heat-affected zone; requires a vacuum; expensive equipment	1–2 mm^3/min
Water-jet machining (WJM)	Cutting all types of nonmetallic materials; suitable for contour cutting of flexible materials; no thermal damage; noisy	Varies considerably with material
Abrasive water-jet machining (AWJM)	Single-layer or multilayer cutting of metallic and nonmetallic materials	Up to 7.5 m/min
Abrasive-jet machining (AJM)	Cutting, slotting, deburring, etching, and cleaning of metallic and nonmetallic materials; tends to round off sharp edges; can be hazardous	Varies considerably with material

Beginning in the 1950s, these difficulties led to the development of chemical, electrical, laser, and high-energy beams as energy sources for removing material from metallic and nonmetallic workpieces (Table 27.1). Also called *nontraditional* or *unconventional machining*, these processes remove material not by producing chips, as in traditional machining and grinding, but by means such as chemical dissolution, etching, melting, evaporation, and hydrodynamic action, at times with the assistance of fine abrasive particles.

A major advantage of these processes is that their efficiency is independent of workpiece hardness. When selected and applied properly, advanced machining processes offer major technical and economic advantages over more traditional methods. This chapter describes these processes, including their characteristics, typical applications, limitations, product quality, dimensional accuracy, surface finish, and economics.

27.2 Chemical Machining

Chemical machining (CM) is based on the fact that chemicals attack and etch most materials, thereby removing small amounts of material from workpiece surfaces. The CM process is carried out by *chemical dissolution* using **reagents** or **etchants**, such as acids and alkaline solutions. Developed in the 1950s, chemical machin-

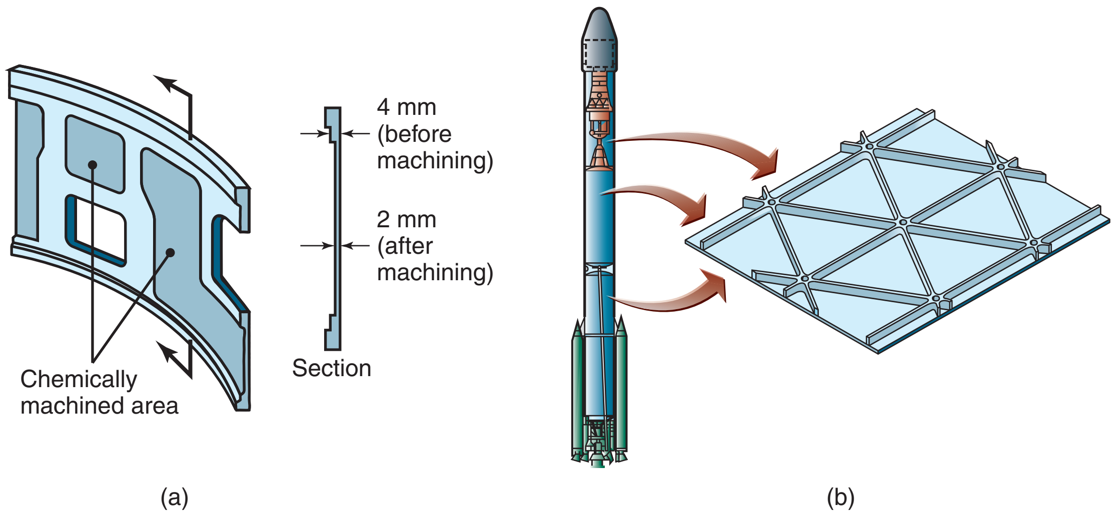


Figure 27.2: (a) Missile skin-panel section contoured by chemical milling to improve the stiffness-to-weight ratio of the part. (b) Weight reduction of space-launch vehicles by the chemical milling of aluminum-alloy plates. The plates are chemically milled after they have been formed into shape by a process such as roll forming or stretch forming. The design of the chemically machined rib patterns can readily be modified at minimal cost.

ing is the oldest of the advanced machining processes, and has been used in engraving metals and stones, in deburring, and in the production of printed-circuit boards and microelectronic devices (Chapters 28 and 29).

Chemical Milling. In *chemical milling*, shallow cavities are produced on plates, sheets, forgings, and extrusions, generally for overall reduction of weight, as can be seen in Fig. 27.2. The process has been used on a wide variety of metals, with depths of removal up to 12 mm. *Selective attack* by a chemical reagent on different areas of the workpiece surfaces is accomplished by removing layers of material from areas that are not **masked** (Fig. 27.3a). Material removal may also be done by *partial immersion* of the part in a reagent.

The procedure for chemical milling consists of the following steps:

1. If the part to be machined has residual stresses left from prior processing, the stresses should first be relieved (Section 4.11) in order to prevent warping after chemical milling (see also Case Study 24.3).

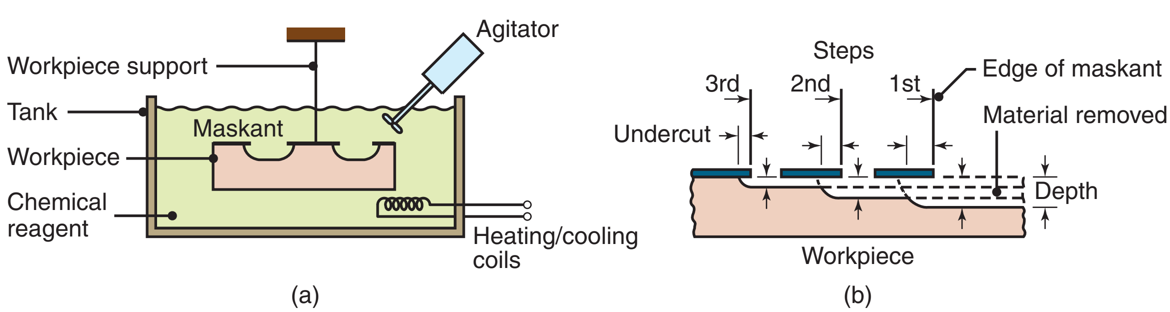


Figure 27.3: (a) Schematic illustration of the chemical-machining process; note that no forces or machine tools are involved in this process. (b) Stages in producing a profiled cavity by chemical machining; note the undercut.

2. The surfaces are thoroughly degreased and cleaned, to ensure both good adhesion of the masking material and maintaining uniform rate of material removal. Scale from prior heat treatment also should be removed.

3. The masking material (**maskant**) is applied, typically using tapes or paints, but also elastomers (rubber and neoprene) and plastics (polyvinyl chloride, polyethylene, and polystyrene). It is essential that the maskant not react with the chemical reagent.

4. Areas in the maskant that will require etching are peeled off, using the scribe-and-peel technique.

5. The exposed surfaces are machined chemically, using etchants such as sodium hydroxide (for aluminum), solutions of hydrochloric and nitric acids (for steels), and iron chloride (for stainless steels). Temperature control and agitation (stirring) of the etchant during milling is important, in order to remove a uniform depth of material from the part surfaces.

6. The parts are then washed thoroughly with water, to prevent further reactions with or exposure to any remaining etchant residues.

7. The rest of the masking material is removed, and the part is cleaned and inspected. Note that although the maskant is unaffected by the reagent, it can easily be dissolved by a different and appropriate type of solvent, such as acetone or Piranha (see Table 28.3).

8. Additional finishing operations may be performed on the milled parts, such as abrasive flow machining (Section 26.9) or electroplating (Section 34.9).

9. This sequence of operations can be repeated to also produce stepped cavities and various contours on parts (Fig. 27.3b).

Chemical milling is used in the aerospace industry to remove shallow layers of material from large aircraft components, missile skin panels (Fig. 27.2), and extruded parts for airframes. Tank capacities for reagents are as large as 3.7 m $\times$ 15 m. The process is also used to fabricate microelectronic devices, and is often referred to as **wet etching**, as described in Section 28.8.1. The ranges of surface finish and tolerances obtained by chemical machining and other machining processes are given in Fig. 33.5.

Because of **preferential etching** and **intergranular attack**, some surface damage may result from chemical milling, adversely affecting surface properties. Chemical milling of welded and brazed structures also may result in uneven material removal, and castings may result in uneven surfaces, caused by porosity and property nonuniformities in the material.

Chemical Blanking. This process is similar to blanking of sheet metals (Fig. 16.4). Typical applications are burr-free etching of printed circuit boards (Section 28.13), decorative panels, and thin sheet-metal stampings, as well as the production of complex or very small parts.

Photochemical Blanking. Also called *photoetching* or *photochemical machining*, this is a modification of the chemical milling process. Material is removed, usually from flat thin sheet, by photographic techniques to first produce a mask, followed by chemical machining. Complex, burr-free shapes can be blanked on metal foil as thin as 0.0025 mm. This process is also used for etching, such as for electrical connectors or pattern plates for reflow or paste soldering (Section 32.3).

The procedure in photochemical blanking consists of the following steps:

1. The design of the part to be blanked is prepared at a magnification of up to 100$\times$. A photographic negative is then made and reduced to the size of the finished part, called **artwork**; note that the original (enlarged) drawing allows inherent design errors to be reduced in size by the amount of reduction, such as 100$\times$, for the final artwork image.

2. The sheet blank is coated with a photosensitive material (**photoresist**, and often called *emulsion*), by dipping, spraying, spin casting, or roller coating; it is then dried in an oven.

3. The negative is placed over the coated blank and is exposed to ultraviolet light, which hardens the exposed areas.

4. The blank is developed, dissolving the unexposed mask areas; it is then immersed into a bath of reagent (as in chemical milling), or is sprayed with the reagent, which etches away the exposed metal areas.

5. The masking material is removed, and the part is thoroughly washed with water, to remove all chemical residues.

Handling of chemical reagents requires precautions and special safety considerations to protect the workers against exposure to both liquid chemicals and volatile chemicals. Furthermore, the disposal of chemical by-products from this process is a major drawback, although some by-products can be recycled.

Although skilled labor is required, tooling costs are low, the process can be automated, and is economical for medium- to high-production volume. Photochemical blanking is capable of making very small parts in cases when traditional blanking dies (Section 16.2) are too difficult to make. The process is also effective for blanking fragile workpieces and materials; tolerances are on the order of 10% of the sheet thickness. Typical applications for photochemical blanking include fine metal screens, printed-circuit boards, electric-motor laminations, flat springs, and various components of miniaturized systems.

Design Considerations for Chemical Machining. General design guidelines for chemical machining are:

- Designs with sharp corners, deep and narrow cavities, severe tapers, folded seams, or porous part materials should be avoided, because the etchant continuously attacks all exposed surfaces.

- Because the etchant attacks the material in both vertical and horizontal directions, **undercuts** may develop, as shown in Fig. 27.3 by the areas under the edges of the maskant.

- To improve production rate, the bulk of the workpiece preferably should be first shaped by other and higher volume rate processes, such as machining, prior to chemical machining.

- Because of size changes in the deposited mask pattern due to humidity and temperature, dimensional variations can occur. These variations can be minimized by properly selecting artwork media and by controlling both the environment in which the artwork is generated and in the production area in the plant.

- Product designs are produced with computer-aided design systems (Chapter 38), and they can be translated into a useful format for etching machinery.

27.3 Electrochemical Machining

Electrochemical machining (ECM) is basically the reverse of electroplating (Section 34.9). An **electrolyte** acts as the current carrier (Fig. 27.4), and the high flow rate of electrolyte in the tool-workpiece gap (typically 0.1 to 0.6 mm) washes metal ions away from the workpiece (*anode*) before they have a chance to plate onto the tool (*cathode*). Note that the cavity produced is the mating image of the tool shape.

The tool, either in solid or tubular shape, is generally made of brass, copper, bronze, or stainless steel. The electrolyte is a highly conductive inorganic fluid, such as an aqueous solution of sodium nitrate, and is pumped through the passages in the tool at rates of 10 to 16 m/s. A DC supply in the range from 10 to 25 V maintains current densities, which, for most applications, are 20 to 200 A/cm^2 of active machined surface.

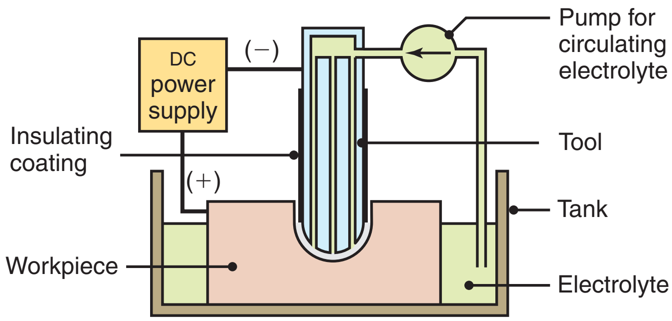


Figure 27.4: Schematic illustration of the electrochemical machining process.

For a current efficiency of 100%, the *material-removal rate* (MRR) in electrochemical machining may be estimated from

$$\text{MRR} = CI \tag{27.1}$$

where MRR is in mm^3/min, I is the current in amperes, and C is a material constant with the unit of $\text{mm}^3/\text{A-min}$. For pure metals, C depends on the valence: the higher the valence, the lower is the value of C.

Machines having current capacities as high as 40,000 A and as low as 5 A are available. The penetration rate of the tool is proportional to the current density, and the material removal rate typically ranges between 1.5 and 4 mm^3 per A-min. Because the metal-removal rate is a function only of the ion exchange rate, it is not affected by the strength, hardness, or toughness of the workpiece, a characteristic that is common to the processes described in this chapter.

Process Capabilities. The basic concept of electrochemical machining developed rapidly beginning with the 1950s, whereupon it became an important manufacturing process. It is generally used to machine complex cavities and shapes in high-strength materials, particularly in the aerospace industry for the mass production of turbine blades, jet-engine parts, and nozzles (Fig. 27.5); other applications include the automotive (engines castings and gears) and medical industries.

Electrochemical machining also is used for machining and finishing forging-die cavities (*die sinking*) and to produce small holes. Modifications of this process are used for turning, facing, milling, slotting, drilling, trepanning, and profiling operations, and in the production of continuous metal strips and webs. More recent applications of ECM include *micromachining* (Chapters 28 and 29) for the electronics industry.

An advance in ECM is *shaped-tube electrolytic machining* (STEM), and used for producing small-diameter deep holes, as in turbine blades (Fig. 27.6). The electrolyte is acid-based, to ensure that the worn metal is dissolved and carried away by the solution. The tool is a titanium tube for corrosion resistance, coated with an electrically-insulating resin to restrict the electrolytic action to the front surface of the electrode. Holes as small as 0.5 mm can be made, and at depth-to-diameter ratios as high as 300:1. Larger holes can be produced by *electrolytic trepanning*, as shown in Fig. 27.6b.

The ECM process leaves a burr-free, bright surface; it can also be used as a deburring operation. The operation does not cause any thermal damage to the part, and the absence of tool forces prevents distortion, especially in thin, flexible parts. Furthermore, there is no tool wear, since only hydrogen is generated at the cathode, and the process is capable of producing complex shapes. However, the mechanical properties of components made by ECM should be compared with those of components made by other processes, to ensure that there has not been a significant compromise due to chemical reactions.

Electrochemical-machining systems are available as *numerically controlled machining centers*, with capability of high production rates, high flexibility of operation, and the maintenance of fine dimensional tolerances. The ECM process also can be combined with electrical-discharge machining (EDM) on the same machine, called **hybrid machining** (see Section 27.10).

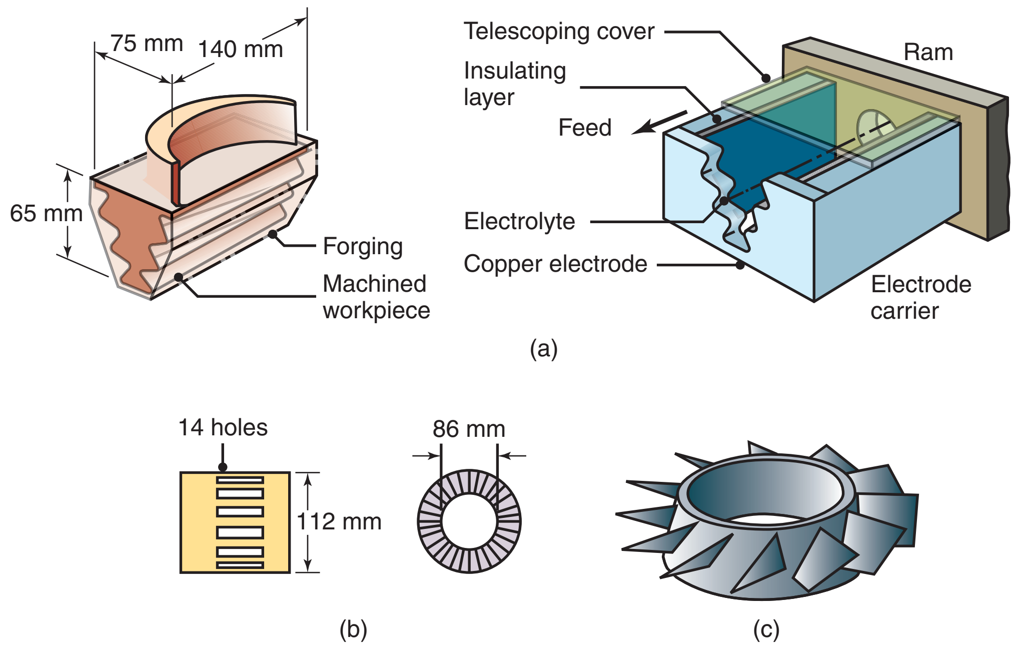


Figure 27.5: Typical parts made by electrochemical machining. (a) Turbine blade made of a nickel alloy of 360 HB; note the shape of the electrode on the right. (b) Thin slots on a 4340-steel roller-bearing cage. (c) Integral airfoils on a compressor disk.

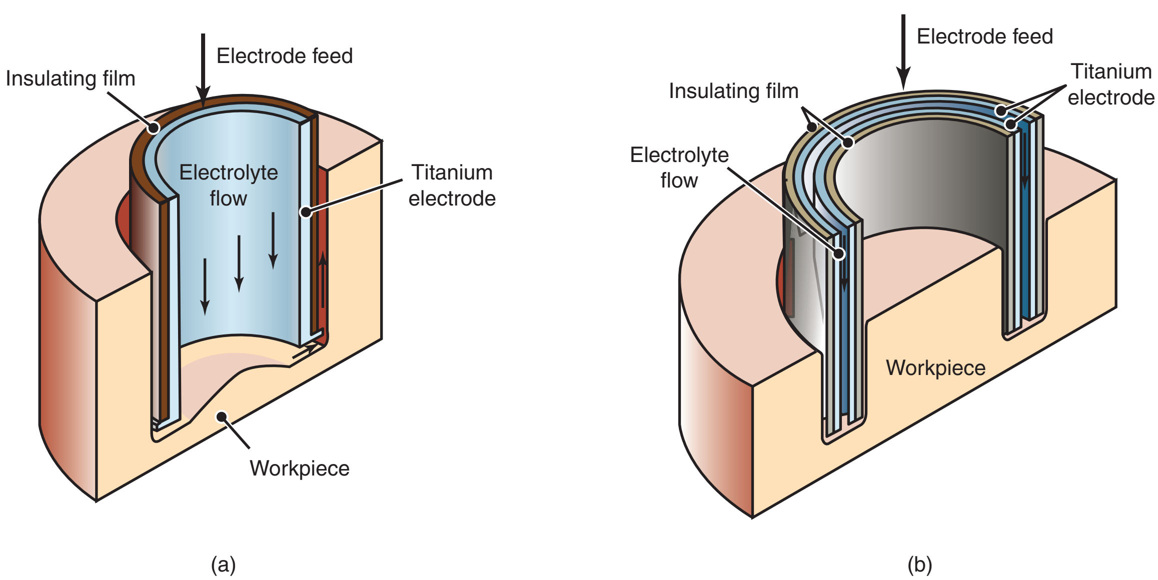


Figure 27.6: Shaped electrolytic machining operations. (a) Shaped-tube electrolytic machining, used to make small holes with aspect ratios as large as 300:1; (b) electrolytic trepanning, used for larger diameter holes.

Design Considerations for Electrochemical Machining. The following are general design guidelines for electrochemical machining:

- Part designs should make provision for a small taper for holes and cavities.
- Sharp square corners or flat bottoms cannot be produced because of the tendency for the electrolyte to erode away sharp profiles.
- Irregular cavities may not be produced to the desired shape and with acceptable dimensional accuracy, because of the difficulty in controlling electrolyte flow.

27.3.1 Pulsed Electrochemical Machining

The *pulsed electrochemical machining* (PECM) process is a refinement of ECM; it uses very high current densities (on the order of 1 A/mm^2), but the current is *pulsed*, rather than direct current. The purpose of pulsing is to eliminate the need for high electrolyte flow rates, which limit the usefulness of ECM in die and mold making. Investigations have shown that PECM improves fatigue life compared to ECM, and the process does not have a characteristic recast layer on die and mold surfaces. The tolerances obtained typically are in the range from 20 to 100 μm.

Case Study 27.1 Electrochemical Machining of a Biomedical Implant

A total knee-replacement system consists of a femoral and tibial implant, combined with an ultrahigh-molecular-weight polyethylene (UHMWPE) insert (Fig. 27.7a). Polyethylene has superior wear resistance and low friction against the cobalt-chrome alloy femoral implant. The UHMWPE insert is compression molded (Section 19.7), and the metal implant is cast and ground on its external mating surfaces.

Designers of implants, manufacturing engineers, and clinicians have long been concerned particularly with the contact surface in the cavity of the metal implant that mates with a protrusion on the polyethylene insert. As the knee articulates during its normal motions, the polyethylene implant slides against the metal part, becoming a potentially serious wear site (Section 33.5). This geometry is necessary to ensure lateral stability of the knee, preventing the knee from buckling sideways.

To produce a smooth surface, grinding of the bearing surfaces of the metal implant, using both hand-held and cam-mounted grinders, was a procedure that had been followed for many years. However, grinding produced marginal repeatability and part quality. The interior surfaces of this part are extremely difficult to access for grinding, and the cobalt–chrome alloy is difficult to grind consistently. Consequently, advanced machining processes, particularly electrochemical machining, were considered to be ideal candidates for this operation.

As shown in Fig. 27.7b, the current procedure consists of placing the metal implant in a fixture to bring a tungsten electrode of the desired final contour in close proximity to the implant. The electrolyte is a sodium nitrate and water mixture, and it is pumped through the tool, filling the gap between the tool and the implant. A power source, typically 10 V and 225 A, is applied, causing local electrochemical machining of the high spots on the implant surface and producing a polished surface.

The electrolyte flow rate can be controlled so as to maximize surface quality. If the rate is too low, defects appear on the machined surface as localized dimples, and if the flow rate is too high, machining times become longer, reducing production times. Typical machining times for this part are four to six minutes. The ECM process can be effective for *micromachining* as well (Section 29.2). Because of complete absence of tool wear, this process also can be used for making precision electronic components.

Source: Courtesy of T. Hershberger and R. Redman, Zimmer Biomet, Inc.

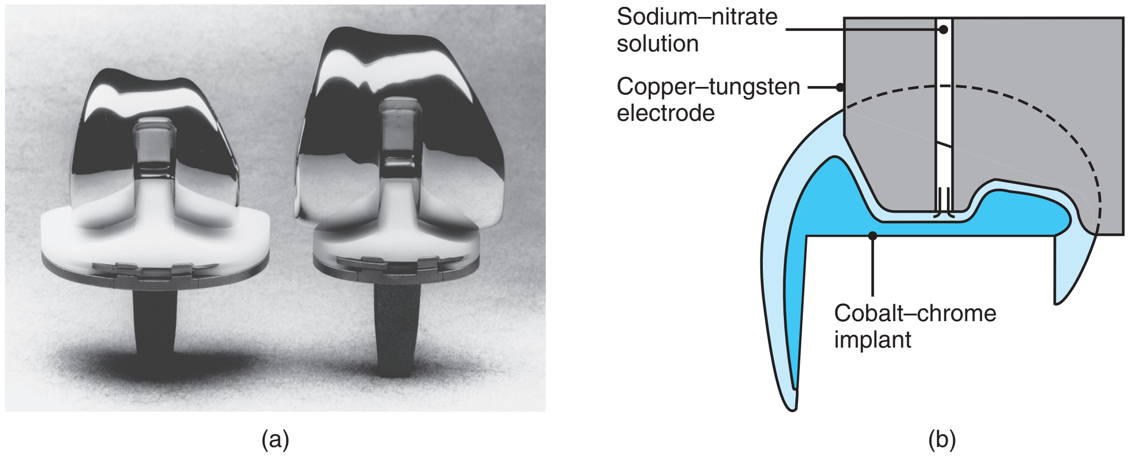


Figure 27.7: (a) Two total knee-replacement systems, showing metal implants (top pieces) with an ultrahigh-molecular-weight polyethylene insert (bottom pieces). (b) Cross section of the ECM process, as applied to the metal implant. *Source:* Courtesy of Zimmer Biomet, Inc.

Machines can perform a combination of both EDM and PECM, thus the need to move the tool and workpiece between the two processes is eliminated. If these operations occur on separate machines, it is difficult to maintain precise alignment when moving the workpiece from the EDM to the PECM operation. If misaligned significantly, all polishing will occur at a location where the gap is smallest, and passivation (Section 3.8) will occur where the gap is largest. Also, this process leaves metal residues suspended in the aqueous solution, which is harmful to the environment if disposed of without proper treatment.

27.4 Electrochemical Grinding

Electrochemical grinding (ECG) combines electrochemical machining with conventional grinding. The equipment used is similar to a conventional grinder, with the exception that the wheel is now a rotating *cathode*, embedded with abrasive particles (Fig. 27.8a). The wheel is metal bonded with diamond or aluminum-oxide abrasives, and rotates at a surface speed from 1200 to 2000 m/min.

The abrasives have two functions: (a) serve as insulators between the wheel and the workpiece and (b) mechanically remove electrolytic products from the working area. A flow of electrolyte solution, usually sodium nitrate, is provided for the *electrochemical* machining phase of the operation where current densities range from 1 to 3 A/mm^2. The majority of metal removal in ECG is by electrolytic action, and, typically, less than 5% of the metal is removed by the *abrasive* action of the wheel; consequently, wheel wear is very low and the part remains cool. Finishing cuts usually are made by the grinding action, but only to produce a surface with good finish and dimensional accuracy.

The ECG process is suitable for applications similar to those for milling, grinding, and sawing (Fig. 27.8b), but it is not adaptable to cavity sinking for die making. The process can be applied successfully to carbides and high-strength alloys, and it offers a distinct advantage over traditional diamond-wheel grinding of very hard materials, where wheel wear can be high. ECG machines are equipped with numerical controls, improving dimensional accuracy and repeatability, and increased productivity.

Electrochemical honing combines the fine abrasive action of honing (Section 26.7) with electrochemical action. Although the equipment is costly, this process is as much as 5 times faster than conventional honing, and the tool lasts as much as 10 times longer. Electrochemical honing is used primarily for finishing internal cylindrical surfaces.

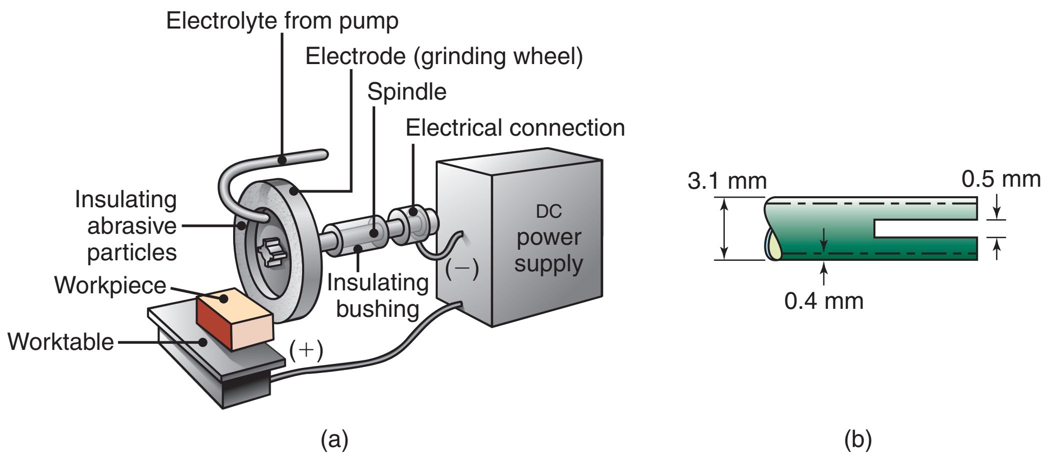


Figure 27.8: (a) Schematic illustration of the electrochemical-grinding process. (b) Thin slot produced on a round nickel-alloy (Inconel) tube by this process.

Design Considerations for Electrochemical Grinding. In addition to the design considerations listed above for electrochemical machining, ECG requires two more:

- Designs should avoid sharp inside radii.
- If a surface is to be flat, it should be narrower than the width of the grinding wheel.

27.5 Electrical-discharge Machining

The principle of *electrical-discharge machining* (EDM), also called *electrodischarge* or *spark-erosion machining*, is based on the erosion of metals by *spark discharges*. Recall that when two current-conducting wires are allowed to touch each other, an arc is produced. When the contact between the two wires is closely examined it will be noted that a small portion of the metal has been eroded away, leaving a small crater on the surface. Although this phenomenon has been known since the discovery of electricity, it was not until the 1940s that a machining process based on that principle was developed. The EDM process is one of the most important and widely used production technologies in manufacturing.

Principle of Operation. The basic EDM system consists of a shaped tool (*electrode*) and the part, connected to a DC power supply and placed in a **dielectric** (electrically nonconducting) fluid, as shown in Fig. 27.9a. When the potential difference between the two is sufficiently high, the dielectric breaks down and a transient spark discharges through the fluid, removing a very small amount of metal from the workpiece surface. The capacitor discharge is repeated continuously, at rates between 200 and 500 kHz, with voltages usually ranging between 50 and 380 V and currents from 0.1 to 500 A. The volume of material removed *per spark discharge* is typically in the range from 10^{-6} to 10^{-4} mm^3.

The EDM process can be used on any material that is an electrical conductor. Two important physical properties determine the volume of metal removed per discharge: melting point and the latent heat of melting of the workpiece material. As these quantities increase, the rate of material removal decreases. The material-removal rate can be estimated from the empirical formula

$$\text{MRR} = 4 \times 10^4 I T_w^{-1.23}, \tag{27.2}$$

where MRR is in mm^3/min, I is the current in amperes, and T_w is the melting point of the workpiece in °C.

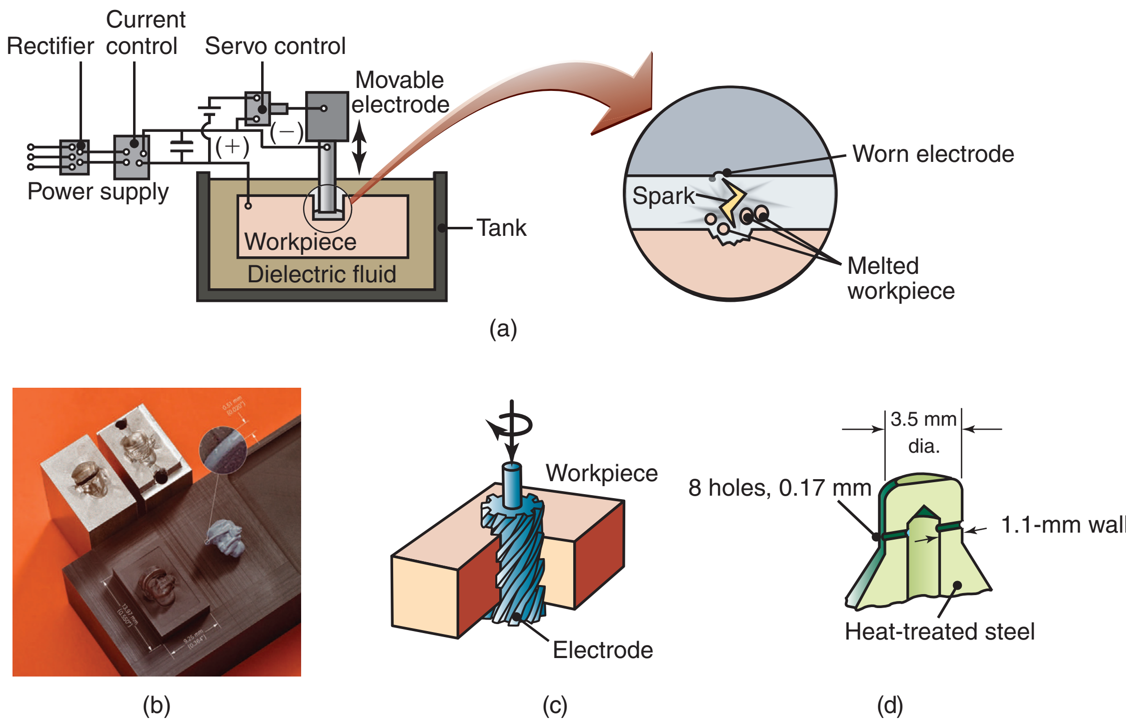


Figure 27.9: (a) Schematic illustration of the electrical-discharge machining process; this is one of the most widely used machining processes, particularly for die-sinking applications. (b) Examples of cavities produced by EDM, using shaped electrodes. The two round parts (rear) are the set of dies used in extruding the aluminum piece shown in front (see also Fig. 15.9b). (c) A spiral cavity produced by EDM using a slowly rotating electrode similar to a screw thread. (d) Holes in a fuel-injection nozzle made by EDM; the material is heat-treated steel. *Source:* (b) Courtesy of AGIE USA, Ltd.

The workpiece is fixtured within a tank containing the dielectric fluid, and its movements are controlled by numerically controlled systems. The gap between the tool and the workpiece (*overcut*) is critical; the downward feed of the tool is thus controlled by a servomechanism, automatically maintaining a constant gap. The frequency of discharge or the energy per discharge, the voltage, and the current are varied to control the removal rate. The rate and surface roughness produced increase with increasing current density and with decreasing frequency of sparks.

Dielectric Fluids. The functions of the dielectric fluid are to:

1. Act as an insulator until the potential is sufficiently high
2. Provide a cooling medium
3. Act as a flushing medium and carry away the debris in the gap.

EDM machines are equipped with a pump and filtering system for the dielectric fluid. The most common dielectric fluids are mineral oils, although kerosene and distilled or deionized water also are used in specialized applications. Although more expensive, low-viscosity transparent fluids that make cleaning easier are also available.

Electrodes. Electrodes for EDM usually are made of graphite, although brass, copper, or copper–tungsten alloys also are used. The tools can be shaped by forming, casting, powder metallurgy, or by CNC machining techniques. Tungsten-wire electrodes, as small as 0.1 mm in diameter, have been used to make holes with depth-to-hole diameter ratios of up to 400:1, a ratio that is much higher than those available by conventional methods (Table 23.11).

The sparks in EDM also erode away the electrode, thus changing its geometry and adversely affecting the shape produced and its dimensional accuracy. *Wear ratio* is defined as the ratio of the volume of workpiece material removed to the volume of tool wear. It ranges from about 3:1 for metallic electrodes to as high as 100:1 for graphite electrodes. Tool wear is related to the melting points of the materials involved: the higher the melting point of the electrode, the lower is the wear rate; consequently, graphite electrodes have the highest wear resistance. Also, the higher the current, the higher is the wear. Tool wear can be minimized by reversing the polarity and using copper tools, a process called **no-wear EDM**. Care must be taken to control this process; it is, for instance, possible for the workpiece material to coat the electrode and thus change its shape.

Process Capabilities. Electrical-discharge machining has numerous applications: dies for forging, extrusion, die casting, injection molding, and large sheet-metal automotive-body components (produced in **die-sinking machining centers**, with computer numerical control). Other applications include machining small-diameter deep holes, using tungsten wire as the electrode; narrow slots in parts; cooling holes in superalloy turbine blades; and various intricate shapes (see Figs. 27.9b and c). Stepped cavities also can be produced by controlling the relative movements of the workpiece in relation to the electrode.

Blue Arc Machining. One variation of electrical discharge machining is the *blue arc* process, developed for roughing cuts of difficult-to-machine materials, especially nickel-based superalloys. The shape of bladed disks, called *blisks*, used in aircraft engines can be challenging to machine; the blue arc process removes most of the material for a rough shape, which then is finish machined through conventional CNC milling. This process uses an electrode and electrical discharge machining to remove material, also adding high pressure fluid flushing to remove chips from the cutting zone. Variations of this technique are available also for turning and grinding.

Because of the molten and resolidified (recast) surface structure developed, high rates of material removal may produce a very rough surface finish, with poor surface integrity and low fatigue properties. Finishing cuts are therefore made at low removal rates, or the recast layer is subsequently removed by various finishing operations. Surface finish can be improved by *oscillating* the electrode in a planetary motion, at amplitudes of 10 to 100 μm.

Design Considerations for EDM. The general design guidelines for electrical-discharge machining are the following:

- Parts should be designed so that the required electrodes can be shaped economically.
- Deep slots and narrow openings should be avoided.
- For economic production, the surface finish specified should not be too fine.
- In order to achieve high production rate, the bulk of material removal should be done by conventional processes, called *roughing out*.

27.5.1 Wire EDM

An important variation of EDM is *wire EDM* or *electrical-discharge wire cutting*. This process is similar to contour cutting with a band saw (Fig. 24.28), in which a slowly moving wire travels along a prescribed path and cuts the workpiece by the EDM action. Figure 27.11a shows a thick plate being cut by this process, on a machine similar to that shown in Fig. 27.11b. Plates as thick as 300 mm, and punches, tools, and dies, made of hard metals, and intricate components for the electronics industry, can be cut by this process.

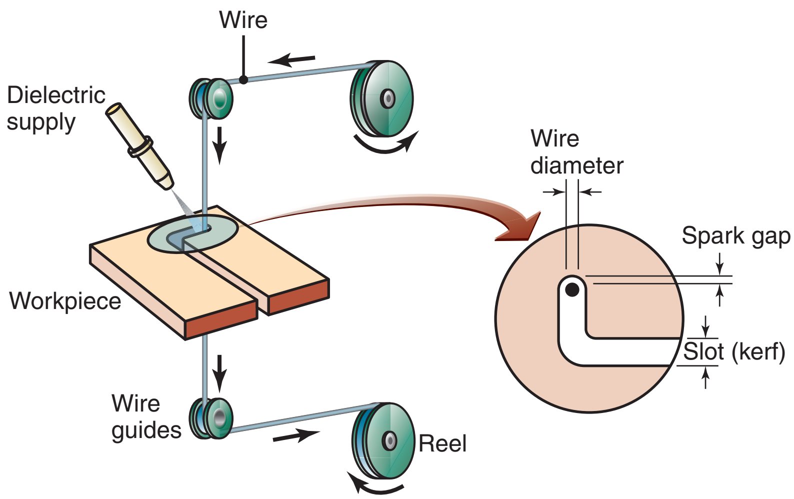


Figure 27.10: Schematic illustration of the wire EDM process. As many as 50 hours of machining can be performed with one reel of wire, which is then discarded.

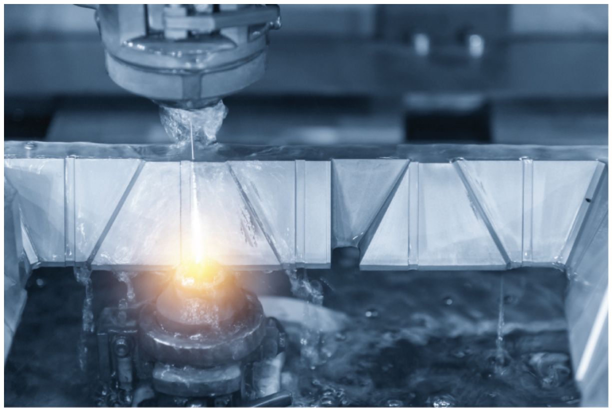
The wire travels at a constant velocity in the range from 0.15 to 9 m/min, and a constant gap (*kerf*) is maintained during the cut. The cutting speed generally is given in terms of the cross-sectional area cut per unit time. Typical examples are 32,000 mm^2/hr for 50-mm thick D2 tool steel and 80,000 mm^2/hr for 150-mm thick aluminum. These removal rates indicate a linear cutting speed of $32{,}000/50 = 640$ mm/hr = 10.7 mm/min and $80{,}000/150 = 533$ mm/hr = 8.9 mm/min, respectively.

The wire is usually made of brass, copper, tungsten, or molybdenum; zinc- or brass-coated, multicoated and steel-cored wires also are used. The wire diameter is typically about 0.30 mm for roughing cuts, and 0.20 mm for finishing cuts. The wire should have high electrical conductivity and tensile strength, as the tension on it is typically 60% of its tensile strength. It usually is used only once, as it is relatively inexpensive compared with the type of operation it performs.

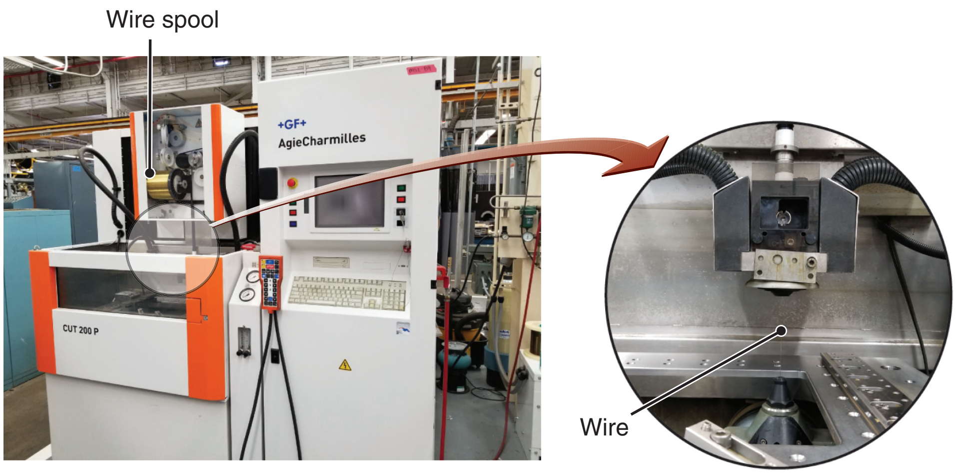
Multiaxis EDM wire-cutting machining centers are capable of producing three-dimensional shapes and are equipped with such features as:

- Computer controls, for controlling the cutting path of the wire and its angle with respect to the workpiece plane
- Multiheads, for cutting two parts at the same time
- Controls for preventing wire breakage
- Automatic self-threading capability, in case of wire breakage
- Programmed machining strategies, to optimize the operation.

Two-axis computer-controlled machines can produce cylindrical shapes, in a manner similar to a turning operation or cylindrical grinding. Modern wire EDM machines allow the control of the feed and take-up ends of the wire, in order to traverse independently in two principal directions, so that tapered parts also can be made.



(a)



(b)

Figure 27.11: (a) Cutting a thick plate with wire EDM. (b) A computer-controlled wire EDM machine. *Source:* Courtesy of L. Love, Oak Ridge National Laboratory.

27.5.2 Electrical-discharge Grinding

The grinding wheel in *electrical-discharge grinding* (EDG) is made of graphite or brass, and contains no abrasives. Material is removed from the workpiece surface by *spark discharges* between the wheel and the workpiece. Although this process is used primarily for grinding carbide tools and dies, it can also be used with fragile parts, such as surgical needles, thin-walled tubes, and honeycomb structures.

The electrical discharges from the graphite wheel break up the oxide film on the workpiece, and is washed away by the flow of the electrolyte. The material-removal rate can be estimated from the equation

$$\text{MRR} = KI, \tag{27.3}$$

where MRR is in mm^3/min, I is the current in amperes, and K is a workpiece material factor in units of mm^3/A-min; for example, $K = 4$ for tungsten carbide, and $K = 16$ for steel.

In **EDM sawing**, a setup similar to a band or circular saw, but without any teeth, is used with the same electrical circuit as for EDM. Narrow cuts can be made in this way and at high rates of metal removal. Because the cutting forces are negligible, the process can also be used on thin and slender components.

The EDG process can be combined with electrochemical grinding. Called **electrochemical-discharge grinding** (ECDG), the process uses a graphite wheel and intermittent spark discharges, from alternating current or pulsed direct current. ECDC also commonly uses a highly conductive electrolyte, instead of a dielectric fluid, and lower voltages. The process is faster than EDG, but power consumption is higher.

27.6 Laser-beam Machining

In *laser-beam machining* (LBM), the source of energy is a **laser** (an acronym for *l*ight *a*mplification by *s*timulated *e*mission of *r*adiation), which focuses optical energy on the workpiece surface (Fig. 27.12a). The highly focused, high-density energy source melts and evaporates portions of the workpiece in a controlled manner. This process, which does not require a vacuum, is used to machine a variety of metallic and nonmetallic materials.

There are several types of lasers used in manufacturing operations (Table 27.2):

1. **CO_2** (*pulsed* or *continuous wave*)
2. **Nd:YAG** (neodymium: yttrium–aluminum–garnet)

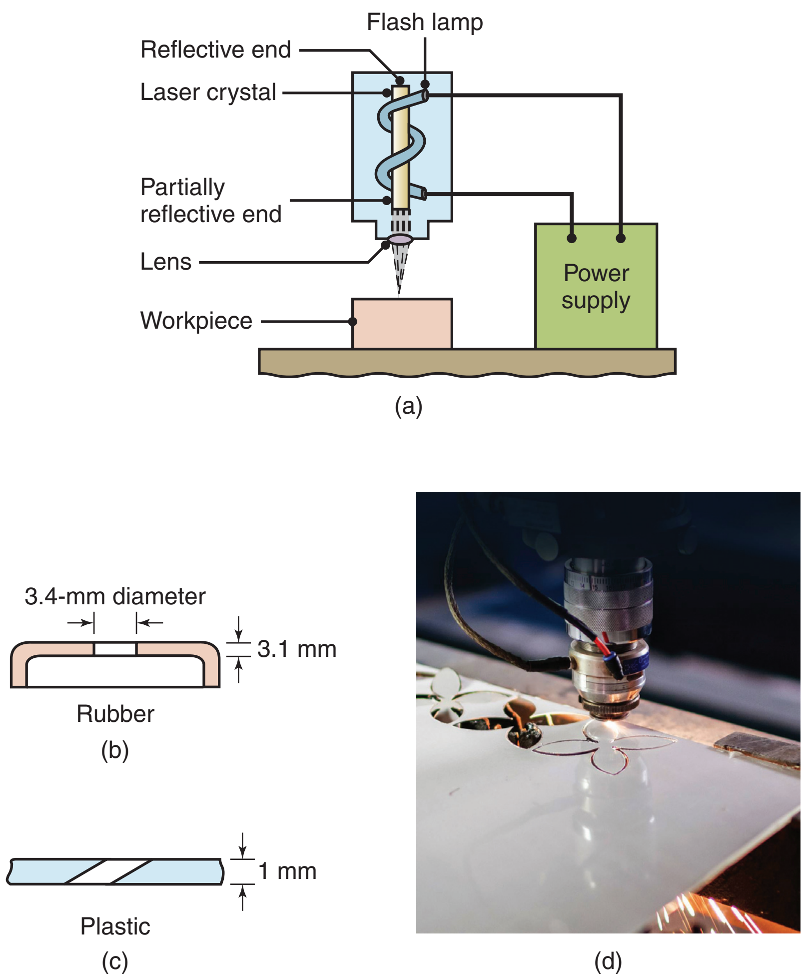


Figure 27.12: (a) Schematic illustration of the laser-beam machining process. (b) and (c) Examples of holes produced in nonmetallic parts by LBM. (d) Cutting sheet metal with a laser beam. *Source:* (d) Courtesy of SPI Lasers UK Ltd.

3. **Nd:glass, ruby**
4. **Diode**
5. **Excimer**, from the words *exc*ited and di*mer*, meaning two mers, or two molecules of the same chemical composition.

Important physical parameters in LBM are the *reflectivity* and *thermal conductivity* of the workpiece surface and its specific heat and latent heats of melting and evaporation (Chapter 3). The lower these quantities, the more efficient is the process. The cutting depth may be expressed as

$$t = \frac{CP}{vd}, \tag{27.4}$$

where t is the depth, C is a constant for the process, P is the power input, v is the cutting speed, and d is the laser-spot diameter. Peak energy densities of laser beams are in the range from 5 to 200 kW/mm^2.

The surface produced by LBM is usually relatively rough and with a *heat-affected zone* (Section 30.9); the zone may have to be removed or heat treated for critical applications. Kerf width is an important consideration, as it is in other cutting processes, such as sawing, wire EDM, and electron-beam machining. In general, the smaller the kerf, the greater is the dimensional accuracy and material utilization, and the lower is the heat affected zone.

Laser beams may be used in combination with a gas, such as an oxygen stream, called **laser-beam torch**, to increase energy absorption for cutting sheet metals. *High-pressure*, *inert-gas assisted laser cutting* is used for stainless steel and aluminum, leaving an oxide-free edge which can improve weldability of these metals. Gas streams (nitrogen or argon) also have the important function of blowing away molten and vaporized material from the workpiece surface.

Process Capabilities. Laser-beam machining is widely used for hole making, trepanning, and cutting metals, nonmetallic materials, ceramics, and composite materials (Fig. 27.12b and c). The cleanliness of the operation has made laser-beam machining an attractive alternative to traditional machining methods. Holes as small as 0.005 mm, with depth-to-diameter ratios of 50:1, have been made with various materials, although a more practical minimum is 0.025 mm. Steel plates as thick as 32 mm also can be cut with laser beams.

Table 27.2: General Applications of Lasers in Manufacturing.

Application	Laser type
Cutting	
Metals	PCO_2, $CWCO_2$, Nd:YAG, ruby
Plastics	$CWCO_2$
Ceramics	PCO_2
Drilling	
Metals	PCO_2, Nd:YAG, Nd:glass, ruby
Plastics	Excimer
Marking	
Metals	PCO_2, Nd:YAG
Plastics	Excimer
Ceramics	Excimer
Surface treatment	$CWCO_2$
Welding	
Metals	PCO_2, $CWCO_2$, Nd:YAG, Nd:glass, ruby, diode
Plastics	Diode, Nd:YAG
Lithography	Excimer

Note: P = pulsed, CW = continuous wave,
Nd:YAG = neodynmium: yttrium–aluminum–garnet.

Significant cost savings can be achieved by LBM, a process that competes with electrical-discharge machining. It is used increasingly in the electronics and automotive industries, and for composite materials. Two typical examples of laser machining are (a) the cooling holes in some vanes for the Boeing 747 jet engines, and (b) bleeder holes for fuel-pump covers and lubrication holes in transmission hubs.

Lasers are also used for the following applications:

- **Welding** (Section 30.7)
- Small-scale and localized **heat treating** of metal and ceramic parts, to modify their surface mechanical and tribological properties
- **Laser forming** and **laser peen forming** (Section 16.12)
- **Marking** parts, such as letters, numbers, and codes; note that marking also can be done by (a) punches, pins, styluses, and scroll rolls, (b) stamping, and (c) etching. Although the equipment is more expensive than that used in other methods, laser marking and engraving has increasingly become common due to its accuracy, reproducibility, flexibility, ease of automation, and online application in manufacturing.

The inherent *flexibility* of the laser-cutting process, including its *fiber-optic beam* delivery, simple fixturing, low setup times, availability of multi-kW machines, and two- and three-dimensional computer-controlled robotic laser-cutting systems are competitive and attractive features of laser-beam machining. Laser cutting of sheets, for example, successfully replace traditional punching processes (Chapter 16). Laser beams can be combined with other processes for improved overall efficiency (Section 27.10; see also Example 27.1.

Example 27.1 Combining Laser-beam Cutting and Punching of Sheet Metal

Laser cutting and punching processes have their respective advantages and limitations regarding both technical and economic aspects (see *hybrid machining*, Section 27.10). The advantages of *laser-beam cutting* generally are (a) flexibility of the operation, because hard tooling is not needed and there is no limitation to part size, (b) wide range of material thicknesses, (c) prototyping capability, and lot sizes that can be as low as one, (d) materials and composites that otherwise might be cut with difficulty, and (e) complex geometries that can easily be programmed.

Drawbacks and advantages of *punching* include (a) large lot sizes that economically justify the purchase of tooling and equipment, (b) relatively simple shapes, (c) small range of part thicknesses, (c) fixed and limited punch geometries, even when using turrets, and (d) high production rate.

The two processes cover different but complementary ranges. It is not difficult to visualize parts with some features that can be produced best by one process and other features that are best produced by the other process.

Machines have been designed and built in such a manner that the processes and fixturing can be utilized jointly (**hybrid machines**) to their full extent, without interfering with each other's operational boundaries. The purpose of combining them is to increase the overall efficiency and productivity of the manufacturing process, for parts that are within the capabilities of each of the two processes, similar to the concept of machining centers (Section 25.2). For example, turret-punch presses can be equipped with an integrated laser head; the machine can either punch or laser cut, but it cannot do both simultaneously.

Several factors must be taken into account in such a combination of processes with respect to the characteristics of each operation: (a) ranges of sizes, thicknesses, and shapes to be produced, and how they are to be nested (see Fig. 16.59); (b) processing and setup times, including loading, fixturing, and unloading of parts; (c) programming for cutting; and (d) process capabilities of each method, including system dynamics, vibrations, and shock from mechanical punching that may disturb adjustments and alignments of the laser components.

Design Considerations for LBM. General design guidelines for laser-beam machining are:

- Sharp corners should be avoided, because they can be difficult to produce.
- Deep cuts will produce tapered walls.
- Dull and unpolished surfaces are preferable.
- There can be adverse effects on the properties of the machined materials, caused by high local temperatures and the heat-affected zone.

Laser microjet®. Laser microjet®, illustrated in Fig. 27.13, uses a low-pressure, laminar water stream to serve as a variable-length fiber-optic cable to direct the laser and deliver laser power at the bottom of the kerf. This has an advantage in that the laser focus is very deep, and cuts with large aspect ratios can be made. The water jet is produced by a sapphire or diamond nozzle, with an opening of 25 to 100 μm, and exerting a force of less than 0.1 N. In laser microjet® machining, material removal is due to the action of the laser, and the water provides cooling of the heat affected zone, and prevents weld splatter from attaching to the workpiece. The laser is typically a Nd:Yag laser, with micro- or nano-second pulse duration and power between 10 and 200 W.

27.7 Electron-beam Machining

The energy source in *electron-beam machining* (EBM) is high-velocity electrons, striking the workpiece surface and generating heat (Fig. 27.14). The machines utilize voltages in the range from 150 to 200 kV, to accelerate the electrons to 50% to 80% of the speed of light (300,000 km/s). Applications of this process are similar to those of laser-beam machining, except that, unlike lasers, EBM requires a *vacuum*; consequently, it is used much less frequently than laser-beam machining.

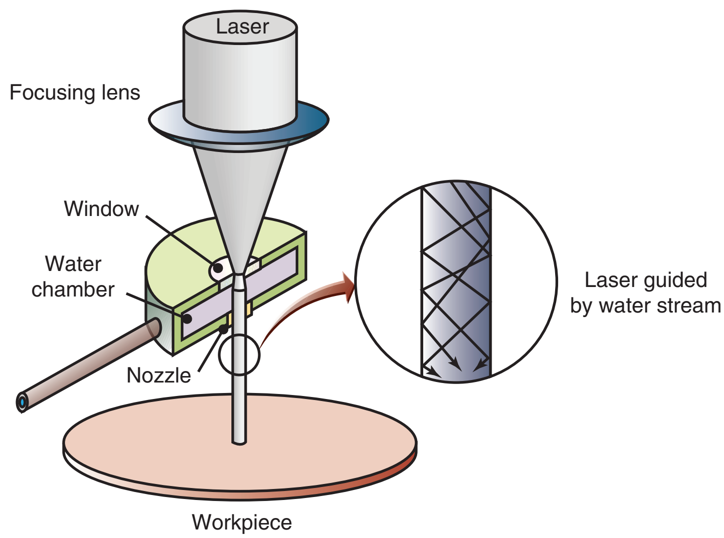


Figure 27.13: Schematic illustration of the laser microjet® process.

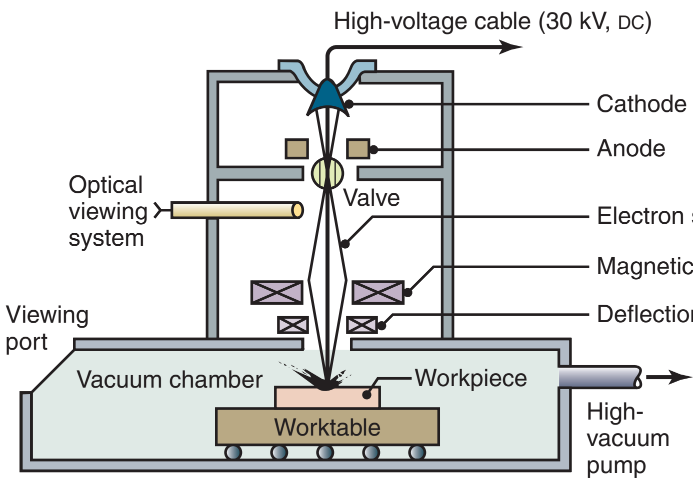


Figure 27.14: Schematic illustration of the electron-beam machining process. Unlike LBM, this process requires a vacuum, so the workpiece size is limited to the size of the vacuum chamber.

Electron-beam machining can be used for very accurate cutting of a wide variety of metals, with surface finish that is better and the kerf is narrower than in other thermal cutting processes (see also Section 30.6 on *electron-beam welding*). However, the interaction of the electron beam with the workpiece surface produces hazardous X-rays, thus the equipment should be used only by highly trained personnel.

Design Considerations for EBM. The guidelines for EBM generally are similar to those for LBM; additional considerations are:

- Because vacuum chambers have limited capacity, individual parts or batches should closely match the size of the vacuum chamber.

- If a part requires electron-beam machining on only a small portion of its volume, consideration should be given to making a number of smaller components, then assembling them.

Plasma-arc Cutting. In plasma-arc cutting (PAC), *plasma beams* (ionized gas) are used to rapidly cut ferrous and nonferrous sheets and plates (Section 30.3). The temperatures generated in the torch are on the order of 9400°C, when using oxygen as a plasma gas. Material-removal rates are thus much higher than those associated with the EDM and LBM processes. The process is rapid, kerf width is small, parts can be machined with good reproducibility, and the surface finish is good; parts as thick as 150 mm can be cut. Plasma-arc cutting is highly automated, using programmable controllers.

27.8 Water-jet Machining

The principle of *water-jet machining* (WJM), also called **hydrodynamic machining**, is based on the force resulting from the momentum change of a stream of water. This force is sufficiently high to cut metallic and nonmetallic materials (Fig. 27.15). The water jet acts like a saw and cuts a narrow groove in the material (Fig. 27.15b; see also *water-jet peening*, Section 34.2).

A wide variety of materials can be cut, including plastics, rubber, wood products, paper, fabrics, leather, insulating materials, brick, and composite materials (Fig. 27.15c). A pressure level of 400 MPa is generally used for efficient operation, although pressures as high as 1400 MPa are available. Jet-nozzle diameters typically range between 0.05 and 1 mm. The process also can be used for deburring operations.

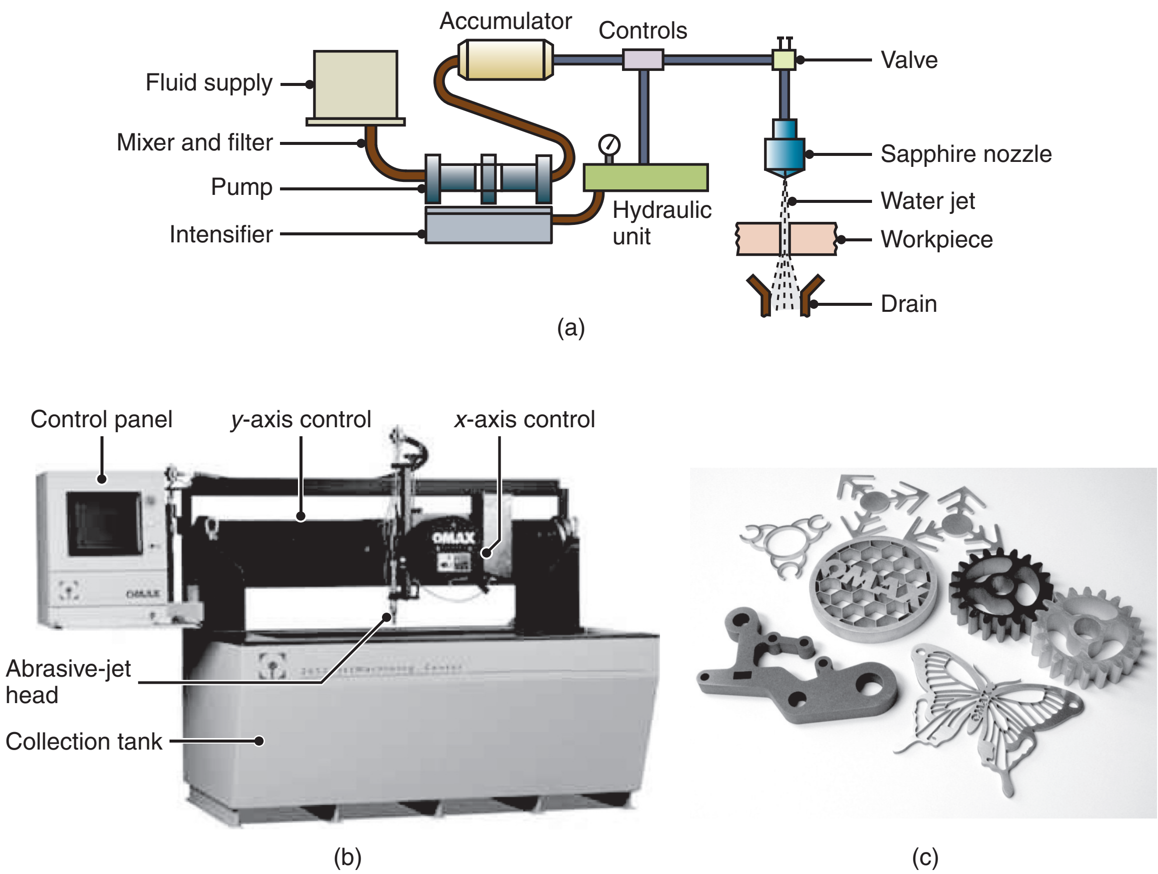


Figure 27.15: (a) Schematic illustration of the water-jet machining process. (b) A computer-controlled water-jet cutting machine. (c) Examples of various nonmetallic parts produced by the water-jet cutting process. *Source:* Courtesy of OMAX Corporation.

Depending on the materials, thickness can range up to 25 mm and higher. Vinyl and foam coverings for automobile dashboards, as well as some body panels, can be cut by *multiple-axis, robot-guided water-jet* equipment. Because it is an efficient and clean operation, as compared to most other cutting processes, it is also used in the food-processing industry for cutting and slicing food products.

The advantages of WJM are:

- Cuts can be started at any location without the need for predrilled holes.
- No heat is produced.
- No deflection of the rest of the workpiece takes place, thus making the process suitable for flexible materials.
- Little wetting of the workpiece takes place.
- The burr produced is minimal.
- It is an environmentally safe manufacturing operation.

Abrasive Water-jet Machining. In *abrasive water-jet machining* (AWJM), the water jet contains abrasive particles, such as silicon carbide or aluminum oxide, which greatly increase the material-removal rate. In modern machines, the optimum level of abrasives in the jet stream is controlled automatically. Nozzles are typically of rubies, sapphires, and diamond (Fig. 27.15a).

AWJM is particularly suitable for heat-sensitive materials that cannot be machined by processes in which heat is produced. Cutting speeds can be as high as 7.5 m/min for reinforced plastics, but are much lower for metals; consequently, the process may not be economical for applications requiring high production rates. Metallic, nonmetallic, and composite materials of various thicknesses, and in single layer or multilayers, can be cut. With multiple-axis and robot-controlled machines, complex three-dimensional parts can be machined economically to finish dimensions.

27.9 Abrasive-jet Machining

In *abrasive-jet machining* (AJM), abrasive particles are propelled at the workpiece by a high-velocity jet of dry air, nitrogen, or carbon dioxide (Fig. 27.16). The particles impact the surface with a concentrated force (see also Section 26.6) that is sufficiently high to chip away materials. Typical applications are (a) cutting small holes, slots, and intricate patterns in very hard or brittle metallic and nonmetallic materials, (b) deburring or removing small flash from parts, (c) trimming and beveling of edges on parts, (d) removing oxides and other surface films, and (e) cleaning of parts with irregular surfaces.

The gas pressure is on the order of 850 kPa; the abrasive-jet velocity can be as high as 300 m/s. Nozzles are usually made of tungsten carbide or sapphire, both of which have abrasive wear resistance. Abrasive particle size is in the range from 10 to 50 μm. Because the flow of the free abrasives tends to round off corners, designs for abrasive-jet machining should avoid sharp corners; also, holes made tend to be tapered, because abrasives preferentially wear the inlet side. There is some hazard involved in using this process, because of airborne particulates.

27.10 Hybrid Machining Systems

Two or more individual machining processes can be *combined* into one system, thus taking advantage of the capabilities of each process. Examples of hybrid machining systems include combinations and integration of the following processes:

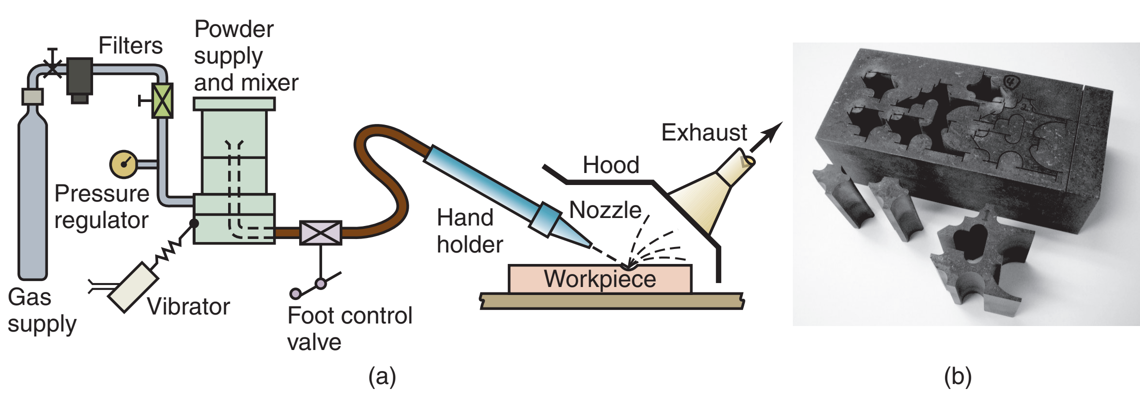


Figure 27.16: (a) Schematic illustration of the abrasive-jet machining process. (b) Examples of parts made by abrasive-jet machining, produced in 50-mm thick 304 stainless steel. *Source:* Courtesy of OMAX Corporation.

1. Abrasive machining and electrochemical machining
2. Abrasive machining and electrical discharge machining
3. Abrasive machining and electrochemical finishing
4. Water-jet cutting and wire EDM
5. High-speed milling, laser ablation, and abrasive blasting, as an example of *three* integrated processes
6. Machining and blasting
7. Electrochemical machining and electrical discharge machining (ECDM), also called electrochemical spark machining (ECSM)
8. Machining and forming processes, such as laser cutting and punching of sheet metal, described in Example 27.1
9. Combinations of various other forming, machining, and joining processes.

The system is able to handle a variety of materials, including metals, ceramics, polymers, and composites. Implementation of these concepts and the development of appropriate machinery and control systems present significant challenges. Important considerations include such factors as:

1. The workpiece material and its manufacturing characteristics (see, for example, Table I.3 in the General Introduction).
2. Compatibility of processing parameters among the two or more processes to be integrated, such as speed, temperature, size, force, and energy.
3. Cycle times for each individual operation involved and their synchronization.
4. Safety considerations and possible adverse effects of the presence of various elements, such as abrasives, chips, chemicals, wear particles, and contaminants.
5. Consequence of a failure in one of the stages in the system, since the operation involves sequential processes.

27.11 Economics of Advanced Machining Processes

Advanced machining processes have unique applications, and are important particularly for difficult-to-machine materials and for parts with complex internal and external features. The economic production run for a particular process depends on such factors as the (a) costs of tooling and equipment, (b) operating costs, (c) material-removal rate, (d) level of operator skill required, and (e) secondary and finishing operations that subsequently may be necessary.

Case Study 27.2 Manufacturing of Small Satellites

Satellites built in the early days of the Space Age (1960s) were very large, and those smaller than 1000 kg were very rare. Table 27.3 shows the classification of modern satellites by their mass. This case study describes the manufacture of propulsion systems for micro- and nanosatellites.

There are several compelling reasons for reducing the size of satellites, none greater than the cost of putting the satellite into orbit. One of the main contributors to weight in a satellite is the propulsion system, essential for changing its orbit or correcting for any drift. Figure 27.18a shows the propulsion system for a microsatellite, incorporating several cold-gas microthrusters, a propellant storage tank, filters, and temperature and pressure sensors.

Selected components of the propulsion system are shown in Fig. 27.17b. Note that production of these miniature parts would be extremely difficult and costly if made through conventional forming, casting, or machining technologies. Moreover, connecting the plumbing for all these components would be very difficult, even with larger components, and almost impossible to perform inside a clean-room environment.

An attractive alternative is the production of an integrated system, with fluid connections made internally through a photochemically etched and diffusion-bonded support, on which components are welded or fastened mechanically. Such a support is shown in Fig. 27.19, along with valve springs and filters that are made through a combination of photochemical blanking diffusion-bonding processes.

Figure 27.19 depicts the manufacturing sequence involved. Titanium is commonly used for propulsion-system components, because it has a high strength-to-weight ratio, thus making possible lightweight designs. A mask is first prepared (Section 27.2), and the titanium is etched or blanked in a solution of hydrofluoric and nitric acid. Multiple layers of titanium are then diffusion bonded (Section 31.7), to produce internal features, such as flow channels.

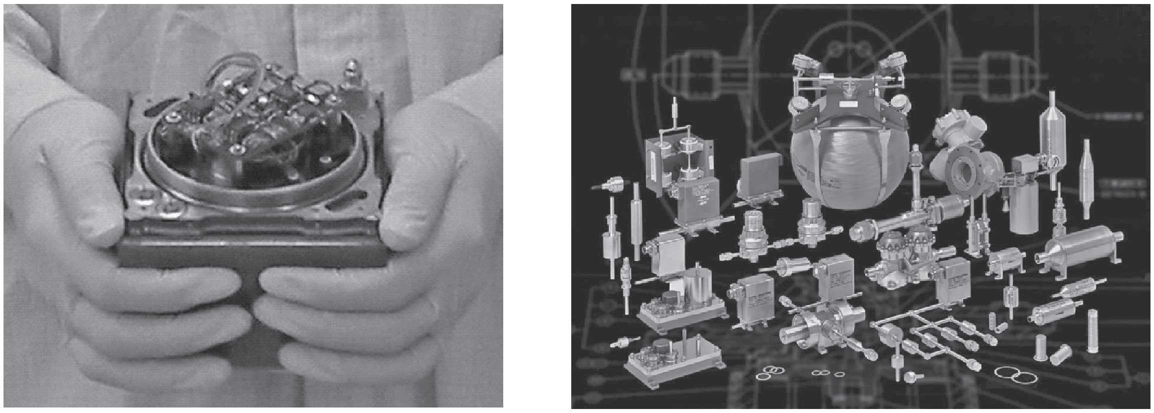
Such fully-integrated systems have resulted in making satellite propulsion systems that are less complex, more robust, and less massive than those in previous designs.

Source: Courtesy of R. Hoppe, VACCO Industries, Inc.

Table 27.3: Satellite Classification.

Group name	Mass kg or g
Large satellite	> 1000 kg
Medium satellite	500–1000
Minisatellite	100–500
Small satellites	
Microsatellite	10–100
Nanosatellite	1–10
Picosatellite	0.1–1
Femtosatellite	< 100 g

In chemical machining, an inherently slow process, an important factor is the cost of reagents, maskants, and their disposal, together with the cost of cleaning the parts. In electrical-discharge machining, the cost of electrodes and the need to periodically replace them can be significant. The rate of material removal and the production rate can vary significantly, as can be seen in Table 27.1. The cost of tooling and equipment varies considerably, as does the operator skill required. The high capital investment for some machinery (Table 40.6), especially when equipped with robotic control, has to be justified in terms of the production runs and the feasibility of making the same part by other methods.



(a) (b)

Figure 27.17: Propulsion system for a small satellite. (a) Miniaturized system suitable for a micro- or nanosatellite, and (b) selected propulsion system components. *Source:* Courtesy of R. Hoppe, VACCO Industries, Inc.

(a) (b) (c)

Figure 27.18: Photochemically etched and blanked components for micro- and nanosatellites. (a) Mounting board incorporating fluid flow channels in an integrated package, (b) microscale valve spring placed next to a U.S. penny, and (c) fuel filter. *Source:* Courtesy of R. Hoppe, VACCO Industries, Inc.

Summary

- Advanced machining processes have unique capabilities, utilizing chemical, electrochemical, electrical, and high-energy-beam sources of energy.

- Mechanical properties of the workpiece material are not significant, because these processes rely on mechanisms that do not involve strength, hardness, ductility, or toughness of the material; rather, they involve physical, chemical, and electrical properties.

- Chemical and electrical methods of machining are particularly suitable for hard materials and complex part shapes. They do not exert forces on the workpiece, and therefore can be used for thin, slender, and flexible workpieces. However, their effects on surface integrity must be considered, as they can damage surfaces, reducing the fatigue life of the parts made.

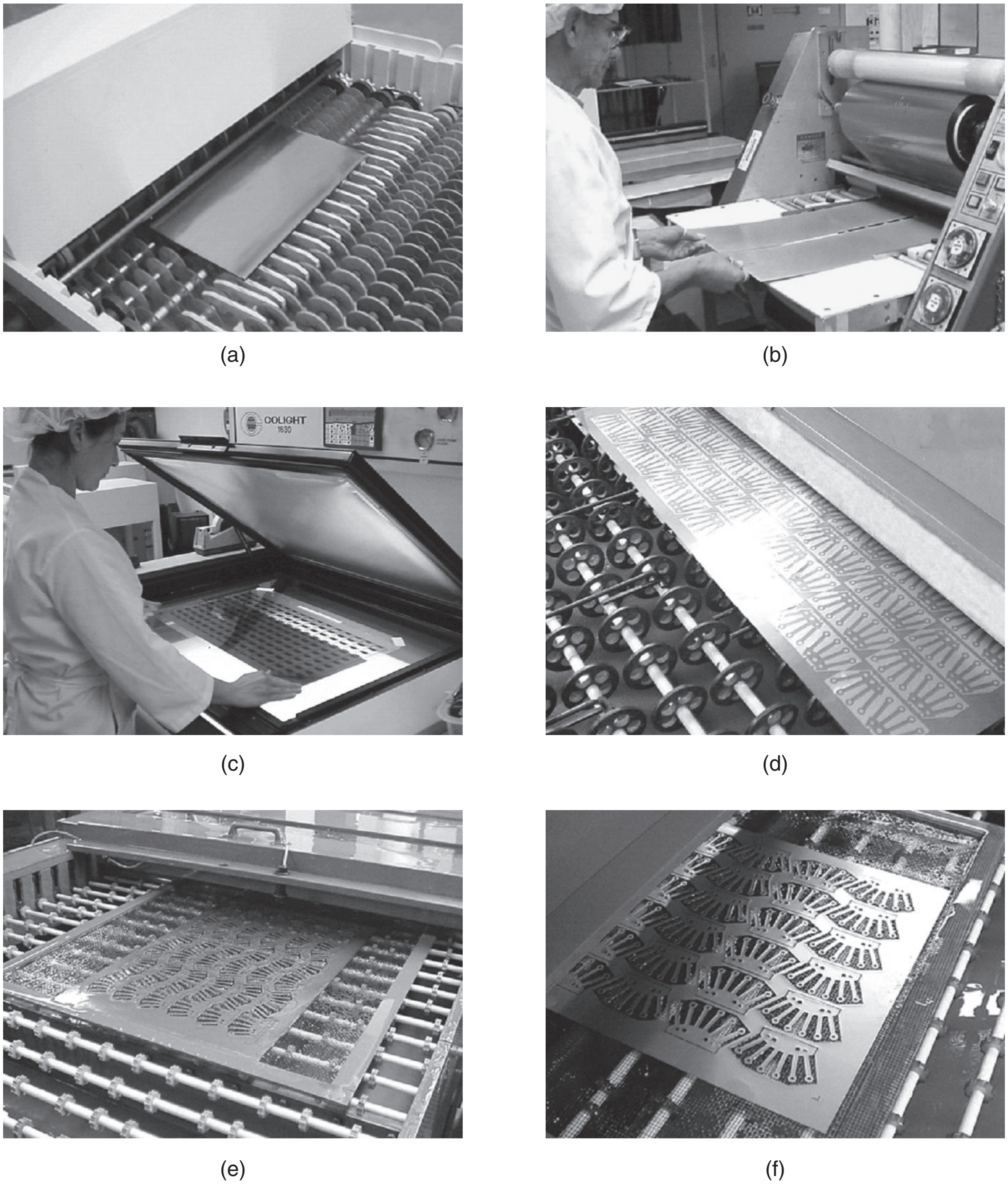


Figure 27.19: Processing sequence for photochemical etching of microsatellite components: (a) clean the raw material, (b) coat with photosensitive material, (c) expose with photographic tool, (d) develop a resist image, (e) etch, and (f) remove the resist. *Source:* Courtesy of R. Hoppe, VACCO Industries, Inc.

- High-energy-beam machining processes basically utilize laser beams, electron beams, and plasma beams. They have important industrial applications, possess high flexibility of operation, with robotic controls, and are economically competitive with other processes.
- Water-jet machining, abrasive water-jet machining, and abrasive-jet machining processes can be used for cutting as well as for deburring operations. Because they do not utilize hard tooling, they have an inherent flexibility of operation.
- Hybrid machining processes offer possibilities for more efficient production of complex parts.

Key Terms

Abrasive-jet machining
Abrasive water-jet machining
Blue arc machining
Chemical blanking
Chemical machining
Chemical milling
Dielectric
Die sinking
Electrical-discharge grinding
Electrical-discharge machining
Electrochemical-discharge grinding
Electrochemical grinding
Electrochemical honing
Electrochemical machining
Electrode
Electrolyte
Electrolytic trepanning
Electron-beam machining
Etchant
Hybrid machining
Hydrodynamic machining
Laser
Laser-beam machining
Laser microjet
No-wear EDM
Photochemical blanking
Photochemical machining
Photoetching
Photoresist
Plasma-arc cutting
Plasma beams
Pulsed electrochemical machining
Reagent
Shaped-tube electrolytic machining
Undercut
Water-jet machining
Wire EDM

Bibliography

Asibu, Jr., E.K., **Principles of Laser Materials Processing**, Wiley, 2009.
Bass, M., **Laser Material Processing**, Elsevier, 2012.
Dahotre, N.B., and Samant, A., **Laser Machining of Advanced Materials**, CRC Press, 2011.
El-Hofy, H., **Advanced Machining Processes: Nontraditional and Hybrid Machining Processes**, McGraw-Hill, 2005.
Grzesik, W., **Advanced Machining Processes of Metallic Materials: Theory, Modelling and Applications**, Elsevier, 2008.
Kibria, G., Bhattacharyya, B., and Davim, J.P., **Non-traditional Micromachining Processes**, Springer, 2017.
Ion, J.C., **Laser Processing of Engineering Materials**, Butterworth-Heinemann, 2011.
Luo, X., and Qin, Y., **Hybrid Machining**, Academic Press, 2018.

Marinescou, I.D., Uhlmann, E., and Doi, T., **Handbook of Lapping and Polishing**, CRC Press, 2006.
Schaaf, P., (ed.), **Laser Processing of Materials**, Springer, 2010.
Schaeffer, R.D., **Fundamentals of Laser Micromachining**, CRC Press, 2012.
Stafe, N., Marcu, A., and Puscas, N.N., **Pulsed Laser Ablation of Solids**, Springer, 2013.
Steen, W.M., and Mazumder, J., **Laser Material Processing**, 4th ed., Springer, 2010.

Review Questions

27.1. Describe the similarities and differences between chemical blanking and conventional blanking using dies.

27.2. Name the processes involved in chemical machining. Describe briefly their principles.

27.3. Explain the difference between chemical machining and electrochemical machining.

27.4. What is the underlying principle of electrochemical grinding?

27.5. Explain how the EDM process is capable of producing complex shapes.

27.6. What are the important features of the blue arc machining process?

27.7. What are the capabilities of wire EDM? Could this process be used to make tapered parts? Explain.

27.8. Explain why laser microjet has a large depth of field.

27.9. Describe the advantages of water-jet machining.

27.10. What is the difference between photochemical blanking and chemical blanking?

27.11. What is kerf?

27.12. What type of workpiece is not suitable for laser-beam machining?

27.13. What is an undercut? Why must it be considered in chemical machining?

27.14. Explain the principle of hybrid machining.

Qualitative Problems

27.15. Give technical and economic reasons that the processes described in this chapter might be preferred over those described in the preceding chapters.

27.16. Why is the preshaping or premachining of parts sometimes desirable in the processes described in this chapter?

27.17. Why is the material removal rate in electrical-discharge machining a function of the melting point of the workpiece material?

27.18. Explain why the mechanical properties of workpiece materials are not significant in most of the processes described in this chapter.

27.19. List the processes that can produce shaped holes, that is, holes that are not circular.

27.20. List the advantages of laser microjet over conventional laser machining.

27.21. Why has electrical-discharge machining become so widely used in industry?

27.22. Describe the types of parts that are suitable for wire EDM.

27.23. Which of the advanced machining processes would cause thermal damage? What is the consequence of such damage to workpieces?

27.24. Which of the processes described in this chapter require a vacuum? Explain why?

27.25. Describe your thoughts regarding the laser-beam machining of nonmetallic materials. Give several possible applications, including their advantages compared with other processes.

27.26. Are deburring operations still necessary for some parts made by advanced machining processes? Explain and give several specific examples.

27.27. List and explain factors that contribute to a poor surface finish in the processes described in this chapter.

27.28. What is the purpose of the abrasives in electrochemical grinding?

27.29. Which of the processes described in this chapter are suitable for producing very small and deep holes? Explain.

27.30. Is kerf width important in wire EDM? Explain.

27.31. Why may different advanced machining processes affect the fatigue strength of materials to different degrees?

27.32. What are the functions of the fluid in EDM?

Quantitative Problems

27.33. A 60-mm-deep hole, 30 mm in diameter, is being produced by electrochemical machining. A high production rate is more important than the quality of the machined surface. Estimate the maximum current and the time required to perform this operation.

27.34. If the operation in Problem 27.33 were performed on an electrical-discharge machine, what would be the estimated machining time?

27.35. A cutting-off operation is being performed with a laser beam. The workpiece being cut is 5 mm thick and 100 mm long. If the kerf width is 3 mm, estimate the time required to perform this operation.

27.36. A 20-mm-thick copper plate is being machined by wire EDM. The wire moves at a speed of 1.2 m/min and the kerf width is 1.6 mm. What is the required power? Note that it takes 1550 J to melt one gram of copper.

Synthesis, Design, and Projects

27.37. It was stated that graphite is the preferred material for EDM tooling. Would graphite be useful in wire EDM? Explain.

27.38. Explain why it is difficult to produce sharp profiles and corners with some of the processes described in this chapter.

27.39. Make a list of the processes described in this chapter in which the following properties are relevant: (a) mechanical, (b) chemical, (c) thermal, and (d) electrical. Are there processes in which two or more of these properties are important? Explain.

27.40. Would the processes described in this chapter be difficult to perform on various nonmetallic or rubberlike materials? Explain your thoughts, commenting on the influence of various physical and mechanical properties of workpiece materials, part geometries, etc.

27.41. Describe the types of parts that would be suitable for hybrid machining. Consider one such part and make a preliminary sketch for a hybrid machine to produce that part.

27.42. Describe your thoughts as to whether the processes described in (a) Chapters 13 through 16, and (b) Chapters 23 and 24 can be suitable for a hybrid system of making parts. Give a preliminary sketch of a machine for the two groups of processes listed.

27.43. Make a list of machining processes that may be suitable for each of the following materials: (a) ceramics, (b) cast iron, (c) thermoplastics, (d) thermosets, (e) diamond, and (f) annealed copper.

27.44. At what stage is the abrasive in abrasive water-jet machining introduced into the water jet? Survey the available literature, and then prepare a schematic illustration of the equipment involved.

27.45. How would you manufacture a large-diameter, conical, round metal disk with a thickness that decreases from the center outward? Make appropriate sketches.

27.46. Describe the similarities and differences among the various design guidelines for the processes described in this chapter.

27.47. Describe any workpiece size limitations in advanced machining processes. Give examples.

27.48. Suggest several design applications for the types of parts shown in Fig. 27.4.

27.49. Based on the topics covered in Parts III and IV, make a comprehensive table of hole-making processes. Describe the advantages and limitations of each method, and comment on the quality and surface integrity of the holes produced.

27.50. Review Example 27.1 and explain the relevant parameters involved; then design a system whereby both processes can be used in combination to produce parts from sheet metal.

27.51. Marking surfaces with numbers and letters for part-identification purposes can be done with a variety of mechanical and nonmechanical methods. Based on the processes described throughout this book thus far, make a list of these methods, explaining their advantages, limitations, and typical applications.

27.52. *Precision engineering* is a term that is used to describe manufacturing high-quality parts with close dimensional tolerances and good surface finish. Based on their process capabilities, make a list of advanced machining processes with decreasing order of the quality of parts produced. Comment on your observations.

27.53. With appropriate sketches, describe the principles of various work-holding methods and work-holding devices that can be used for the processes described in this chapter.

27.54. Make a table of the process capabilities of the advanced machining processes described in this chapter. Use several columns and describe the machines involved, the type of tools and tool materials used, the shapes of blanks and parts produced, the typical maximum and minimum sizes, surface finish, tolerances, and production rates.

27.55. One of the general concerns regarding advanced machining processes is that, in spite of their many advantages, they generally are slower than conventional machining operations. Conduct a survey of the speeds, machining times, and production rates involved, and prepare a table comparing their respective process capabilities.

27.56. It can be seen that several of the processes described in Part IV of this book can be employed, either singly or in combination, to make or finish dies for metalworking operations. Write a brief technical paper on these methods, describing their advantages, limitations, and typical applications.

PART V
Micromanufacturing and Fabrication of Microelectronic Devices

The importance of the topics covered in the following two chapters can best be appreciated by considering manufacturing of a simple metal spur gear. It is important to also recall that some gears are designed to transmit *power*, such as those in gear boxes, and others transmit motion, such as in clocks. If the gear is, say, 100 mm in diameter, it can be produced by traditional methods, such as starting with a cast or forged blank, and machining and grinding it to its final shape and dimensions.

On the other hand, if a gear is only 2 mm in diameter, it can be difficult to produce by these methods; if sufficiently thin, the gear could be made from sheet metal, by fine blanking, chemical etching, or electroforming. If the gear is only a *few micrometers* in size, it has to be produced by such techniques as optical lithography, wet and dry chemical etching, or related processes. If only a *nanometer* in diameter, then the necessity of the very concept of a gear has to be questioned.

The challenges faced in producing gears of increasingly smaller sizes is highly informative, and can be put into proper perspective by referring to the illustration of length scales shown in Fig. V.1. Conventional manufacturing processes, described in Chapters 11 through 27, typically produce parts that are larger than one mm or so, and can be described as visible to the naked eye. The sizes of such parts generally are referred to as **macroscale**, the word “macro” being derived from the Greek *makros*, meaning “long,” and the processing of such parts is known as **macromanufacturing**. Macroscale is the most developed and best understood size range from a design and manufacturing standpoint, with a wide variety of processes available for producing components of that size. All of the examples and case studies given thus far in this book have been examples of macromanufacturing.

Note that the gear shown in Fig. V.1 is the size of a few tens of micrometers across, and thus fits into the category of micromanufacture, developed mostly for electronic devices of all types, including computer processors and memory chips, sensors, and magnetic storage devices. For the most part, this type of manufacturing relies heavily on lithography, wet and dry etching, and coating technologies.

Examples of products that rely on micromanufacturing techniques are a wide variety of sensors and probes (see Fig. V.2), ink-jet and jet fusion printing heads, microactuators, and such devices as computer processors and memory chips. Microscale mechanical devices have become commonplace; examples are smartphone sensors and cameras.

Mesomanufacturing overlaps macro- and micromanufacturing, as seen by the illustrations given in Fig. V.1. Examples of mesomanufacturing are extremely small motors, bearings, and components for miniature devices, such as hearing aids, stents, heart valves, and mechanical watches, with components the same as the gear shown in Fig. V.1.

In **nanomanufacturing**, parts are produced at scales of one billionth of a meter, typically between 10^{-6} and 10^{-9} m in length; many of the features in integrated circuits are at this length scale. *Biomanufacturing*, covering such areas as molecularly engineered pharmaceutical products, genetic testing, gene therapy, and agricultural products, are at nanoscale level. It is now recognized that many physical and biological processes act at this scale and thus nanomanufacturing holds much promise for future innovations.

In Chapter 28, manufacturing of silicon wafers and microelectronic devices is described, which include a wide variety of computer processors, memory devices, and integrated circuits. Controls, transportation, communications, engineering design and manufacturing, medicine, and entertainment all have been changed greatly by the wide availability of **metal-oxide-semiconductor** (MOS) *devices*, generally based on **single-crystal silicon**. Microelectronics are the best known and commercially important example of *Micromanufacturing*, with some aspects of the applications exemplifying nanomanufacturing. The chapter also covers the techniques used in packaging and assembling integrated circuits onto printed circuit boards.

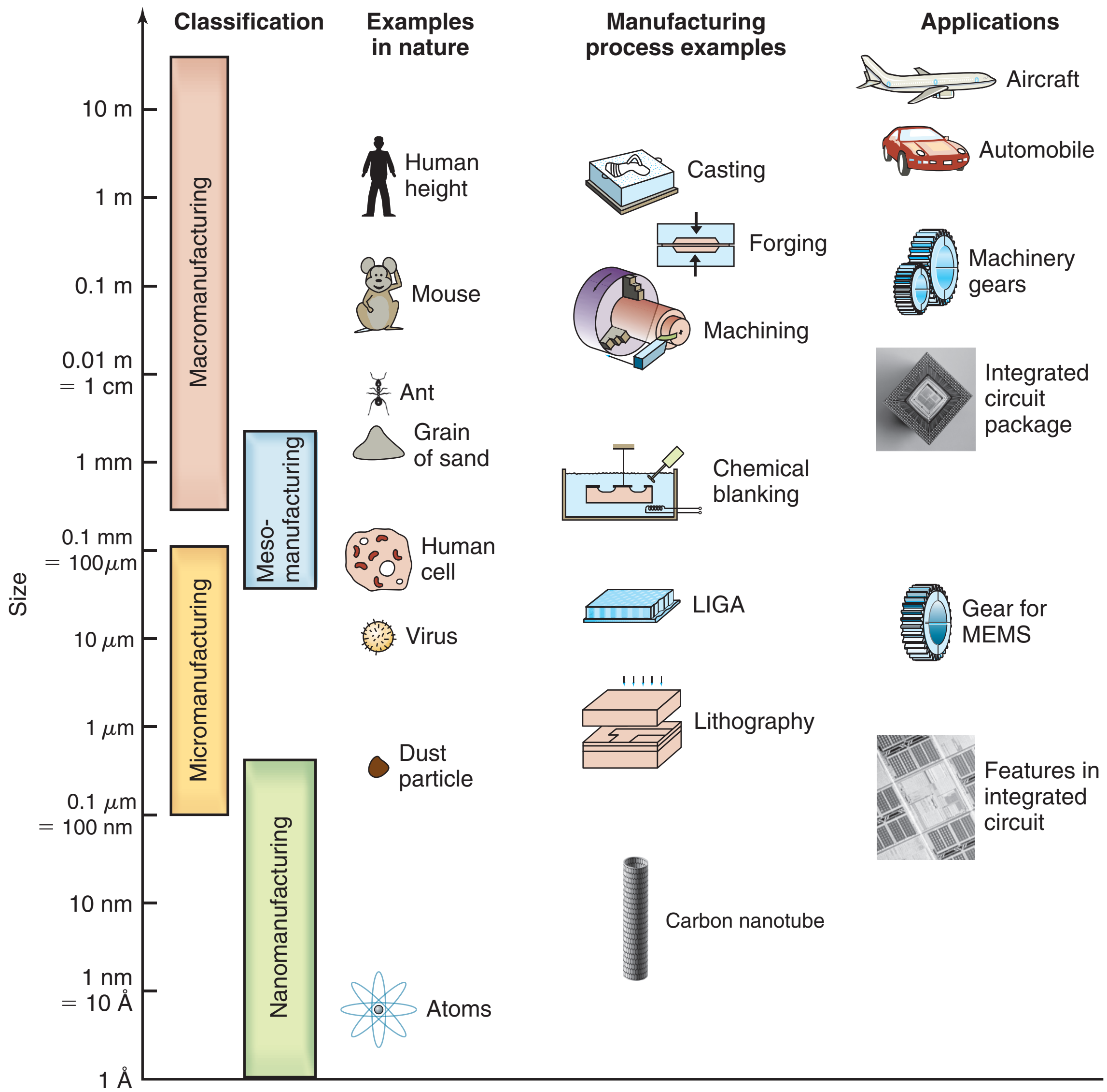


Figure V.1: Illustration of the regimes of macro-, meso-, micro-, and nanomanufacturing, the range of common sizes of parts, and the capabilities of manufacturing processes in producing those parts.

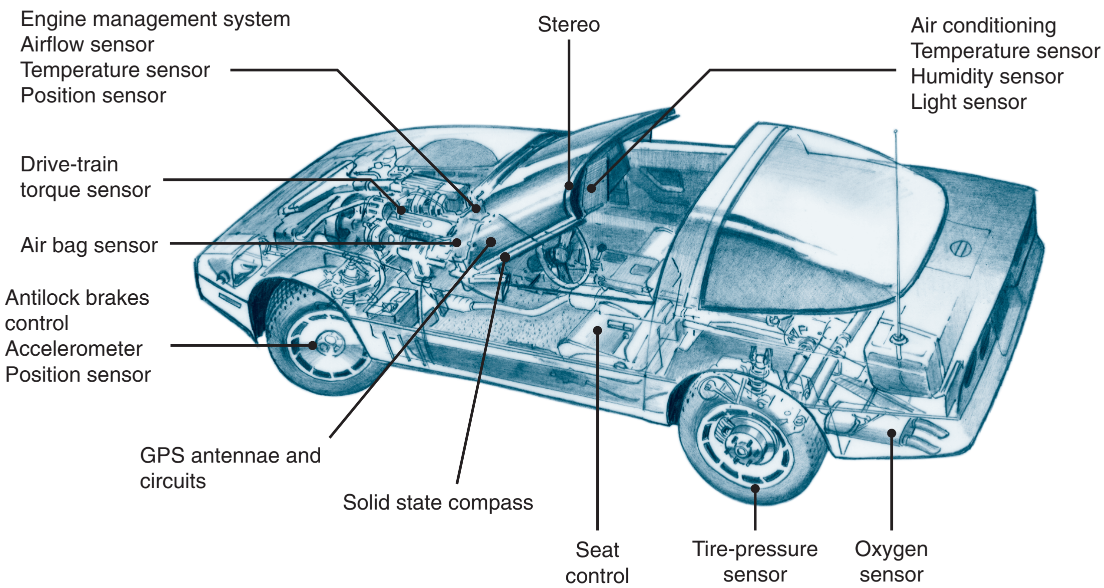


Figure V.2: Microelectronic and microelectromechanical devices and parts used in a typical automobile.

Production of microscale devices that are mechanical and electrical in nature is described in Chapter 29. Depending on their level of integration, these devices are called **micromechanical devices** or **microelectromechanical systems** (MEMS). While the historical origins of MEMS manufacturing stem from the same processes used for microelectronic systems and from identical processes and production sequences still in use, several unique approaches have been developed.

Chapter 28

Fabrication of Microelectronic Devices

- This chapter presents the science and the technologies involved in the production of integrated circuits, a product that has fundamentally changed society.
- It begins by describing silicon, the preferred host material for most integrated circuits, and its unique properties that make it attractive. Beginning with a cast ingot, the operations required to produce a wafer are then described.

- The production of patterns on wafers is next described, including the processes of lithography, wet and dry etching, and doping.
- Metallization and testing are then described, including approaches for obtaining electrical connections from integrated circuits to circuit boards.
- The different packages used for integrated circuits are described, including different mounting methods and materials.
- The chapter concludes with a discussion of flexible electronics and hybrid electronics, including manufacturing approaches such as roll-to-roll printing.

Typical parts produced: Computer processors, memory chips, printed circuit boards, and integrated circuits of all types.

28.1 Introduction

Although semiconducting materials have been used in electronics for a long time (the word **semiconductor** first appeared in 1838), it was the invention of the **transistor** in 1947 that set the stage for what would become one of the greatest technological achievements in all of history. **Microelectronics** has played an increasing role ever since the **integrated circuit** (IC) technology became the foundation for calculators, wrist watches, controls for home appliances and automobiles, information systems, telecommunications, robotics, space travel, weaponry, personal computers, cell phones, and tablets.

The major advantages of today's ICs are their very small size and low cost. As their fabrication technology has become more advanced, the size and cost of such devices as transistors, diodes, resistors, and capacitors continue to decrease, and the global market has become highly competitive. More and more components can now be placed onto a **chip**, a very small piece of semiconducting material on which the circuit is fabricated.

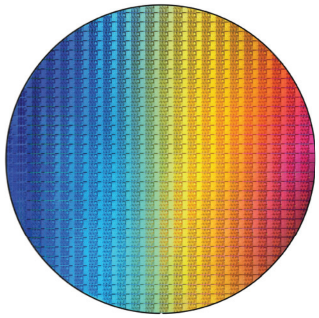
Typical chips produced today have sizes as small as 0.5 mm $\times$ 0.5 mm and, in some rare cases, can be more than 50 mm $\times$ 50 mm, if not an entire wafer. New technologies now allow densities in the range of 10 million devices per chip (Fig. 28.1), a magnitude that has been called **very large scale integration** (VLSI). Some of the advanced ICs may contain more than 100 million devices, called **ultralarge-scale integration** (ULSI). The Intel Itanium® processors, for example, has surpassed 2 billion transistors, and the Advanced Micro Devices Tahiti® graphic processing unit has surpassed 4.3 billion transistors.

This chapter describes the processes currently in use in the fabrication of microelectronic devices and integrated circuits, following the basic sequence shown in Fig. 28.2. The major steps in fabricating a **metal-oxide-semiconductor field-effect transistor** (MOSFET), one of the dominant devices used in modern IC technology, are shown in Fig. 28.3.

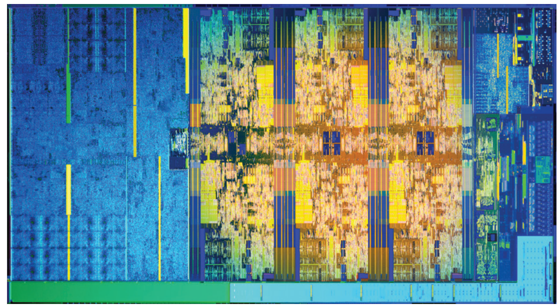
28.2 Clean Rooms

Clean rooms are essential for the production of integrated circuits, a fact that can be appreciated by noting the scale of manufacturing to be performed. Integrated circuits are typically a few mm in length, and the smallest features in a transistor on the circuit may be as small as a few tens of nm. This size range is smaller than particles that generally are not considered harmful, such as dust, smoke, and perfume. However, if these contaminants are present on the surface of a silicon wafer during its processing, they can seriously compromise the performance of the entire device.

There are several levels of clean rooms, defined by the **class** of the room. The size and the number of particles are significant in defining the class of a clean room (Fig. 28.4). The traditional classification



(a)



(b)

(c)

(d)

Figure 28.1: The 8th generation Intel® i7 core processor using 14 nm technology. (a) A 300-mm wafer with a large number of i7 core processor dies fabricated onto its surface; (b) detail view of an Intel i7 core die; (c) image of the Intel i7 die in package; and (d) motherboard of the Intel Hades Canyon Next Unit of Computing (NUC). The large IC package in the middle is an Intel core i7 processor. *Source:* Courtesy of Intel Corporation.

system refers to the number of 0.5-μm or larger particles within a cubic foot of air; thus, a Class-10 clean room has 10 or fewer such particles per cubic foot. Although this standard has now been superseded by an ISO standard, the traditional classification scheme is still used. Most clean rooms for microelectronics manufacturing range from Class 1 to Class 10. In comparison, the contamination level in modern hospitals is on the order of 10,000 particles per cubic foot.

For controlled atmospheres that are free from particulate contamination, all ventilating air is passed through a *high-efficiency particulate air* (HEPA) filter. In addition, the air usually is conditioned so that it is at 21°C and 45% relative humidity.

The largest sources of contaminants in a clean room are the workers themselves. Skin particles, hair, perfume, makeup, clothing, bacteria, and viruses are given off naturally by humans and in sufficiently large numbers to quickly compromise a Class-100 clean room. For these reasons, most clean rooms require special coverings, such as white laboratory coats, gloves, and hairnets, as well as avoiding perfumes and makeup. The most stringent clean rooms require full-body coverings, called *clean room smocks*. Other precautions include using (a) a ballpoint pen, instead of a pencil, to avoid objectionable graphite particles from pencils and (b) special clean-room paper, to prevent the accumulation of paper particles in the air.

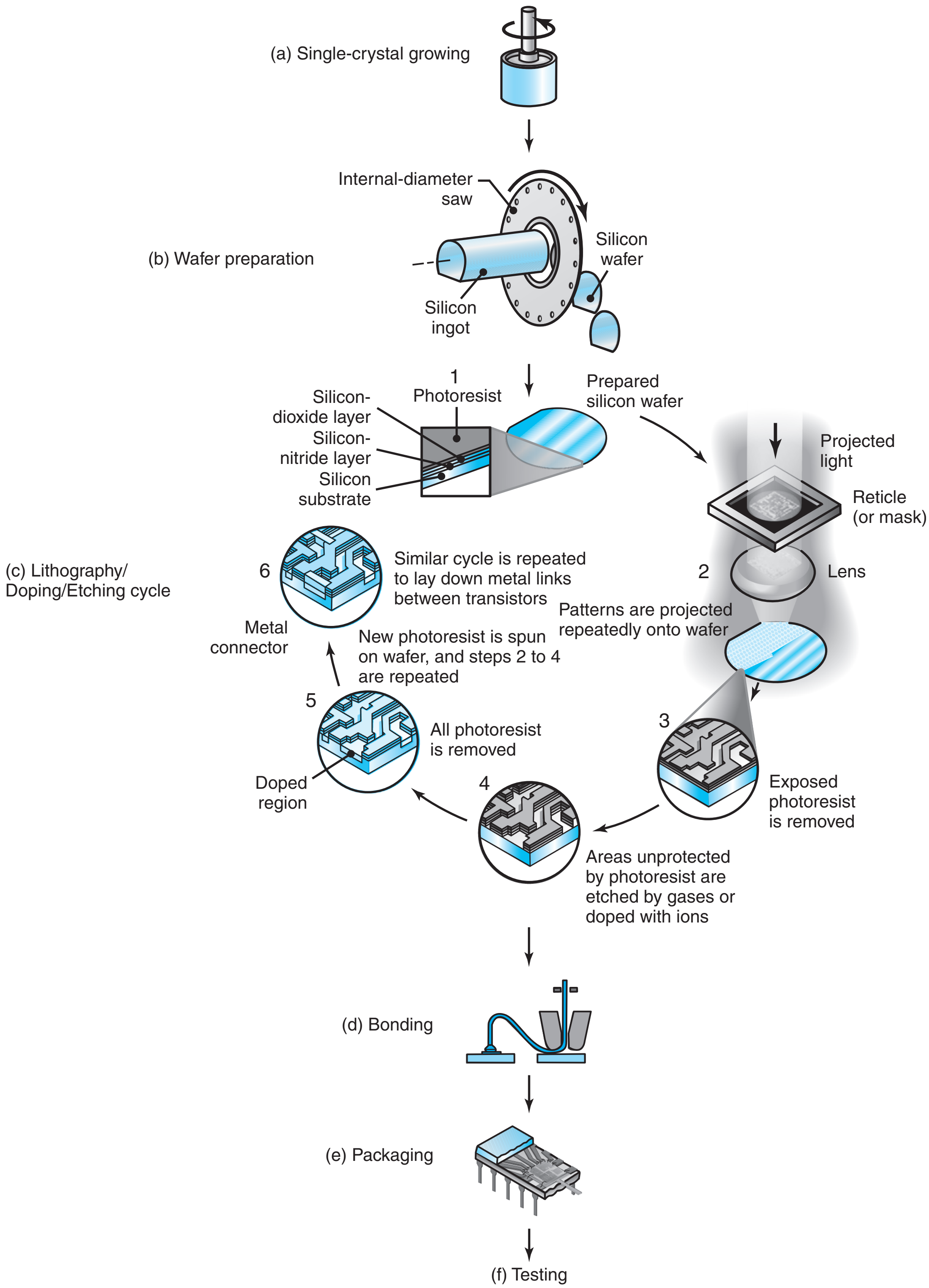


Figure 28.2: Outline of the general fabrication sequence for integrated circuits.

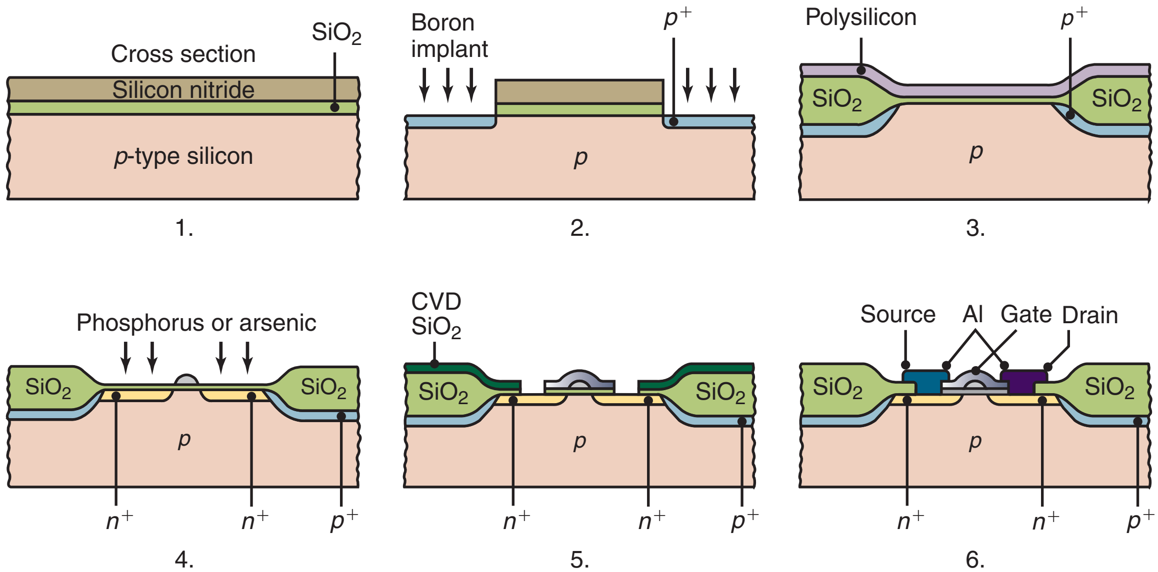


Figure 28.3: Cross-sectional views of the fabrication of a MOSFET transistor. *Source:* After R.C. Jaeger.

Clean rooms are designed such that the cleanliness at critical *processing areas* is better than in the clean room in general. This is accomplished by always directing the filtered air *from top down* in the clean rooms and from *floor to ceiling* in the service aisle, a goal that can be facilitated by laminar-flow hooded work areas. To minimize defects, no product is allowed in the service aisle.

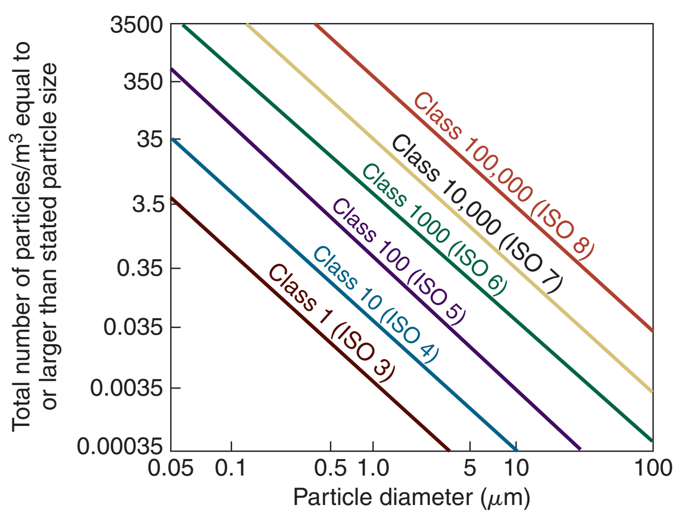


Figure 28.4: Allowable particle size counts for various clean-room classes.

28.3 Semiconductors and Silicon

As the name suggests, **semiconductor materials** have electrical properties that lie between those of conductors and insulators, and that they exhibit resistivities between 10^{-3} and 10^{8} Ω-cm. Semiconductors have become the foundation for electronic devices because their electrical properties can be altered, when controlled amounts of selected impurity atoms are added to their crystal structures. The impurity atoms, also known as **dopants**, have either one more valence electron (n-type, or negative, dopant) or one less valence electron (p-type, or positive, dopant) than the atoms in the semiconductor lattice.

For silicon, which is a Group IV element in the Periodic Table, typical n-type and p-type dopants include, respectively, phosphorus (Group V) and boron (Group III). The electrical operation of semiconductor devices can thus be controlled through the creation of regions with different doping types and concentrations.

The earliest electronic devices were fabricated on *germanium*, but **silicon** has become the industry standard. The abundance of its alternative forms in the crust of the Earth is second only to that of oxygen, making silicon economically attractive. Its main advantage over germanium is its large energy gap (1.1 eV), as compared with that of germanium (0.66 eV). This energy gap allows silicon-based devices to operate at temperatures of about 150°C higher than devices fabricated on germanium, which operate at about 100°C.

Another important processing advantage of silicon is that its oxide (*silicon dioxide*, SiO_2) is an excellent electrical *insulator*, and can be used for both isolation and passivation (see Section 3.8) purposes. By contrast, germanium oxide is water soluble and thus unsuitable for electronic devices. Moreover, the oxidized form of silicon allows the production of **metal-oxide-semiconductor** (MOS) devices, which are the basis for MOS transistors. These materials are used in memory devices, processors, and other devices, and are, by far, the largest volume of semiconductor material produced worldwide.

Structure of Silicon. The crystallographic structure of silicon is a diamond-type fcc structure (Fig. 28.5), along with the *Miller indices* of an fcc material. Miller indices are a notation for identifying planes and

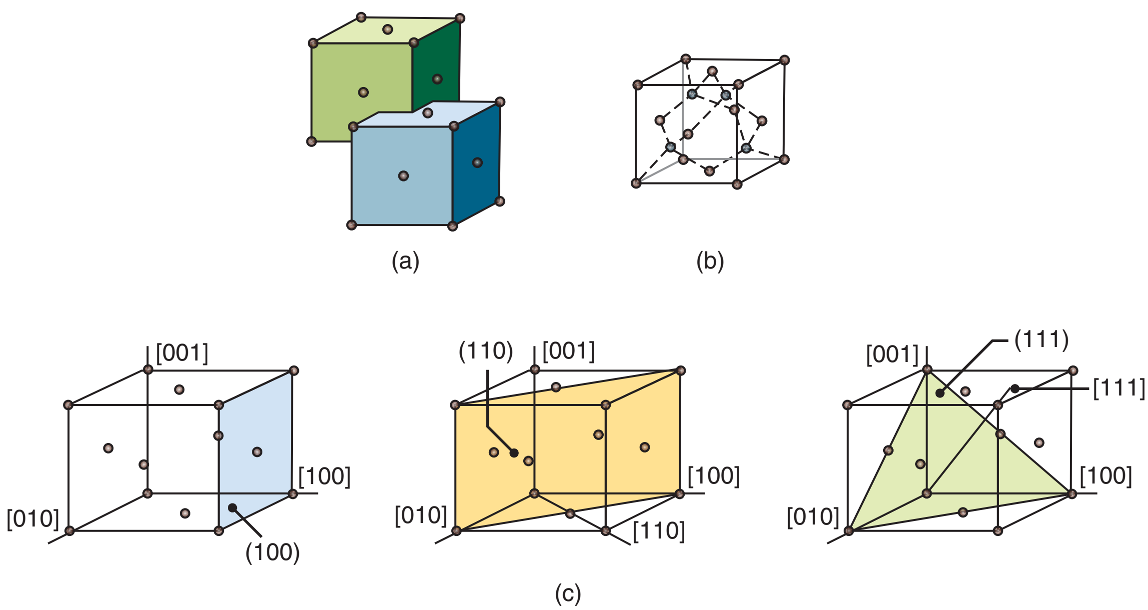


Figure 28.5: Crystallographic structure and Miller indices for silicon. (a) Construction of a diamond-type lattice from interpenetrating face-centered cubic cells; one of eight penetrating cells is shown. (b) Diamond-type lattice of silicon; the interiors have been shaded in color. (c) Miller indices for a cubic lattice.

directions within a unit cell (Section 1.3). A crystallographic plane is defined by the reciprocal of its intercepts with the three axes. Because anisotropic etchants (Section 28.8.1) preferentially remove material in certain crystallographic planes, the orientation of the silicon crystal in a wafer is an important consideration.

The relatively low operating temperature of silicon devices has encouraged the development of *compound semiconductors*, specifically **gallium arsenide**. Its major advantage over silicon is its ability to emit light, thus allowing the fabrication of such devices as *lasers* and *light-emitting diodes* (LEDs).

Devices fabricated on gallium arsenide also have much higher operating speeds than those fabricated on silicon, and thus it can be found in high-frequency communication systems. Some of gallium arsenide's limitations, on the other hand, are its considerably higher cost, greater processing complications, and, most critically, the difficulty of growing high-quality oxide layers, the need for which is emphasized throughout the rest of this chapter.

28.4 Crystal Growing and Wafer Preparation

Silicon occurs naturally in the forms of *silicon dioxide* and various *silicates*. It must, however, undergo a series of purification steps in order to become the high-quality, defect-free, single-crystal material that is necessary for semiconductor device fabrication. The purification process begins by heating silica and carbon together in an electric furnace (Fig. 5.2), resulting in a 95 to 98% pure polycrystalline silicon. This material is then converted to an alternative form, commonly *trichlorosilane* (a compound of silicon, hydrogen, and chlorine), which is then purified and decomposed in a high-temperature hydrogen atmosphere. The resulting product is extremely high-quality *electronic-grade silicon* (EGS).

Single-crystal silicon usually is obtained through the *Czochralski* or **CZ process** (Section 11.5). It utilizes a *seed* crystal that is dipped into a silicon melt, and is then pulled out slowly while being rotated. At this point, controlled amounts of impurities can be added to obtain a uniformly doped crystal. The result is a cylindrical single-crystal ingot, typically 100–300 mm in diameter and over 1 m in length. Because this technique does not allow for exact control of the ingot diameter, ingots are grown a few mm larger than the required size; they are then ground to the desired diameter. Silicon **wafers** are produced from silicon ingots by a sequence of machining and finishing operations (Fig. 28.6).

Next, the crystal is sliced into individual wafers, using an inner-diameter diamond-encrusted blade (Fig. 24.25f), whereby a rotating, ring-shaped blade with its cutting edge on the inner diameter of the ring is utilized. While the substrate depth required for most electronic devices is no more than several microns, wafers typically are cut to a thickness of about 0.5 mm. This thickness provides the physical support necessary for the absorption of temperature variations and the mechanical support needed during subsequent fabrication of the wafer.

The wafer is then ground along its edges using a diamond wheel; this operation gives the wafer a rounded profile which is more resistant to chipping. Finally, the wafers are polished and cleaned, to remove surface damage caused by sawing. This operation is commonly performed by *chemical–mechanical polishing*, also referred to as *chemical–mechanical planarization* (Section 26.7).

In order to properly control the production process, it is important to determine the orientation of the crystal in a wafer, which is done by notches or flats machined into them for identification (Fig. 28.7). Most commonly, the (100) or (111) plane of the crystal defines the wafer surface, although (110) surfaces also can be used for micromachining applications (Section 29.2). Wafers are also identified by a laser *scribe* mark, produced by the particular manufacturer. Laser scribing of information may take place on the front or on the back side of the wafer. The front side of some wafers has an exclusion edge area, 3 to 10 mm in size reserved for the scribe information, such as lot numbers, orientation, and a wafer identification code unique to the particular manufacturer.

Wafers are typically processed in lots of 25 or 50, with 150 to 200 mm diameters each, or lots of 12 to 25 with 300-mm diameters each. In this way, they can easily be handled and transferred during subsequent processing steps. Thousands of individual circuits can be placed on one wafer because of the small device size and large wafer diameter. Once processing is completed, the wafer is then diced into individual **chips**, each containing one complete integrated circuit.

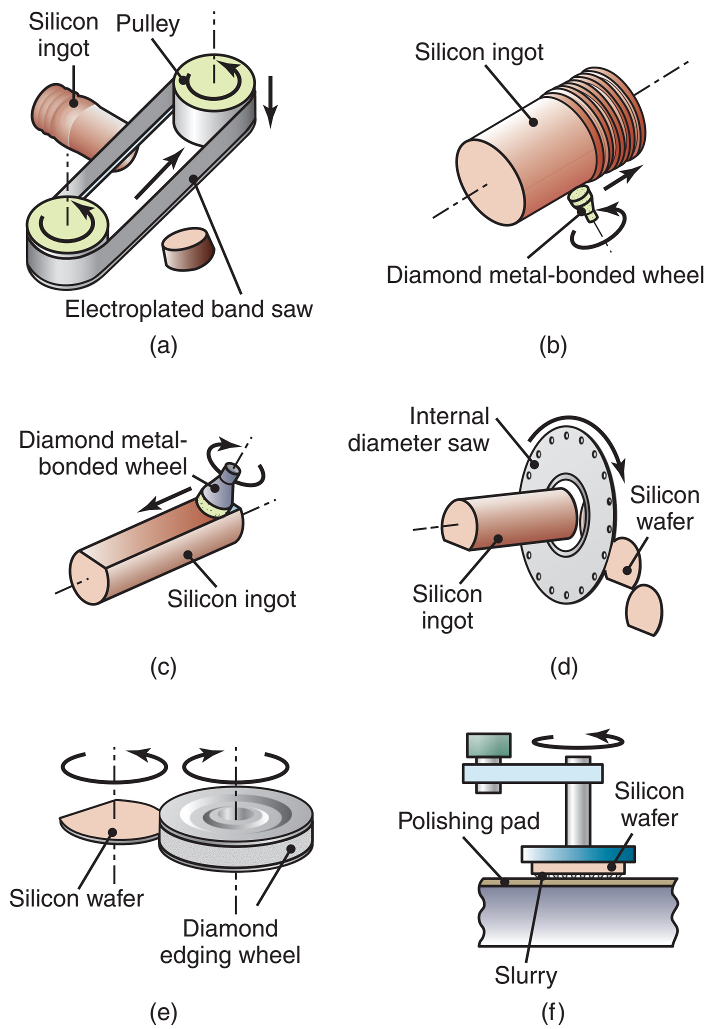


Figure 28.6: Finishing operations on a silicon ingot to produce wafers: (a) sawing the ends off the ingot; (b) grinding of the end and cylindrical surfaces of a silicon ingot; (c) machining of a notch or flat; (d) slicing of wafers; (e) end grinding of wafers; (f) chemical–mechanical polishing of wafers.

At this point, the single-crystal silicon wafer is ready for the fabrication of the integrated circuit or device. Fabrication takes place over the entire wafer surface, thus many chips are produced at the same time (Fig. 28.1a). Because the number of chips that can be produced depends on the surface area of the wafer, advanced-circuit manufacturers have moved towards using larger single-crystal solid cylinders, and 300-mm diameter wafers now are common.

28.5 Film Deposition

Films are used extensively in microelectronic-device processing, particularly for insulating and conducting types. Commonly deposited films include polysilicon, silicon nitride, silicon dioxide, and tungsten, titanium, and aluminum. In some cases, the wafers merely serve as a mechanical support, on which custom *epitaxial layers* are grown (see below).

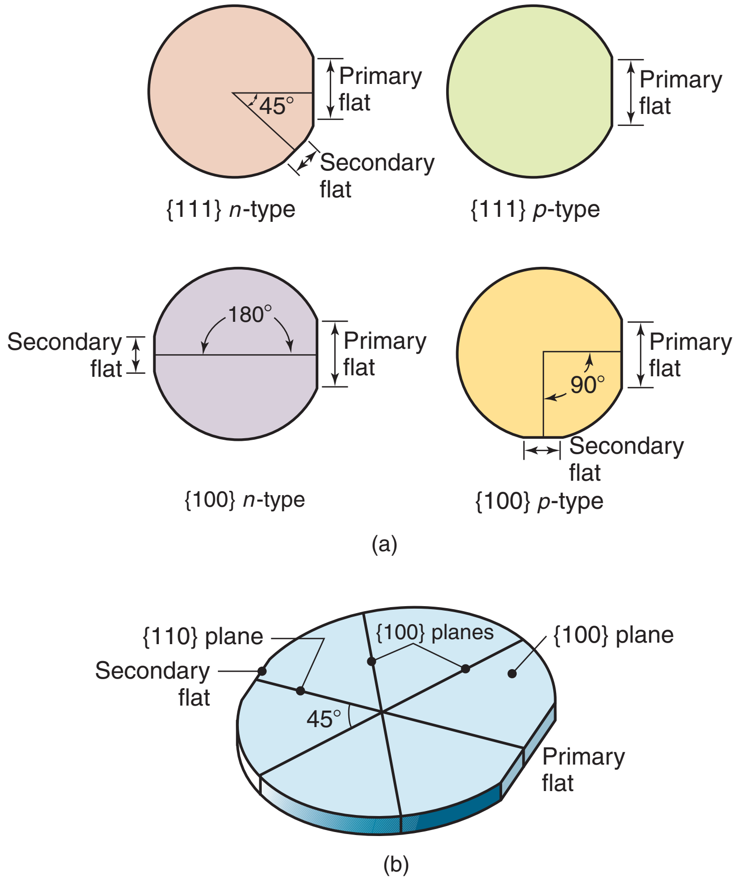


Figure 28.7: Identification of single-crystal wafers of silicon. This identification scheme is common for 150-mm diameter wafers, but notches are more common for larger wafers.

Epitaxy. Defined as the growth of a vapor deposit, *epitaxy*, or *electrodeposit*, occurs when the crystal orientation of the deposit is related directly to the crystal orientation in the underlying crystalline substrate. The advantages of processing on these deposited films, instead of on the actual wafer surface, include fewer impurities (especially carbon and oxygen), improved device performance, and tailoring of material properties (which cannot be done on the wafers themselves).

Some of the major functions of deposited films are **masking** and protecting the semiconductor surface. In masking, the film must both inhibit the passage of dopants and concurrently display a capability to be etched into patterns of high resolution. Upon completion of device fabrication, films are applied to protect the underlying circuitry. Films used for masking and protecting include silicon dioxide, phosphosilicate glass (PSG), and silicon nitride. Each of these materials has distinct advantages, and often are used in combination.

Conductive films are used primarily for device interconnection. These films must have low electrical resistivity, be capable of carrying large currents, and be suitable for connection to terminal packaging leads with wire bonds. To date, aluminum is the more common material, because it can be dry etched more easily; however, copper, with its low resistivity, is used for high-performance processors. Its main drawback is associated with difficulties in plasma etching, requiring development of alternate deposition and patterning techniques. Increasing circuit complexity has required up to 10 levels of conductive layers, all of which must be separated by insulating films.

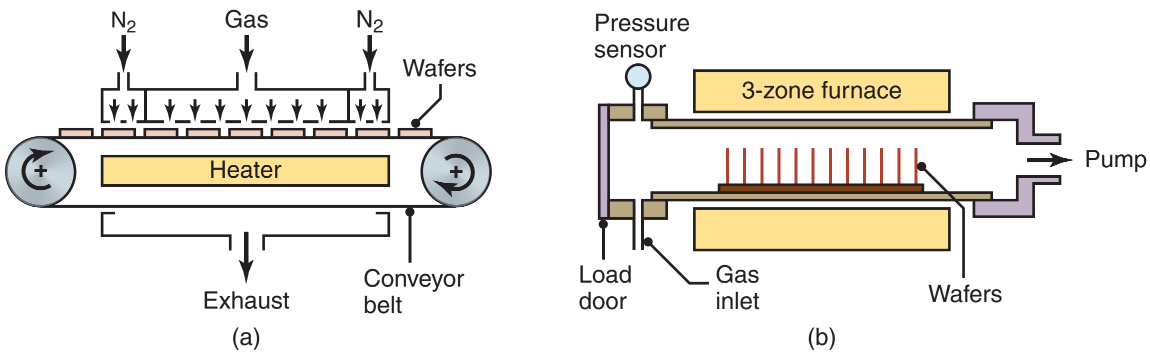


Figure 28.8: Schematic diagrams of (a) a continuous, atmospheric-pressure CVD reactor and (b) a low-pressure CVD. *Source:* After S.M. Sze.

Film Deposition. Films can be *deposited* by several techniques, involving a variety of pressures, temperatures, and vacuum systems (see also Chapter 34):

- One of the oldest and simplest methods is **evaporation**, used primarily for depositing metal films. The metal is first heated in a vacuum to its point of vaporization; upon evaporation, it forms a thin layer on the substrate surface. The heat of evaporation is usually generated by a heating filament or electron beam.

- **Sputtering** involves bombarding a target in a vacuum with high-energy ions, usually argon (Ar^+). Sputtering systems generally include a DC power supply to produce the energized ions. As the ions impinge on the target, atoms are knocked off and are subsequently deposited on wafers mounted within the system. Although some argon may be trapped within the film, sputtering results in highly uniform coverage. Recent advances include using a radio-frequency power source (**RF sputtering**) and introducing magnetic fields (**magnetron sputtering**).

- In one of the most common techniques, **chemical-vapor deposition** (CVD), film is deposited by way of the reaction and/or decomposition of gaseous compounds. Using this technique, silicon dioxide is deposited routinely by the oxidation of silane or a chlorosilane. Figure 28.8a shows a continuous CVD reactor that operates at atmospheric pressure.

  A similar method that operates at lower pressures is **low-pressure chemical-vapor deposition** (LPCVD), shown in Fig. 28.8b. Capable of coating hundreds of wafers at a time, this method has a much higher production rate than that of atmospheric-pressure CVD, and it also provides superior film uniformity and with less consumption of carrier gases. The technique is commonly used for depositing polysilicon, silicon nitride, and silicon dioxide.

- **Plasma-enhanced chemical-vapor deposition** (PECVD) involves the processing of wafers in an RF plasma containing the source gases. This method has the advantage of maintaining a low wafer temperature during deposition.

Atomic Layer Deposition (ALD) is a class of CVD in which the reactant gases are introduced to the deposition system in an alternating, non-overlapping fashion with purge cycles in between each reactant cycle. Each reactant cycle self-limits once all the host sites on the wafer surface have been consumed. In this manner, a slow but very controllable film growth occurs. The thickness of the film is well controlled and determined by the number of cycles (reactant A, purge, reactant B, purge, reactant A, etc.). In addition, ALD films are highly conformal, and closely follow the contour of the underlying wafer topology.

Silicon **epitaxy** layers, in which the crystalline layer is formed using the substrate as a seed crystal can be grown by a variety of methods. If the silicon is deposited from the gaseous phase, the process is known as **vapor-phase epitaxy** (VPE). In another variation, called **liquid-phase epitaxy** (LPE), the heated substrate is brought into contact with a liquid solution containing the material to be deposited. Another high-vacuum process, **molecular-beam epitaxy** (MBE), utilizes evaporation to produce a thermal beam of molecules that are deposited on the heated substrate, resulting in a very high degree of purity. In addition, since the films are grown one atomic layer at a time, it is possible to have excellent control over doping profiles, important especially in gallium-arsenide technology. However, the MBE process has relatively low growth rates as compared to other conventional film-deposition techniques.

28.6 Oxidation

The term *oxidation* refers to the growth of an oxide layer as a result of the reaction of oxygen with the substrate material; oxide films also can be formed by deposition techniques. Thermally grown oxides, described in this section, display a higher level of purity than do deposited oxides, because they are grown directly from the high-quality substrate. However, deposition methods must be used if the composition of the desired film is different from that of the substrate material.

Silicon dioxide is the most widely used oxide in IC technology today, and its excellent characteristics are one of the major reasons for the widespread use of silicon. Aside from its effectiveness in dopant masking and device isolation, silicon dioxide's most critical role is that of the *gate oxide* material. Silicon surfaces have an extremely high affinity for oxygen; thus, a freshly sawed slice of silicon will quickly grow a native oxide of 3 to 4 nm in thickness. Modern IC technology requires oxide thicknesses ranging from a few to a few hundred nm.

- **Dry oxidation** is accomplished by elevating the substrate temperature, typically to about 750° to 1100°C, in an oxygen-rich environment. As a layer of oxide forms, the oxidizing agents must be capable of passing through the oxide and reach the silicon surface, where the actual reaction takes place. Thus, an oxide layer does not continue to grow on top of itself, but rather, it grows from the silicon surface outward. Some of the silicon substrate is consumed in the oxidation process (Fig. 28.9).

  The ratio of oxide thickness to the amount of silicon consumed is found to be 1:0.44. Thus, to obtain an oxide layer 100 nm thick, approximately 44 nm of silicon will be consumed. This requirement does not present a problem, as substrates are always grown sufficiently thick. One important effect of silicon consumption is the rearrangement of dopants in the substrate near the interface. Because different impurities have different segregation coefficients, or mobilities, in silicon dioxide, some dopants become depleted away from the oxide interface while others pile up there. Consequently, the processing parameters must be properly adjusted to compensate for this effect.

- **Wet oxidation** utilizes a water-vapor atmosphere as the agent. This method results in a considerably higher growth rate than that of dry oxidation; however, it suffers from a lower oxide density and, therefore, a lower dielectric strength. The common practice is to combine both the dry and wet oxidation methods, by growing an oxide in a three-part layer: dry–wet–dry. This approach combines the advantages of wet oxidation's much higher growth rate and dry oxidation's high quality.

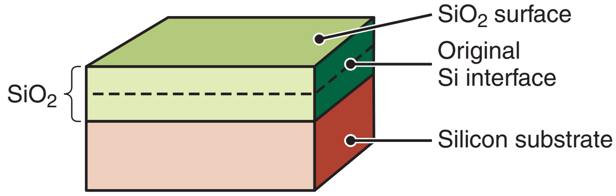


Figure 28.9: Growth of silicon dioxide, showing consumption of silicon. *Source:* After S.M. Sze.

The two oxidation methods described are useful primarily for coating the entire silicon surface with oxide; however, it also may be necessary to oxidize only certain portions of the surface. The procedure is called **selective oxidation**; it uses silicon nitride, which inhibits the passage of oxygen and water vapor. Thus, by covering certain areas with silicon nitride, the silicon under these areas remains unaffected while the uncovered areas are oxidized.

28.7 Lithography

Lithography is the process by which the geometric patterns that define devices are transferred to the substrate surface. A summary of lithographic techniques is given in Table 28.1, and a comparison of the basic lithography methods is shown in Fig. 28.10. The most common technique used today is **photolithography**. *Electron-beam* and *X-ray lithography* are of great interest because of their ability to transfer patterns with higher resolution, a necessary feature for the increased miniaturization of integrated circuits.

Photolithography. Photolithography uses a **reticle**, also called a **mask** or **photomask**, which is a glass or quartz plate with a pattern of the chip deposited onto it with a chromium film. The reticle image can be the same size as the desired structure on the chip, but it is often an enlarged image, usually 4× to 20× larger, although 10× magnification is the most common. The enlarged images are then focused onto a wafer through a lens system, an operation referred to as *reduction lithography*.

In current practice, the lithographic process is applied to each microelectronic circuit as many as 75 times, each time using a different reticle to define the different areas of the working devices or their interconnection. Typically designed at several thousand times their final size, reticle patterns undergo a series of reductions before being applied permanently to a defect-free quartz plate. Computer-aided design (CAD; Section 38.4) has had a major impact on reticle design and generation.

Cleanliness is especially important in lithography, and manufacturers now use robotics and specialized wafer-handling apparati in order to minimize contamination from dust and dirt. Once the film deposition process is completed and the desired reticle patterns have been generated, the wafer is cleaned and coated with a **photoresist** (PR), an organic polymer.

A photoresist consists of three principal components:

1. A polymer, that changes its structure when exposed to radiation.

2. A sensitizer, that controls the reactions in the polymer.

3. A solvent, to deliver the polymer in liquid form.

Photoresist layers 0.5 to 2.5 μm thick are produced by applying the photoresist to the substrate and then spinning it at several thousand rpm for 30 or 60 seconds to give uniform coverage (Fig. 28.11).

The next step in photolithography is **prebaking** the wafer to remove the solvent from the photoresist and harden it. This step is carried out on a hot plate, heated to around 100°C. The pattern is transferred

Table 28.1: General Characteristics of Lithography Techniques.

Method	Wavelength (nm)	Finest feature size (nm)
Ultraviolet (Photolithography)	365	350
Deep UV	248	250
Extreme UV	10–20	30–100
X-ray	0.01–1	20–100
Electron beam	—	80
Immersion	193	11

Source: After P.K. Wright.

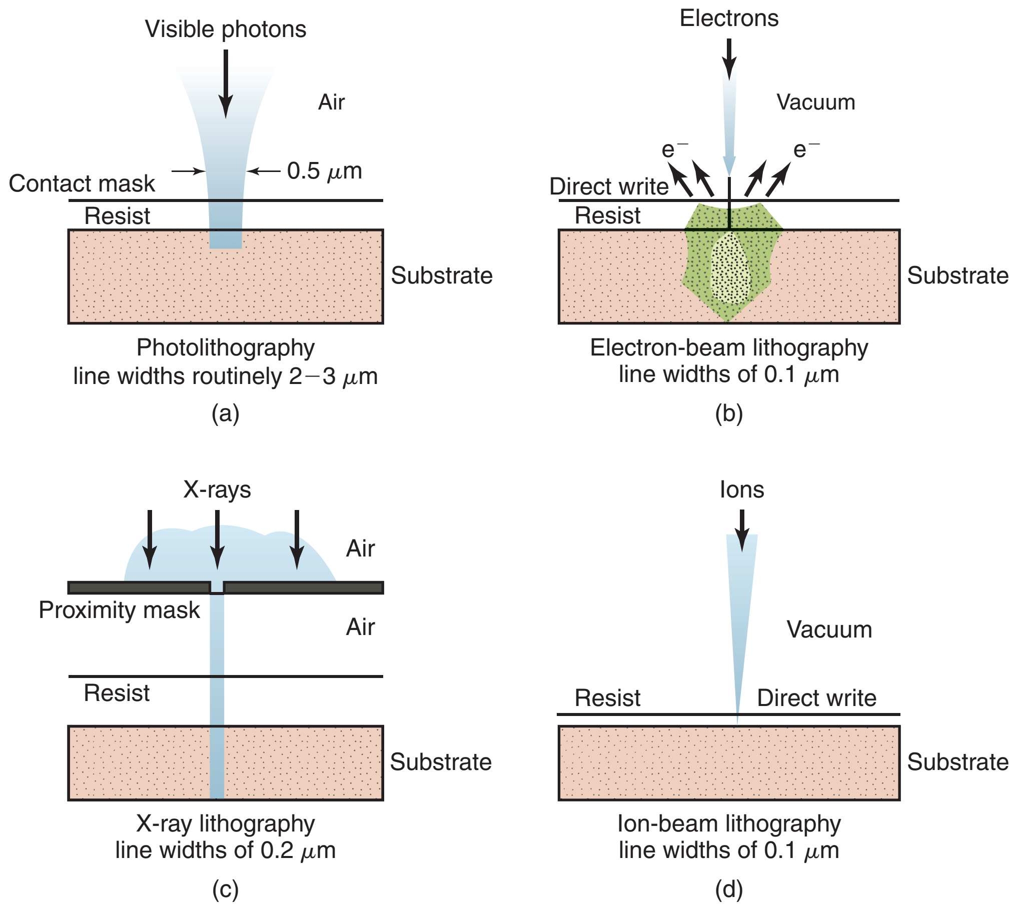


Figure 28.10: Comparison of lithography techniques. Note that (a) and (c) involve masking to achieve pattern transfer, while (b) and (d) scribe the pattern without a mask, known as *direct writing*.

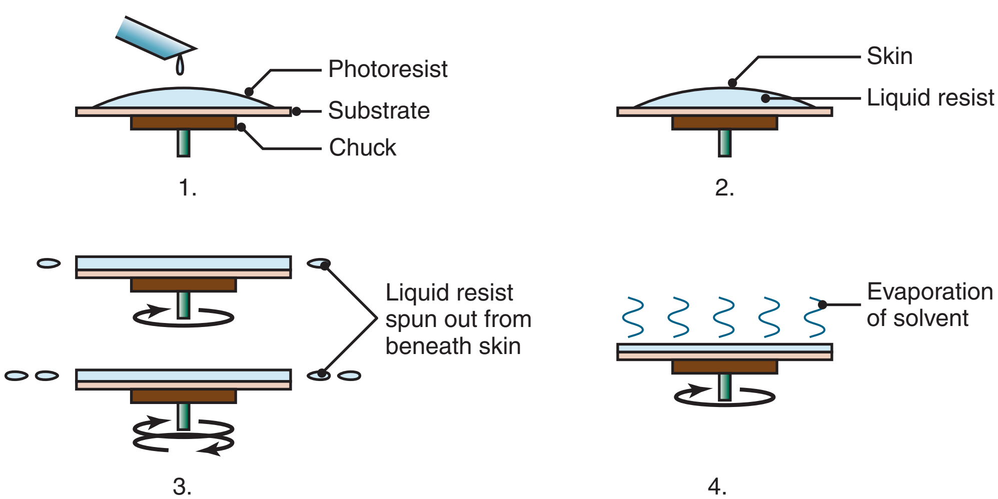


Figure 28.11: Spinning of an organic coating on a wafer.

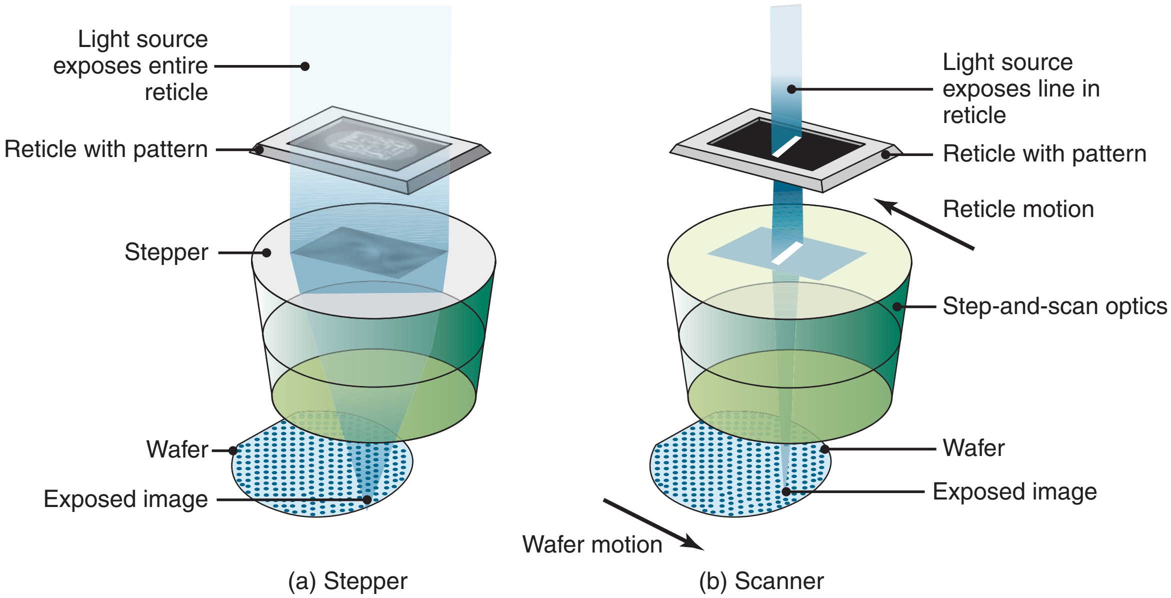


Figure 28.12: Schematic illustration of (a) wafer stepper technique for pattern transfer, and (b) wafer step-and-scan technique.

to the wafer through *stepper* or *step-and-scan* systems. With wafer steppers (Fig. 28.12a), the full image is exposed in one flash, and the reticle pattern is then refocused onto another adjacent section of the wafer. With step-and-scan systems (Fig. 28.12b), the exposing light source is focused into a line, and the reticle and wafer are translated simultaneously in opposite directions to transfer the pattern.

The wafer must be aligned carefully under the desired reticle. In this crucial step, called **registration**, the reticle must be aligned correctly with the previous layer on the wafer. Once the reticle is aligned, it is subjected to exposure. Upon development and removal of the exposed photoresist, a duplicate of the reticle pattern will appear in the photoresist layer. As seen in Fig. 28.13, the reticle can be a negative image or a positive image of the desired pattern. A positive reticle uses the UV radiation to break down the chains in the organic film, so that these films are removed preferentially by the developer. Positive masking is more common than negative masking because the photoresist can swell and distort with negative masking, thus making it unsuitable for small features. Newer negative photoresist materials do not have this problem.

Following the exposure and development sequence, **postbaking** the wafer drives off the solvent, and toughens and improves the adhesion of the remaining resist. In addition, a deep UV treatment (thereby baking the wafer to about 150°C to 200°C in ultraviolet light) can be used to further strengthen the resist against high-energy implants and dry etches. The underlying film not covered by the photoresist is then etched away (Section 28.8) or implanted (Section 28.9).

Following lithography, the developed photoresist must be removed, in a process called **stripping**. In *wet stripping*, the photoresist is dissolved by such solutions as acetone or strong acids; in this method, the solutions tend to lose potency in use. *Dry stripping* involves exposing the photoresist to an oxygen plasma, called **ashing**. Dry stripping has become more common, because it (a) does not involve the disposal of consumed hazardous chemicals, (b) is easier to control, and (c) it can result in exceptional surfaces.

One of the major issues in lithography is **line width**, which is the width of the smallest feature imprintable on the silicon surface, and is called **critical dimension** (CD). As circuit densities have escalated over the years, device sizes and features have become smaller and smaller. Today, the commercially feasible minimum critical dimension is generally under 20 nm. Intel first shipped 14 nm devices in 2014 (see Fig. 28.1), and is aggressively pursuing 10 nm production.

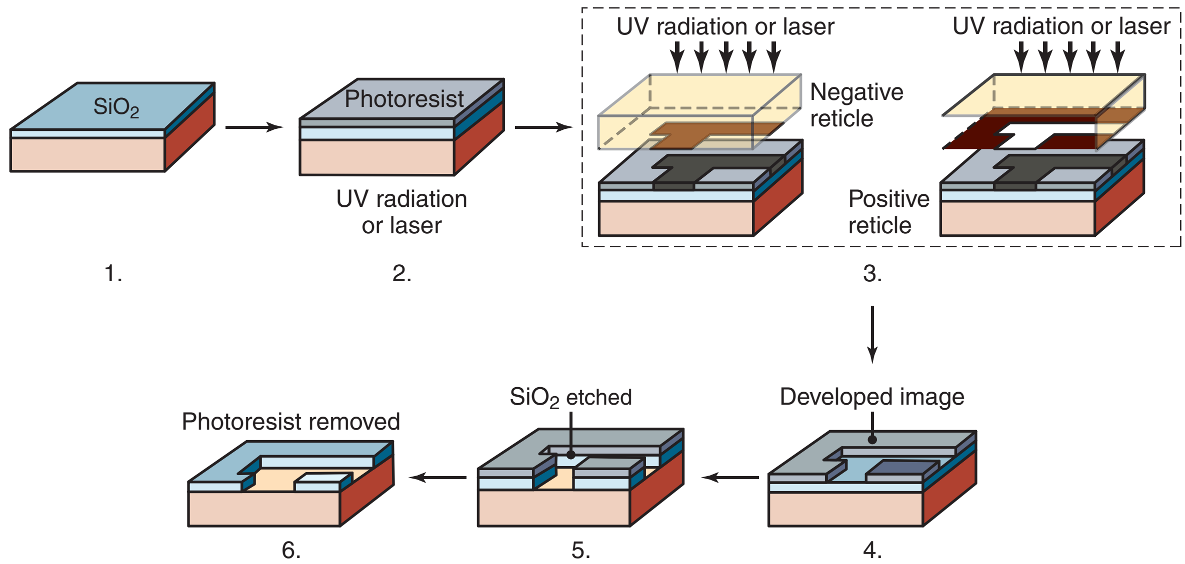


Figure 28.13: Pattern transfer by photolithography; note that the mask in Step 3 can be a positive or a negative image of the pattern.

Because pattern resolution and device miniaturization have been limited by the wavelength of the radiation source used, the need has arisen to move to wavelengths shorter than those in the UV range, such as deep UV wavelengths, extreme UV wavelengths, electron beams, and X-rays. In these technologies, the photoresist is replaced by a similar resist that is sensitive to a specific range of shorter wavelengths.

Pitch Splitting Lithography. Multiexposure techniques have been developed to obtain higher resolution images than can be attained through conventional single-exposure lithography, and has been applied for sub-32 nm feature development. *Pitch splitting* is shown in Fig. 28.14, using conventional lithography in multiple stages. Recognizing that the spacing between features is the limiting dimension, pitch splitting involves breaking up the desired pattern into two complimentary portions and creating corresponding masks. By using two imaging steps, features can be developed in the substrate with twice the resolution of a single imaging step.

There are two forms of pitch splitting. In **double exposure** (DE), a mask exposes some of the desired trenches or regions in the photoresist, then a second mask is used to expose the remaining features (Fig. 28.14a). The photoresist is then exposed and the substrate is etched. **Double patterning** (DP) involves two sequential lithography and etch steps, so that it is sometimes referred to as the **LELE** (lithography-etch-lithography-etch) *process*.

Immersion Lithography. The resolution of lithography systems can be increased by inserting a fluid with a high refractive index between the final lens and the wafer, called *immersion lithography*. Water has been mainly used, and has been the preferred approach used to attain feature sizes below 45 nm. Fluids with a refractive index higher than that of water also are being investigated to increase the resolutions of immersion lithography.

Immersion lithography requires careful process controls, especially thermal controls, since any bubbles that develop in the water will result in defects due to distortion of the light source.

Extreme Ultraviolet Lithography. The pattern resolution in photolithography and immersion lithography is ultimately limited by light diffraction. One of the means of reducing the effects of diffraction is to use ever shorter wavelengths. *Extreme ultraviolet lithography* (EUV) uses light at a wavelength of 13 nm, in order to

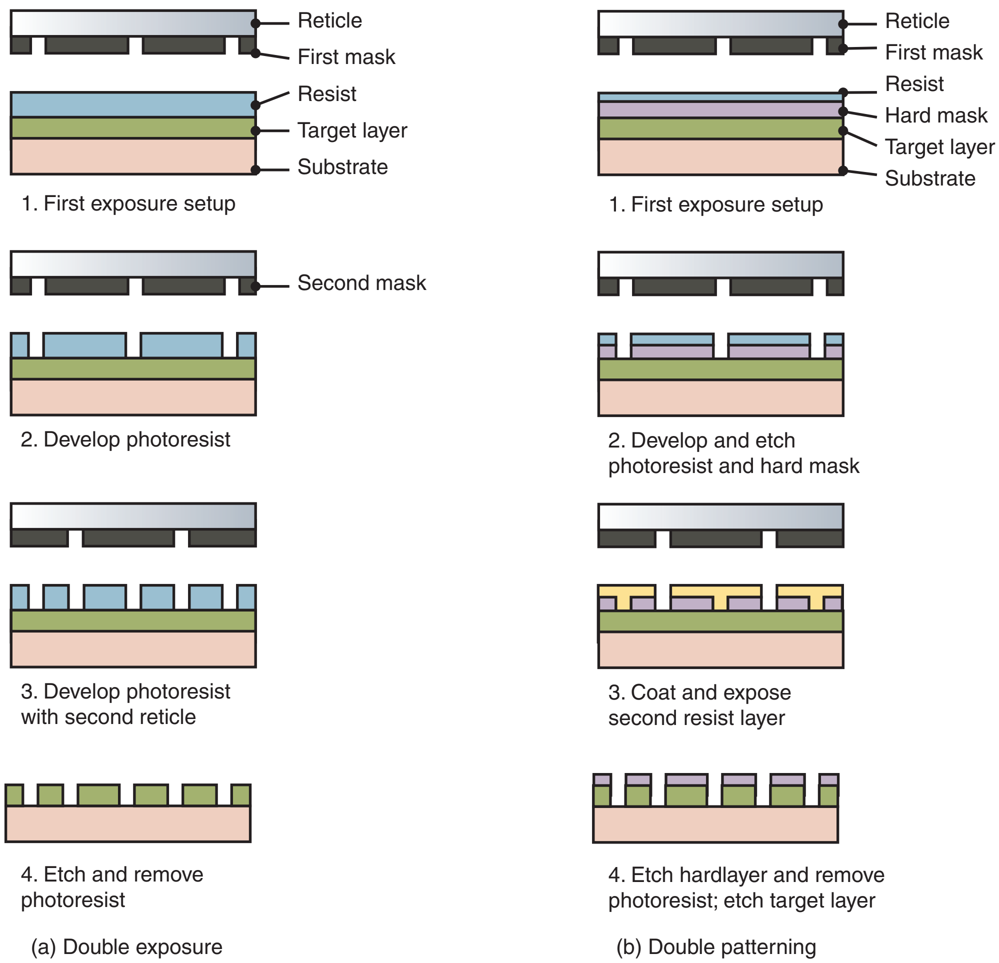


Figure 28.14: Pitch splitting lithography. (a) Double exposure (DE) process; (b) double patterning (DP) process, also known as the LELE (lithography-etch-lithography-etch) process.

obtain features commonly in the range from 30 to 100 nm, but expected to be useful for sub-25 nm features. Because glass lenses absorb some EUV light, the waves are focused through highly reflective molybdenum-silicon mirrors through the mask to the wafer surface.

X-ray Lithography. Although photolithography is the most widely used lithography technique, it has fundamental resolution limitations associated with light diffraction. *X-ray lithography* is superior to photolithography, because of the shorter wavelength of the radiation and its very large depth of focus. These characteristics allow much finer patterns to be resolved, and make X-ray lithography far less susceptible to dust than photolithography. Moreover, the *aspect ratio* (defined as the ratio of depth to lateral dimension) can be higher than 100, whereas it is limited to around 10 with photolithography. However, to achieve this benefit, synchrotron radiation is required, which is expensive and available at only a few research laboratories.

Given the large capital investment required for a manufacturing facility, industry has preferred to refine and improve optical lithography, instead of investing new capital into X-ray-based production. Currently, X-ray lithography is not widespread, although the LIGA process (Section 29.3) fully exploits the benefits of this technique.

Electron-beam and Ion-beam Lithography. As in X-ray lithography, *electron-beam* (e-beam) and *ion-beam* lithography are superior to photolithography in terms of attainable resolutions. These two methods involve high current density in narrow electron or ion beams (known as *pencil sources*), which scan a pattern one pixel at a time onto a wafer. The masking is done by controlling the point-by-point transfer of the stored pattern, called *direct writing*, using software. These techniques have the advantages of accurate control of exposure over small areas of the wafer, large depth of focus, and low defect densities. Resolutions are limited to about 10 nm because of electron scatter, although 2-nm resolutions have been reported for some materials.

It should be noted that the scan time significantly increases as the resolution increases, because more highly focused beams are required. The main drawback of these two techniques is that electron and ion beams have to be maintained in a vacuum, thus significantly increasing equipment complexity. Moreover, the scan time for a wafer is much longer than that for other lithographic methods.

SCALPEL. In the SCALPEL (*sc*attering with *a*ngular *l*imitation *p*rojection *e*lectron-beam *l*ithography) process (Fig. 28.15), a mask is first produced from about a 0.1-μm-thick membrane of silicon nitride; it is then patterned with an approximately 50-nm-thick coating of tungsten. High-energy electrons pass through both the silicon nitride and the tungsten, but the tungsten scatters the electrons widely, whereas the silicon nitride results in very little scattering. An aperture blocks the scattered electrons, resulting in a high-quality image at the wafer.

The limitation to SCALPEL is the small-sized masks that are currently in use, although the process has high potential. Perhaps its most significant advantage is that energy does not need to be absorbed by the reticle; instead, it is blocked by the aperture, which is not as fragile or as expensive as the reticle.

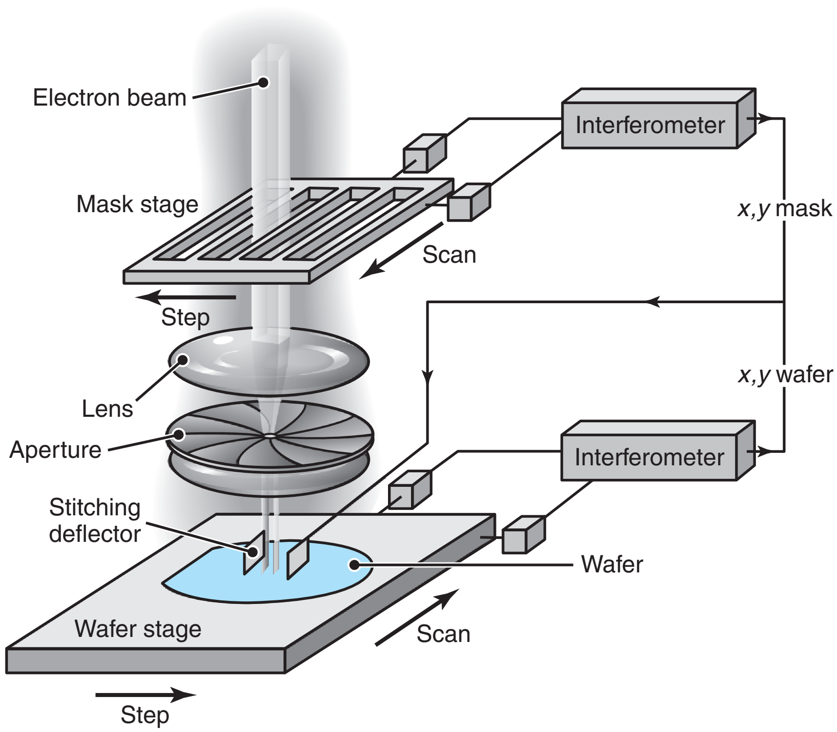


Figure 28.15: Schematic illustration of the SCALPEL process.

Example 28.1 Moore's Law

G. Moore, an inventor of the IC, observed in 1965 that the surface area of a single transistor is reduced by 50% every 12 months. In 1975, he revised this estimate to every two years; the resulting estimate is widely known as Moore's law, and it had been remarkably accurate for around three decades. Figure 28.16 shows the historical progression of feature size in *Dynamic Random Access Memory* (DRAM) bits, as well as projected future developments. Since about 2013, the rate of improvement has slowed down, which can be attributed to several factors. Among the more important ones are:

- To produce ever smaller features in a transistor requires that even more stringent manufacturing tolerances be achieved. For example, 180-nm line widths require ±14-nm dimensional tolerances, whereas 50-nm line widths require ±4 nm. Either requirement is especially problematic for the metal connection lines within the transistor.
- Smaller transistors can operate only if the dopant concentration is increased. Above a certain limit, however, the doping atoms cluster together, with the result that p-type and n-type silicon cannot be produced reliably at small length scales.
- The gate-switching energy of transistors has not been reduced at the same rate as their size; the result is increased power consumption in integrated circuits. This effect has a serious consequence, in that it is very difficult to dissipate the heat produced.
- At smaller length scales, microprocessors require lower voltages for proper operation. However, since the power consumption is still relatively high, very large currents are needed between the power-conversion devices and the central-processing units of modern microprocessors. These large currents result in resistive heating, compounding the heat-extraction problems.

Much research continues to be directed towards overcoming these limitations. Moore's law was intended as a prediction of the short-term future of the semiconductor industry, and was put forward at a time when photolithography was the only option. During the four decades since the law was first stated, researchers often have identified seemingly insurmountable problems which, in turn, have been overcome.

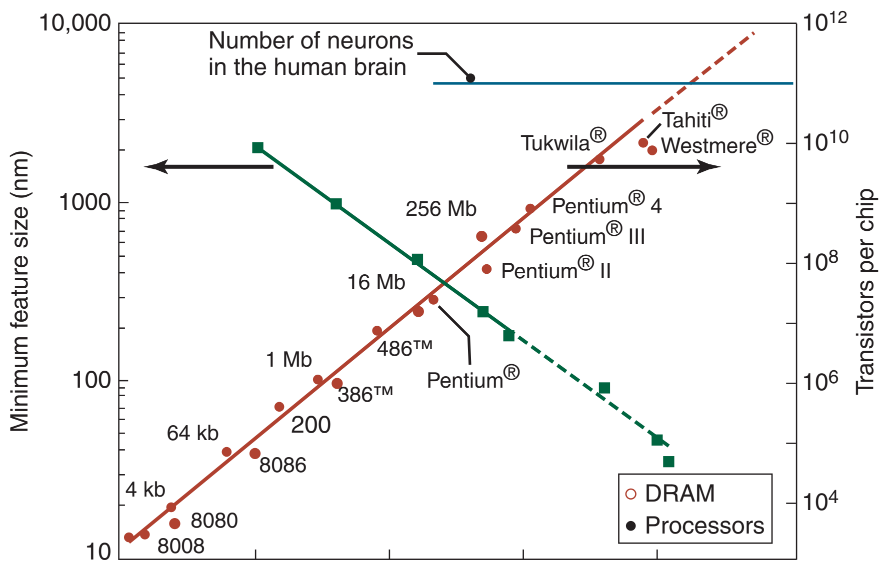


Figure 28.16: Illustration of Moore's law.

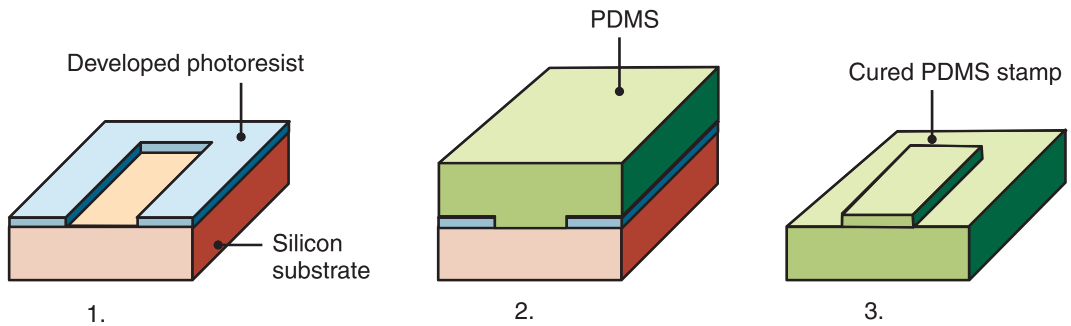


Figure 28.17: Production of a polydimethylsiloxane (PDMS) mold for soft lithography. 1. A developed photoresist is produced through standard lithography (Fig. 28.13). 2. A PDMS stamp is cast over the photoresist. 3. The PDMS stamp is peeled off the substrate to produce a stamp. The stamp shown has been rotated to emphasize the replication of surface features; the master pattern can be used several times. *Source:* After Y. Xia and G.M. Whitesides.

Soft Lithography. *Soft lithography* refers to several processes for pattern transfer, all of which require that a *master mold* be created by one of the standard lithography techniques described above. The master mold is then used to produce an elastomeric pattern, or stamp, as shown in Fig. 28.17. An elastomer that has commonly been used for the stamp is silicone rubber (polydimethylsiloxane, PDMS), because it is chemically inert, is not hygroscopic (does not swell due to humidity), and has good thermal stability, strength, durability, and surface properties.

Several PDMS stamps can be produced using the same pattern, and each stamp can be used several times. Some of the common soft lithography processes are:

1. **Microcontact printing** (μCP). In *microcontact printing*, the PDMS stamp is coated with an "ink" and then pressed against a surface. The peaks of the pattern are in contact with the opposing surface, and a thin layer of the ink is transferred, often only one molecule thick (called a *self-assembled monolayer* or *boundary film*; Section 33.6). This film can serve as a mask for selective wet etching, described below, or it can be used to impart a desired chemistry onto the surface.

2. **Microtransfer molding** (μTM). In this process, shown in Fig. 28.18a, the recesses in the PDMS mold are filled with a liquid polymer precursor, and then pressed against a surface. After the polymer has cured, the mold is peeled off, leaving behind a pattern suitable for further processing.

3. **Micromolding in capillaries.** Called MIMIC, in this technique (Fig. 28.18b), the PDMS stamp pattern consists of channels that use capillary action to wick a liquid into the stamp, either from the side of the stamp or from reservoirs within the stamp itself. The liquid can be a thermosetting polymer, a ceramic sol gel, or suspensions of solids within liquid solvents.

   Good pattern replication can be obtained, as long as the channel aspect ratio is moderate and the actual channel dimensions allow fluid flow. The dimensions required depend on the liquid used. The MIMIC process has been used to produce all-polymer field-effect transistors and diodes, and has various applications in sensors (Section 37.7).

28.8 Etching

Etching is the process by which entire films or particular sections of films are removed. One of the key criteria in this process is **selectivity**: the ability to etch one material without etching another. In silicon technology, an etching process must etch the silicon-dioxide layer effectively, with minimal removal of

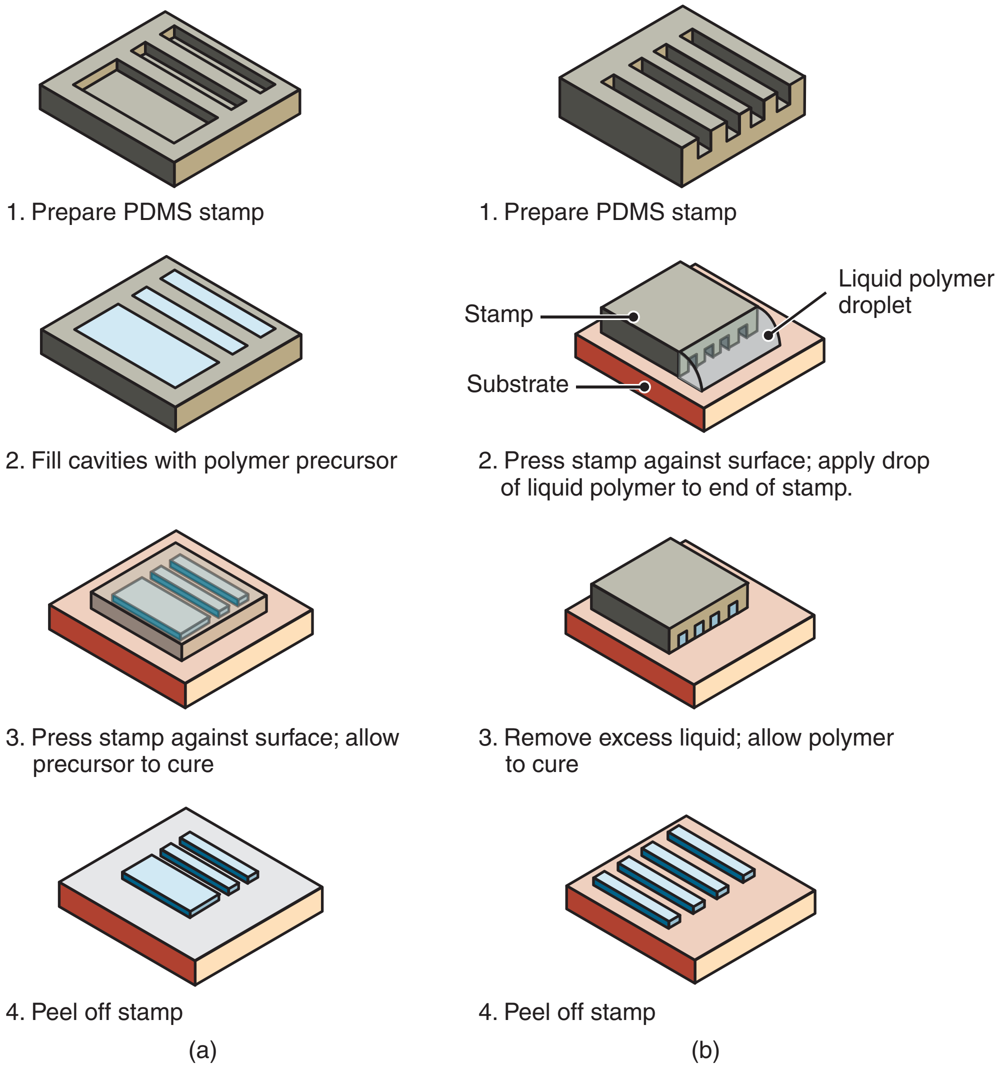


Figure 28.18: Soft lithography techniques. (a) Microtransfer molding (TM), and (b) micromolding in capillaries (MIMIC). *Source:* After Y. Xia and G.M. Whitesides.

either the underlying silicon or the resist material. In addition, polysilicon and metals must be etched into high-resolution lines with vertical wall profiles, and with minimal removal of either the underlying insulating film or the photoresist. Typical etch rates range from hundreds to several thousands of angstroms per minute, and selectivities can range from 1:1 to 100:1. A summary of etching processes and etchants is given in Tables 28.2 and 28.3.

28.8.1 Wet Etching

Wet etching involves immersing the wafers in a liquid solution, usually acidic. A primary feature of most wet-etching operations is that they are *isotropic*; that is, they etch in *all* directions of the workpiece at the same rate. Isotropy results in *undercuts* beneath the mask material (see, for example, Figs. 27.3b and 28.19a), and thus limits the resolution of geometric features in the substrate.

Table 28.2: General Characteristics of Silicon Etching Operations.

Wet etching	Temperature °C	Etch rate (μm/min)	{111} / {100} selectivity	Nitride etch rate (nm/min)	SiO_2 etch rate (nm/min)	p^{++} etch stop
Wet etching						
$HF:HNO_3:CH_3COOH$	25	1–20	—	Low	10–30	No
KOH	70–90	0.5–2	100:1	< 1	10	Yes
Ethylene-diamine pyrocatechol (EDP)	115	0.75	35:1	0.1	0.2	Yes
$N(CH_3)_4OH$ (TMAH)	90	0.5–1.5	50:1	< 0.1	< 0.1	Yes
Dry (plasma) etching						
SF_6	0–100	0.1–0.5	—	200	10	No
SF_6/C_4F_8 (DRIE)	20–80	1–3	—	200	10	No

Source: Adapted from N. Maluf, *An Introduction to Microelectromechanical Systems Engineering*, Artech House, 2000.

Effective etching requires the following conditions:

1. Etchant transport to the surface
2. A chemical reaction
3. Transport of reaction products away from the surface
4. Ability to stop the etching process rapidly, known as *etch stop*, in order to obtain superior pattern transfer, usually by using an underlying layer with high selectivity.

If the first or third condition listed above limits the speed of the process, agitation or stirring of the etchant can be employed to increase etching rates. If the second condition limits the speed of the process, the etching rate will strongly depend on temperature, etching material, and solution composition. Reliable etching thus requires both proper temperature control and repeatable stirring capability.

Isotropic Etchants. These etchants are widely used for

- Removing damaged surfaces
- Rounding sharply etched corners, to avoid stress concentrations

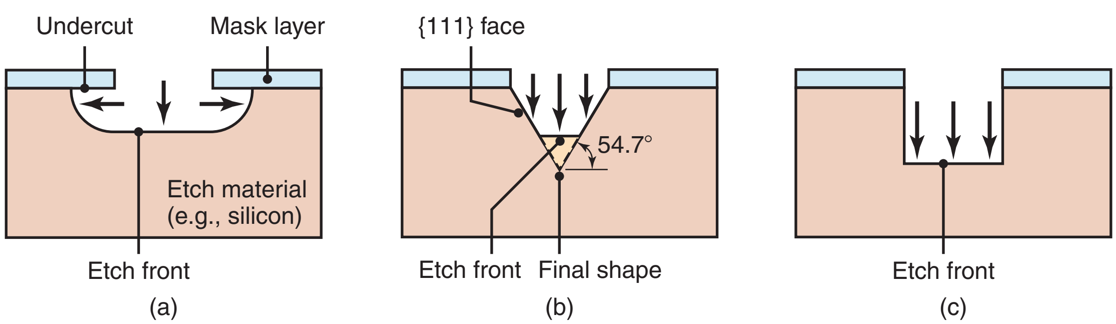


Figure 28.19: Etching directionality. (a) Isotropic etching: Etch proceeds vertically and horizontally at approximately the same rate, with significant mask undercut. (b) Orientation-dependent etching (ODE): Etch proceeds vertically, terminating on {111} crystal planes with little mask undercut. (c) Vertical etching: Etch proceeds vertically with little mask undercut. *Source:* After K.R. Williams, Agilent Laboratories.

Table 28.3: Comparison of Etch Rates for Selected Etchants and Target Materials.

Etchant	Target material	Etch rate (nm/min)[a] Polysilicon n^+	Polysilicon, undoped	Silicon dioxide	Silicon nitride	Phospho-silicate glass, annealed	Alum-inum	Titan-ium	Photoresist (OCG-820PR)
Wet etchants									
Concentrated HF (49%)	Silicon oxides	0	—	2300	14	3600	4.2	>1000	0
25:1 HF:H_2O	Silicon oxides	0	0	9.7	0.6	150	—	—	0
5:1 BHF[b]	Silicon oxides	9	2	100	0.9	440	140	>1000	0
Silicon etchant (126 HNO_3:60 H_2O:5 NH_4F)	Silicon	310	100	9	0.2	170	400	300	0
Aluminum etchant (16 H_3PO_4:1 HNO_3:1 HAc:2 H_2))	Aluminum	<1	<1	0	0	<1	660	0	0
Titanium etchant (20 H_2O:1 H_2O_2:1 HF)	Titanium	1.2	—	12	0.8	210	>10	880	0
Piranha (50 H_2SO_4:1 H_2O_2)	Metals and organics (cleaning off)	0	0	0	0	0	180	240	>10
Acetone (CH_3COOH)	Photoresist	0	0	0	0	0	0	0	>4000
Dry Etchants									
CF_4+CHF_3+He, 450 W	Silicon oxides	190	210	470	180	620	—	>1000	220
SF_6+He, 100 W	Silicon nitrides	73	67	31	82	61	—	>1000	69
SF_6, 12.5 W	Thin silicon nitrides	170	280	110	280	140	—	>1000	310
O_2, 400 W	Ashing photoresist	0	0	0	0	0	0	0	340

Notes:
[a] Results are for fresh solutions at room temperature, unless noted. Actual etch rates will vary with temperature and prior use of solution, area of exposure of film, other materials present, film impurities, and microstructure.
[b] Buffered hydrofluoric acid (33% NH4F and 8.3% HF by weight).
Source: After K. Williams and R. Muller.

- Reducing roughness developed after anisotropic etching
- Creating structures in single-crystal slices
- Evaluating defects.

Fabrication of microelectronic devices as well as microelectromechanical systems (Chapter 29) requires precise machining of structures, done through masking. With isotropic etchants, however, masking can be challenging, because the strong acids (a) etch aggressively, at a rate of up to 50 μm/min with an etchant of 66% HNO_3 and 34% HF, although etch rates of 0.1 to 1 μm/s are more typical, (b) produce rounded cavities, and (c) the etch rate is highly sensitive to agitation, thus lateral and vertical features are difficult to control. Because the size of the features in an integrated circuit determines its performance, there is a strong need to produce extremely small and well-defined structures. Such small features cannot be attained through isotropic etching, because of the poor definition resulting from undercutting of masks.

Anisotropic Etching. This situation occurs when etching is strongly dependent on compositional or structural variations in the material. There are two basic types of anisotropic etching: *orientation-dependent etching* (ODE) and *vertical etching*. Orientation-dependent etching commonly occurs in a single crystal, where etching takes place at different rates in different directions, as shown in Fig. 28.19b. Most vertical etching is done with dry plasmas, as described later.

Anisotropic etchants produce geometric shapes with walls that are defined by the crystallographic planes that resist the etchants. For example, Fig. 28.20 shows the vertical etch rate for silicon as a function of temperature. Note that the etching rate is more than one order of magnitude slower in the [111] crystal direction than in the other directions; therefore, well-defined walls can be obtained along the [111] crystal direction.

The **anisotropy ratio** for etching is defined by

$$\mathrm{AR} = \frac{E_1}{E_2} \tag{28.1}$$

where E is the etch rate in the crystallographic direction of interest, and the two subscripts refer to two directions for the materials of interest. The anisotropy ratio is unity for isotropic etchants, but can be as

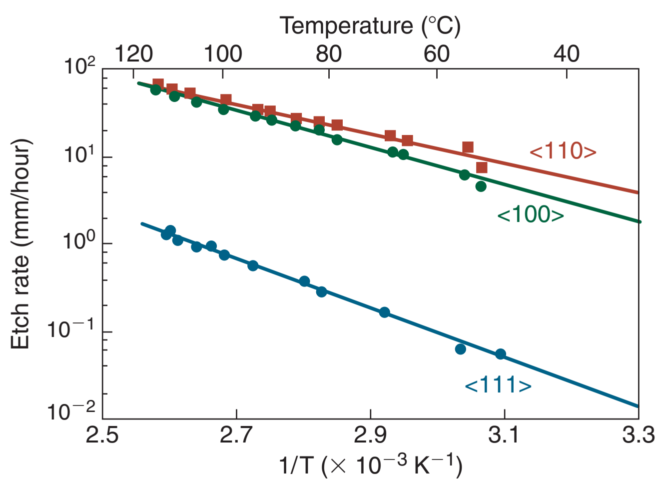


Figure 28.20: Etch rates of silicon in different crystallographic orientations, using ethylene-diamine/pyrocatechol-in-water as the solution. *Source:* After H. Seidel.

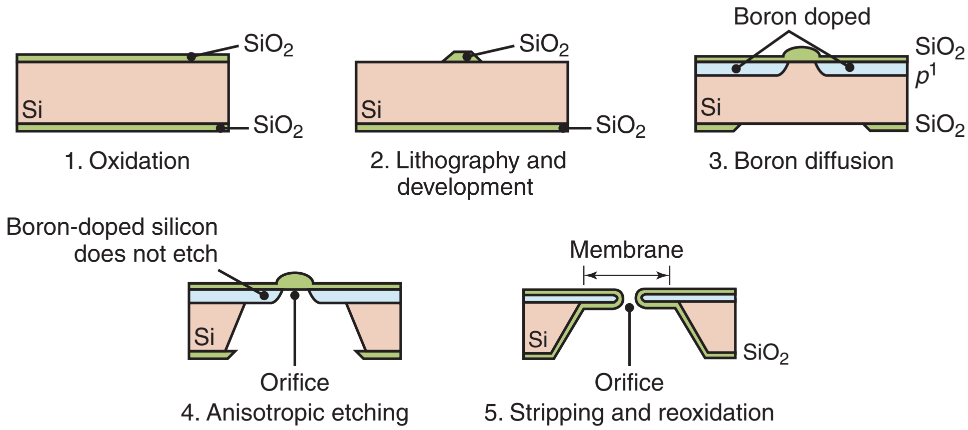


Figure 28.21: Application of a boron etch stop and back etching to form a membrane and orifice. *Source:* After I. Brodie and J.J. Murray.

high as 400/200/1 for (110)/(100)/(111) silicon. The $\langle 111 \rangle$ planes always etch at the slowest rate, but the $\langle 100 \rangle$ and $\langle 110 \rangle$ planes can be controlled through etchant chemistry.

Masking is also a concern in anisotropic etching, but for different reasons than those for isotropic etching. Anisotropic etching is slow, typically 3 μm/min, with anisotropic etching through a wafer taking as much as several hours. Silicon oxide may etch too rapidly to be used as a mask, hence a high-density silicon nitride mask may be needed.

Often, it is important to rapidly halt the etching process, especially when thin membranes are to be made or features with very precise thickness control are required. Conceptually, *rapid halting* can be accomplished by removing the wafer from the etching solution; however, etching depends to a great extent on the ability to circulate fresh etchants to the desired locations. Since the circulation varies across a wafer's surface, this strategy for halting the etching process would lead to large variations in the depth etched.

The most common approach to obtain uniform feature sizes across a wafer is to use a *boron etch stop* (Fig. 28.21), whereby a boron layer is diffused or implanted into the silicon. Examples of common etch stops include the placement of a boron-doped layer beneath silicon or the placement of silicon dioxide (SiO_2) beneath silicon nitride (Si_3N_4). Because anisotropic etchants do not attack boron-doped silicon as aggressively as they do undoped silicon, surface features or membranes can be created by **back etching**.

Several etchant formulations have been developed, including hydrofluoric acid, phosphoric acid, mixtures of nitric acid and hydrofluoric acid, potassium hydrochloride, and mixtures of phosphoric acid, nitric acid, acetic acid, and water. Wafer *cleaning* is done with a solution consisting of sulfuric acid and peroxide, called *Piranha solution*, a trade name. Photoresist can be removed with these solutions, although acetone is commonly used for this purpose.

28.8.2 Dry Etching

Integrated circuits are now etched exclusively by *dry etching*, which involves the use of chemical reactants in a low-pressure system. In contrast to wet etching, dry etching can have a high degree of directionality, resulting in highly anisotropic etching profiles (Fig. 28.19c). Also, dry etching requires only small amounts of the reactant gases, whereas the solutions used in wet etching have to be refreshed periodically.

Dry etching usually involves a plasma or discharge in areas of high electric and magnetic fields; any gases that are present are dissociated to form ions, photons, electrons, or highly reactive molecules. Table 28.2 lists some of the more common dry etchants, their target materials, and typical etch rates. There are several specialized dry-etching techniques.

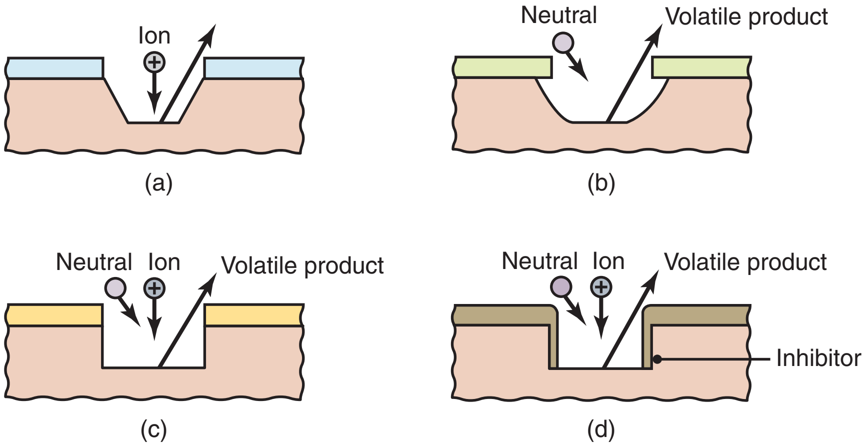


Figure 28.22: Machining profiles associated with different dry-etching techniques: (a) sputtering; (b) chemical; (c) ion-enhanced energetic; and (d) ion-enhanced inhibitor. *Source:* After M. Madou.

Sputter Etching. This process removes material by bombarding the surface with noble gas ions, usually Ar^+. The gas is ionized in the presence of a cathode and an anode (Fig. 28.22). If a silicon wafer is the target, the momentum transfer associated with the bombardment of atoms causes bond breakage and material to be ejected or sputtered. If the silicon chip is the substrate, then the material in the target is deposited onto the silicon, after it has been sputtered by the ionized gas.

The major concerns in sputter etching are:

- The ejected material can be redeposited onto the target, especially with large aspect ratios.
- Sputtering can cause damage or excessive erosion of the material.
- Sputter etching is not material selective, and because most materials sputter at about the same rate, masking is difficult.
- The sputter etching process is slow, with etch rates limited to tens of nm/min.
- The photoresist is difficult to remove.

Reactive Plasma Etching. Also called **dry chemical etching**, this process involves chlorine or fluorine ions (generated by RF excitation), and other molecular species, that diffuse into and chemically react with the substrate. As a result, a volatile compound is formed, which is then removed by a vacuum system. The mechanism of reactive plasma etching is illustrated in Fig. 28.23:

1. A reactive species, such as CF4, is first produced, and it dissociates upon impact with energetic electrons, to produce fluorine atoms.
2. The reactive species then bombards and diffuses into the surface.
3. The reactive species chemically reacts to form a volatile compound.

4. The reactant desorbs from the surface.

5. It then diffuses into the bulk gas, where it is removed by a vacuum system.

Some reactants polymerize on the surface, thus requiring additional removal of material, either with oxygen in the plasma reactor or by an external *ashing* operation (see *dry stripping* in Section 28.7). The electrical charge of the reactive species is not high enough to cause damage through impact on the surface, hence no sputtering occurs. Thus, the etching is isotropic and undercutting of the mask takes place (Fig. 28.19a).

Physical–Chemical Etching. Processes such as *reactive ion-beam etching* (RIBE) and *chemically assisted ion-beam etching* (CAIBE) combine the advantages of physical and chemical etching. These processes use a chemically reactive species to drive material removal, but are assisted physically by the impact of ions onto the surface. In RIBE, also known as *deep reactive-ion etching* (DRIE), vertical trenches hundreds of nanometers deep can be produced by periodically interrupting the etching process and depositing a polymer layer (Fig. 28.23d).

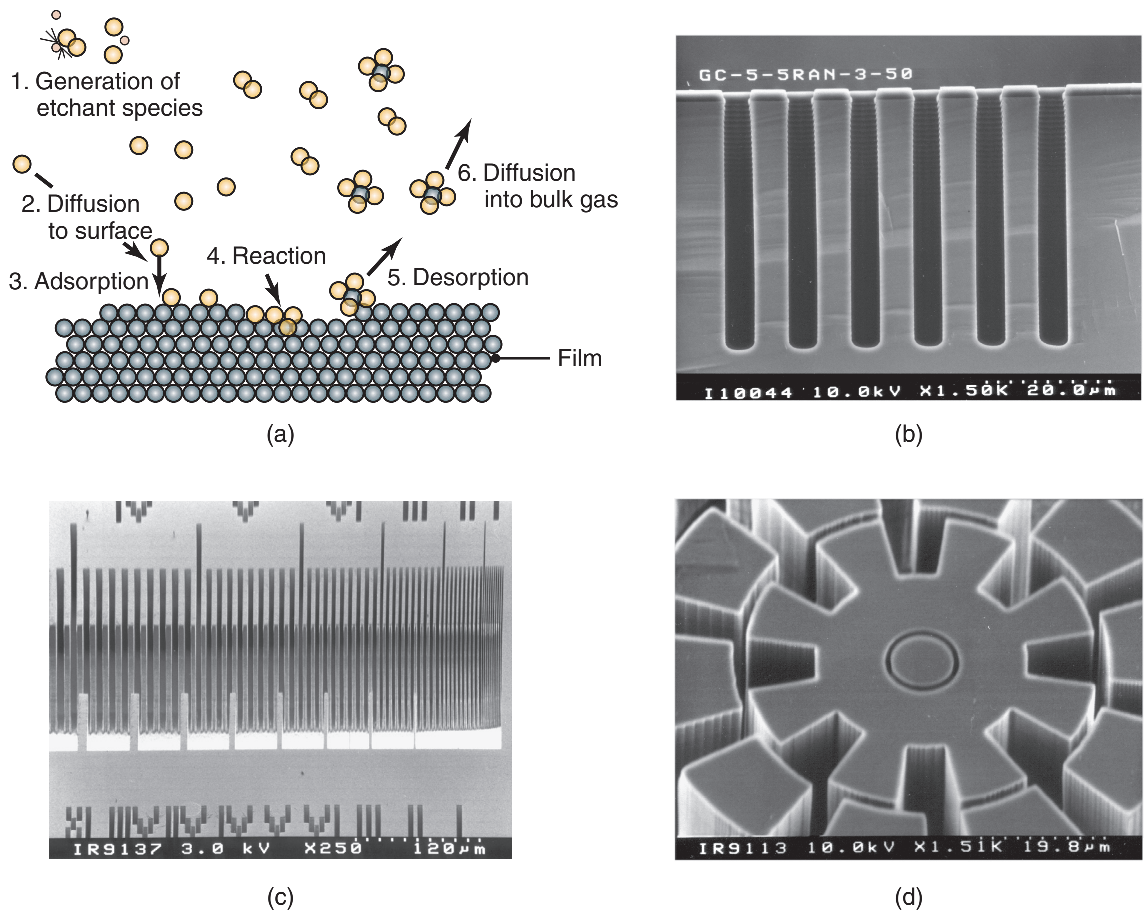


Figure 28.23: (a) Schematic illustration of reactive plasma etching. (b) Example of a deep reactive-ion etched trench; note the periodic undercuts, or scallops. (c) Near-vertical sidewalls produced through deep reactive-ion etching (DRIE), an anisotropic-etching process. (d) An example of cryogenic dry etching, showing a 145-μm deep structure etched into silicon with the use of a 2.0-μm-thick oxide masking layer. The substrate temperature was -140°C during etching. *Source:* (a) After M. Madou. (b) through (d). After R. Kassing and I.W. Rangelow, University of Kassel, Germany.

In CAIBE, ion bombardment can assist dry chemical etching by

- making the surface more reactive;
- clearing the surface of reaction products and allowing the chemically reactive species access to the cleared areas; and
- providing the energy to drive surface chemical reactions; however, the neutral species do most of the etching.

Physical–chemical etching is extremely useful because the ion bombardment is directional, so that etching is anisotropic. Also, the ion-bombardment energy is low and does not contribute much to mask removal, thus allowing the generation of near-vertical walls with very large aspect ratios. Since the ion bombardment does not remove material directly, masks can be used.

Cryogenic Dry Etching. This method is used to obtain very deep features with vertical walls. The workpiece is first lowered to cryogenic temperatures, then the CAIBE process takes place. The very low temperatures involved ensure that the energy available is not sufficient for a surface chemical reaction to take place, unless ion bombardment is normal to the surface. Oblique impacts, such as those occurring on sidewalls in deep crevices, cannot drive the chemical reactions.

Because dry etching is not selective, etch stops cannot be applied directly and dry-etching reactions must be terminated when the target film is removed. This can be done by measuring the wavelength of light being emitted during a reaction; when the target film is removed, the wavelength of light emitted will change and can be detected with proper sensors.

Example 28.2 Comparison of Wet and Dry Etching

Consider the case where a $\langle 100 \rangle$ wafer (see Fig. 28.5) has an oxide mask placed on it, in order to produce square or rectangular holes. The sides of the square are oriented precisely within the $\langle 100 \rangle$ direction of the wafer surface, as shown in Fig. 28.24.

Isotropic etching results in the cavity shown in Fig. 28.24a, and since etching occurs at constant rates in all directions, a rounded cavity that undercuts the mask is produced. An orientation-dependent etchant produces the cavity shown in Fig. 28.24b. Because etching is much faster in the $\langle 100 \rangle$ and $\langle 110 \rangle$ directions than in the $\langle 111 \rangle$ direction, sidewalls defined by the $\langle 111 \rangle$ plane are generated. For silicon, these sidewalls are at an angle of 54.74° to the surface.

The effect of a larger mask or shorter etch time is shown in Fig. 28.24c. The resultant pit is defined by $\langle 111 \rangle$ sidewalls and by a bottom in the $\langle 100 \rangle$ direction parallel to the surface. A rectangular mask and the resulting pit are shown in Fig. 28.24d. Deep reactive-ion etching is depicted in Fig. 28.24e. Note that a polymer layer is deposited periodically onto the hole sidewalls, to allow for deep pockets, but scalloping (greatly exaggerated in the figure) is unavoidable. A hole resulting from CAIBE is shown in Fig. 28.24f.

28.9 Diffusion and Ion Implantation

Recall that the operation of microelectronic devices depends on regions that have different doping types and concentrations. The electrical characteristics of these regions are altered through the introduction of dopants into the substrate, by *diffusion* and *ion-implantation processes*. This step in the fabrication sequence is repeated several times, since many different regions of microelectronic devices must be defined.

In the diffusion process, the movement of atoms is a result of thermal excitation. Dopants can be introduced to the substrate surface in the form of a deposited film or the substrate can be placed in a vapor, containing the dopant source. The process takes place at temperatures usually 800° to 1200°C. Dopant movement within the substrate is strictly a function of temperature, time, and the diffusion coefficient (or *diffusivity*) of the dopant species, as well as the type and quality of the substrate material.

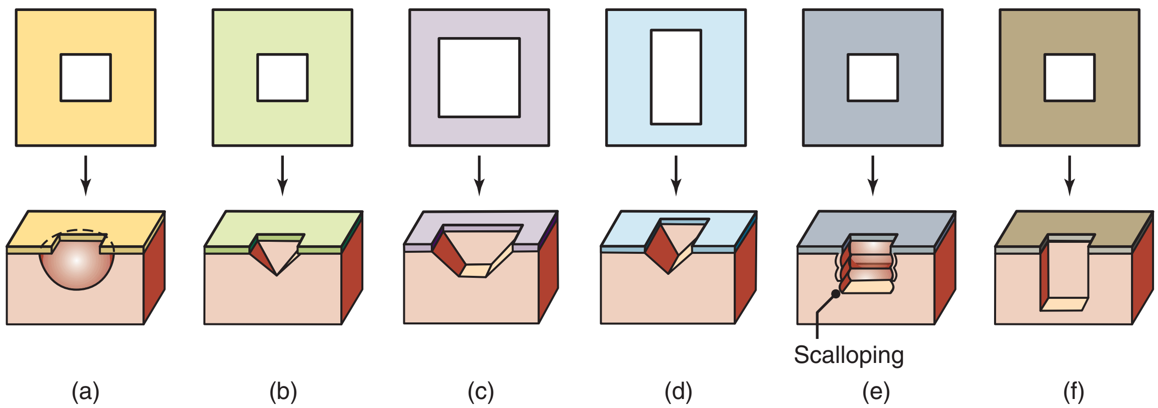


Figure 28.24: Various holes generated from a square mask in (a) isotropic (wet) etching; (b) orientation-dependent etching (ODE); (c) ODE with a larger hole; (d) ODE of a rectangular hole; (e) deep reactive-ion etching; and (f) vertical etching. *Source:* After M. Madou.

Because of the nature of diffusion, the dopant concentration is very high at the substrate surface, and drops off sharply away from the surface. For a more uniform concentration within the substrate, the wafer is further heated to drive in the dopants, in a process called **drive-in diffusion**. Diffusion, whether desired or not, is highly isotropic and always occurs at high temperatures, a phenomenon that has to be taken into account during subsequent processing steps.

Ion implantation is a much more extensive process and requires specialized equipment (Fig. 28.25; see also Section 34.7). Implantation is accomplished by accelerating the ions through a high-voltage field of as much as 1 million electron volts, and then by choosing the desired dopant by means of a mass separator. In a manner similar to that of cathode-ray tubes, the beam is swept across the wafer by sets of deflection plates, thus ensuring uniformity of coverage of the substrate. The whole implantation operation must be performed in a vacuum.

The high-velocity impact of ions on the silicon surface damages the lattice structure, resulting in lower electron mobilities. Although this condition is undesirable, the damage can be repaired by an *annealing* step, which involves heating the substrate to relatively low temperatures, usually 400° to 800°C, for a period of 15 to 30 min. Annealing provides the energy that the silicon lattice needs to rearrange and mend itself. Another important function of annealing is driving in the implanted dopants; implantation alone imbeds the dopants less than half a micron below the silicon surface. Annealing enables the dopants to diffuse to a more desirable depth of a few microns.

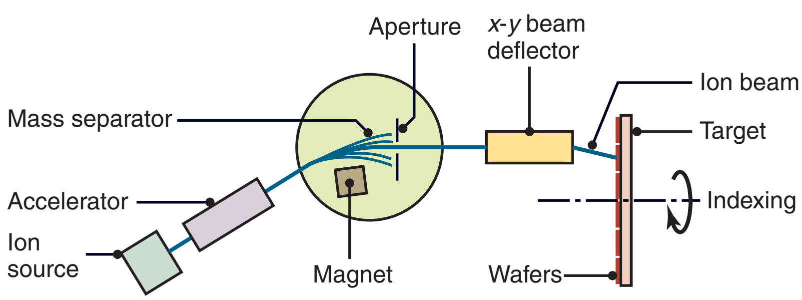


Figure 28.25: Schematic illustration of an apparatus for ion implantation.

Example 28.3 Processing of a *p*-type Region in *n*-type Silicon

Given: It is desired to create a p-type region within a sample of n-type silicon.

Find: Draw cross sections of the sample at each processing step in order to accomplish this task.

Solution: Refer to Fig. 28.26. This simple device, known as a *pn-junction diode*, is the foundation for most semiconductor devices.

28.10 Metallization and Testing

The preceding sections focused on device fabrication. Generating a complete and functional integrated circuit requires that these devices be *interconnected*, which must take place on a number of levels (Fig. 28.27). **Interconnections** are made using metals that exhibit low electrical resistance and good adhesion to dielectric insulator surfaces. Aluminum, copper, and aluminum–copper alloys remain the most commonly used materials for this purpose in VLSI technology today.

Because device dimensions continue to shrink, **electromigration** has become more of a concern with aluminum interconnects, a process by which aluminum atoms are moved physically by the impact of drifting electrons under high currents. In extreme cases, electromigration can lead to severed or shorted metal lines. Solutions to this problem include (a) addition of sandwiched metal layers, such as tungsten and titanium, and (b) using pure copper, which displays lower resistivity and has significantly better electromigration performance than aluminum.

Metals are deposited by standard deposition techniques, called **metallization**. Modern ICs typically have 1 to 10 layers of metallization, each layer of metal being insulated by a dielectric. Interconnection patterns are generated through lithographic and etching processes.

Planarization, producing a planar surface of interlayer dielectrics, is critical to the reduction of metal shorts and the line width variation of the interconnect. A common method to achieve a planar surface is a uniform oxide-etch process, that smoothens out the peaks and valleys of the dielectric layer. Planarizing high-density interconnects has now become the process of **chemical–mechanical polishing** (CMP) (Section 26.7). This process entails physically polishing the wafer surface, in a manner similar to that by which a disc or belt sander flattens the ridges in a piece of wood. A typical CMP process combines an abrasive medium with a polishing compound or slurry; it can polish a wafer to within 0.03 μm of being perfectly flat, with an R_q roughness (Section 33.3) on the order of 0.1 nm for a new silicon wafer.

Layers of metal are connected together vertically by **vias**, and access to the devices on the substrate is achieved through **contacts** (Fig. 28.28). As devices continue to become smaller and faster, the size and speed of some chips have become limited by the metallization process itself. Wafer processing is completed upon application of a *passivation layer*, usually silicon nitride (Si_3N_4). The silicon nitride acts as a barrier to sodium ions and also provides excellent scratch resistance.

The next step is to test each of the individual circuits on the wafer (Fig. 28.29). Each chip, also known as a **die**, is tested by a computer-controlled probe platform, containing needlelike probes that access the bonding pads on the die. The probes are of two forms:

1. **Test patterns or structures.** The probe measures test structures, often outside of the active dice, placed in the scribe line (the empty space between dice). These probes consist of transistors and interconnect structures that measure various processing parameters, such as resistivity, contact resistance, electromigration performance, and leakage currents.

2. **Direct probe.** This approach involves 100% testing on the bond pads of each die.

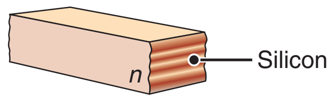
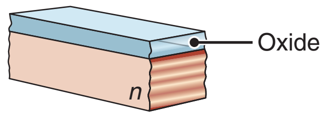
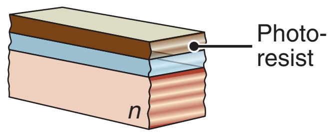
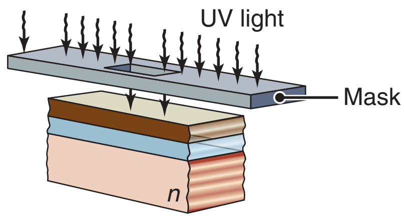
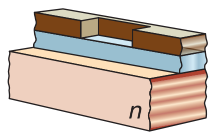
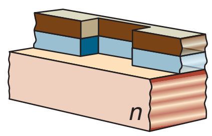
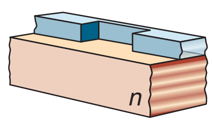
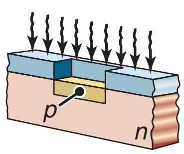
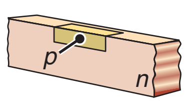
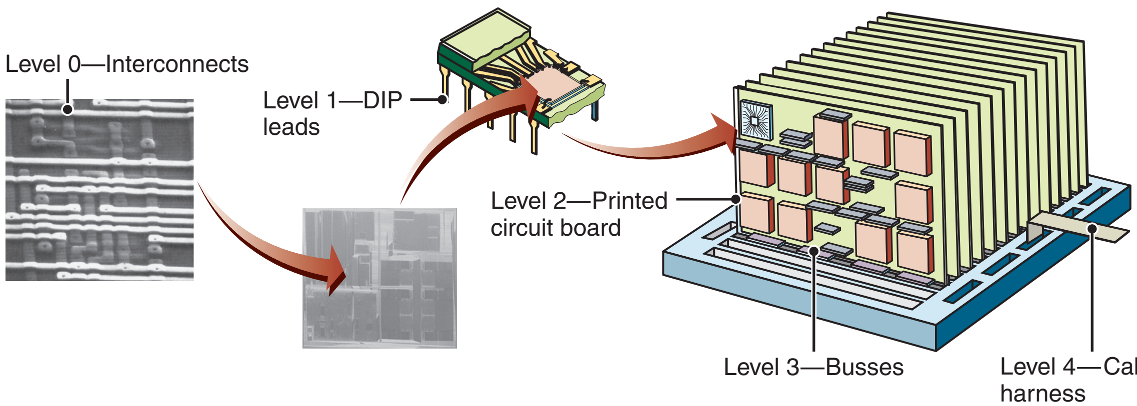
Processing step	Cross section	Description
1.		Sample of *n*-type silicon
2.		Grow silicon dioxide by oxidation
3.		Apply photoresist
4.		Expose photoresist, using appropriate lithographic mask
5.		Develop photoresist
6.		Etch silicon dioxide
7.		Remove photoresist
8.		Implant boron
9.		Remove silicon dioxide

Figure 28.26: Fabrication sequence for a pn-junction diode.

The platform scans across the wafer, and uses computer-generated timing waveforms to test whether each circuit is functioning properly. If a chip is defective, it is marked with a drop of ink. Up to one-third of the cost of a microelectronic circuit can be incurred during this testing period.

After the wafer-level testing is completed, back grinding may be done to remove a large amount of the original substrate. The final die thickness depends on the packaging requirement, and anywhere from 25% to 75% of the wafer thickness may be removed. After back grinding, each die is separated from the wafer.



Level	Element example	Interconnection method
Level 0	Transistor within an IC	IC metallization
Level 1	ICs, other discrete components	Package leads or module interconnections
Level 2	IC packages	Printed circuit board
Level 3	Printed circuit boards	Connectors (busses)
Level 4	Chassis or box	Connectors/cable harnesses
Level 5	System, e.g., computer	

Figure 28.27: Connections between elements in the hierarchy for integrated circuits.

Diamond sawing is a commonly used separation technique, and results in very straight edges with minimal chipping and cracking damage. The chips are then sorted, the functional dice are sent on for packaging, and the inked dice are discarded.

28.11 Wire Bonding and Packaging

The working dice must be attached to a more rugged foundation to ensure reliability. One simple method is to fasten a die to its packaging material with epoxy cement; another method makes use of a eutectic bond, made by heating metal-alloy systems (Section 4.3). A widely used mixture is 96.4% Au and 3.6% Si, which has a eutectic point at 370°C.

Once the chip has been attached to its substrate, it must be connected electrically to the package leads. This is accomplished by *wire bonding* very thin (25 μm diameter) gold wires from the package leads to bonding pads, located around the perimeter or down the center of the die (Fig. 28.30). The bonding pads on the die are typically drawn at 50 to 100 μm per side; the bond wires are attached by means of *thermocompression, ultrasonic,* or *thermosonic* techniques (Fig. 28.31).

The connected circuit is now ready for final packaging. The *packaging* process largely determines the overall cost of each completed IC since the circuits are mass produced on the wafer, but they are then packaged individually. Packages are available in a wide variety of styles; the appropriate one must reflect the operating requirements. Consideration of a circuit's package includes the chip size, number of external

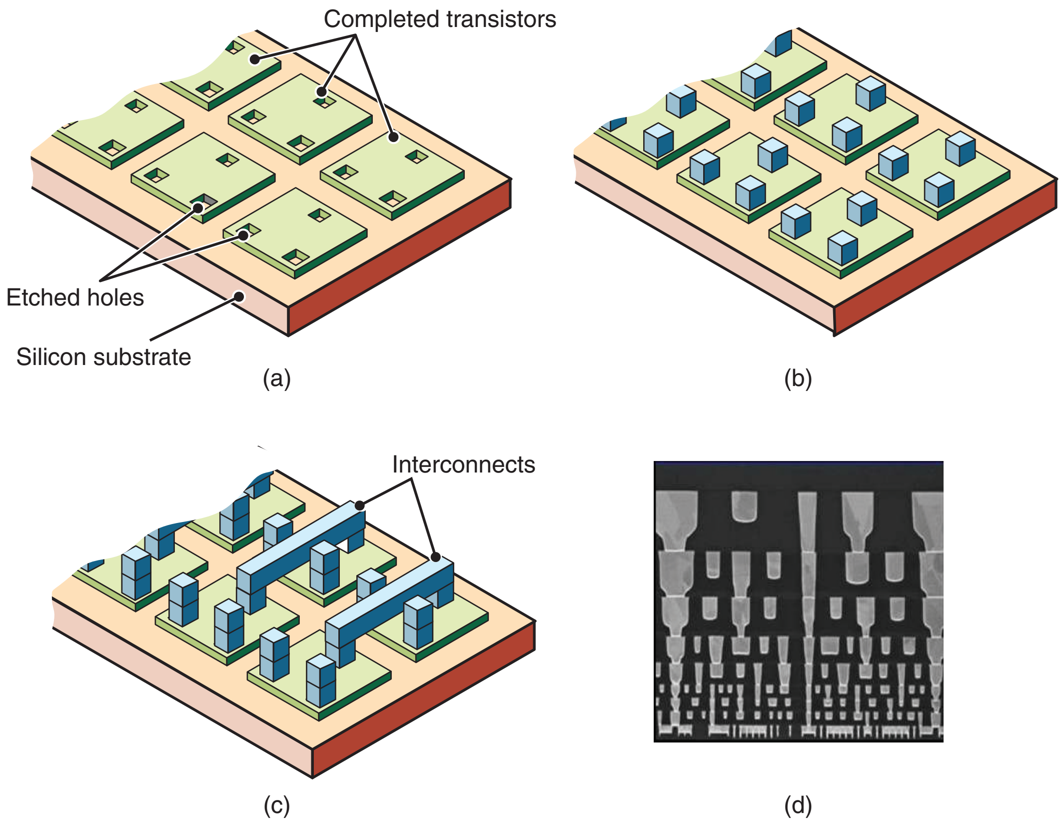


Figure 28.28: Production of metal interconnects. (a) Wafer with fabricated transistors (see Fig. 28.26). (b) First metallization layer, produced by electroplating copper. The dielectric has been removed for clarity; note that the copper extends into the holes in the transistors to produce an electrical connection. (c) After the second layer, showing an interconnect between transistors. Over 50 layers can be required to fully interconnect an IC, although most applications require less than 13. (d) Cross section of a nine-layer interconnect structure; each layer is between 0.25 and 1 μm thick. *Source:* (d) Courtesy of Intel Corporation.

leads, operating environment, heat dissipation, and power requirements. For example, ICs used for military and industrial applications require packages of particularly high strength, toughness, and resistance to temperature.

Packages are produced from polymers, metals, or ceramics. Metal containers are made from alloys such as Kovar (an iron-cobalt-nickel alloy with a low coefficient of thermal expansion; Section 3.6), which provide a hermetic seal and good thermal conductivity, but they are limited in the number of leads that can be used. Ceramic packages usually are produced from aluminum oxide (Al_2O_3); they are hermetic, have good thermal conductivity, but have higher lead counts than metal packages and also are more expensive. Plastic packages are inexpensive and have high lead counts, but they cannot withstand high temperatures and are not hermetic.

An older style of packaging is the **dual-in-line package** (DIP), shown in Fig. 28.32a. Characterized by low cost and ease of handling, DIP packages are made of thermoplastics, epoxies, or ceramics, and can have from 8 to 500 external leads. Ceramic packages are designed for use over a broader temperature range and in high-reliability and military applications; however, they cost considerably more than plastic packages.

Figure 28.32b shows a flat ceramic package in which the package and all of the leads are in the same plane. This package style does not offer the ease of handling or the modular design of the DIP package. Consequently, it usually is affixed permanently to a multiple-level circuit board, in which the low profile of the flat package is necessary.

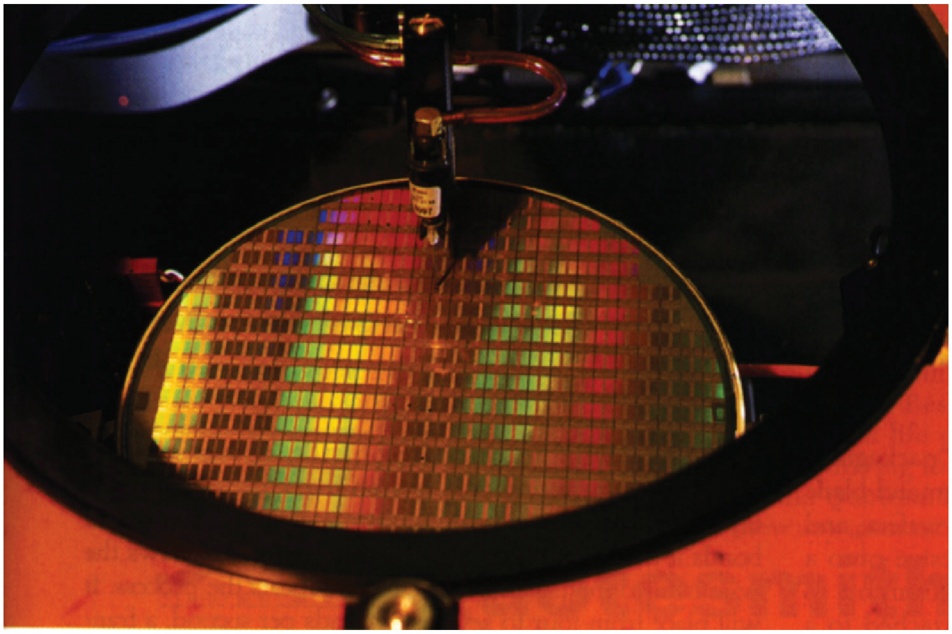


Figure 28.29: A probe (top center) checking for defects in a wafer; an ink mark is placed on each defective die. *Source:* Courtesy of Intel Corp.

Surface-mount packages have become common for today's integrated circuits. Some examples are shown in Fig. 28.32c; note that the main difference among them is in the shape of the connectors. The DIP connection to the surface board is *via prongs*, which are inserted into corresponding holes, whereas a surface mount is soldered onto a specially fabricated pad or *land*, a raised solder platform for interconnections among components in a printed circuit board. Package size and layouts are selected from standard patterns, and usually require adhesive bonding of the package to the board, followed by **wave soldering** of the connections (Section 32.3.3).

Faster and more versatile chips require increasingly tightly spaced connections. **Pin-grid arrays** (PGAs) use tightly packed pins that connect onto printed circuit boards by way of through holes. PGAs and other in-line and surface-mount packages are, however, extremely susceptible to plastic deformation of the wires

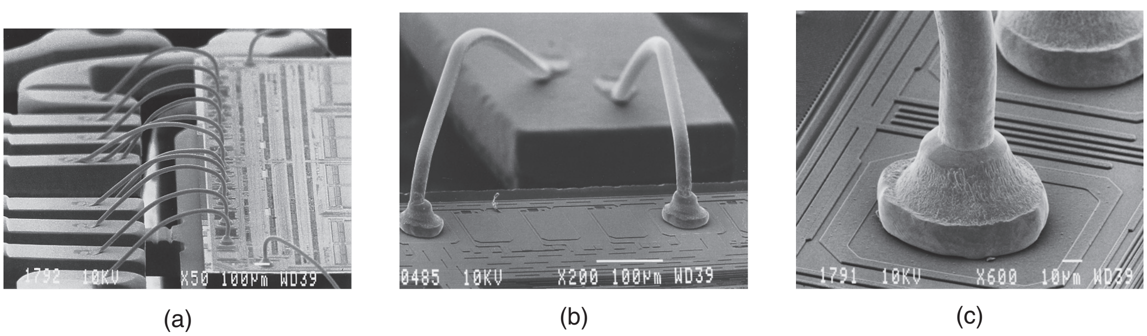


Figure 28.30: (a) SEM photograph of wire bonds connecting package leads (left-hand side) to die bonding pads. (b) and (c) Detailed views of (a). *Source:* Courtesy of Micron Technology, Inc.

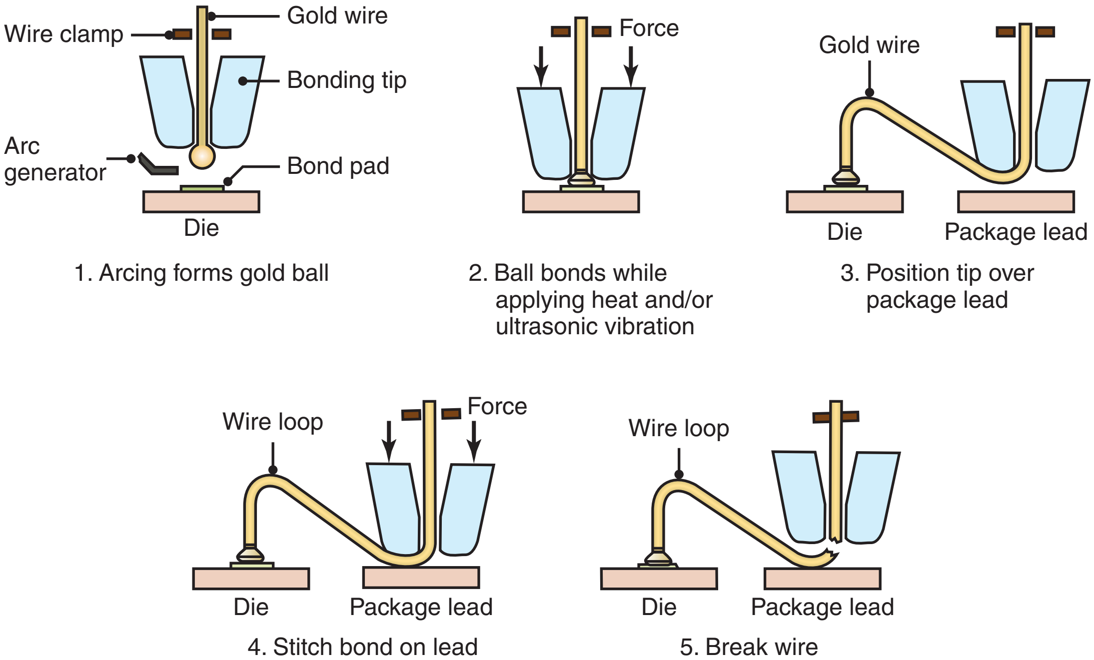


Figure 28.31: Schematic illustration of thermosonic welding of gold wires from package leads to bonding pads.

and legs, especially with small-diameter, closely spaced wires. One way of achieving tight packing of connections and avoiding the difficulties of slender connections is through **ball-grid arrays** (BGAs), shown in Fig. 28.32d. This type of array has a solder-plated coating on a number of closely spaced metal balls on the underside of the package. The spacing between the balls can be as small as 50 μm, but more commonly it is standardized as 1.0 mm, 1.27 mm, or 1.5 mm.

Although BGAs can be designed with over 1000 connections, this is extremely rare and usually 200 to 300 connections are sufficient for demanding applications. By using the **reflow soldering** technique (Section 32.3.3), the solder serves to center the BGA by surface tension, thus resulting in well-defined electrical connections for each ball. After the chip has been sealed in the package, it undergoes a final testing. Because one of the main purposes of packaging is isolation from the environment, testing at this stage usually involves such factors as heat, humidity, mechanical shock, corrosion, and vibration. Destructive tests also are performed to determine the effectiveness of sealing.

Chip on Board. *Chip on board* (COB) designs refer to the direct placement of chips onto an adhesive layer on a circuit board. Electrical connections are then made by wire bonding the chips directly to the pads. After wire bonding, final encapsulation with an epoxy is necessary, not only to attach the IC package more securely to the printed circuit board but also to transfer heat evenly during operation.

Flip-chip on Board. The Flip-chip on board (FOB) technology (Fig. 28.33) involves the direct placement of a chip with solder bumps onto an array of pads on the circuit board. The main advantage to flip chips and ball-grid array packages is that the space around the package, normally reserved for bond pads, is saved. Thus, a higher level of miniaturization can be achieved.

System in Package. A trend that allows for more compact devices involves incorporating more than one integrated circuit into a package. Figure 28.34 illustrates the major categories of *silicon in package* (SiP) designs.

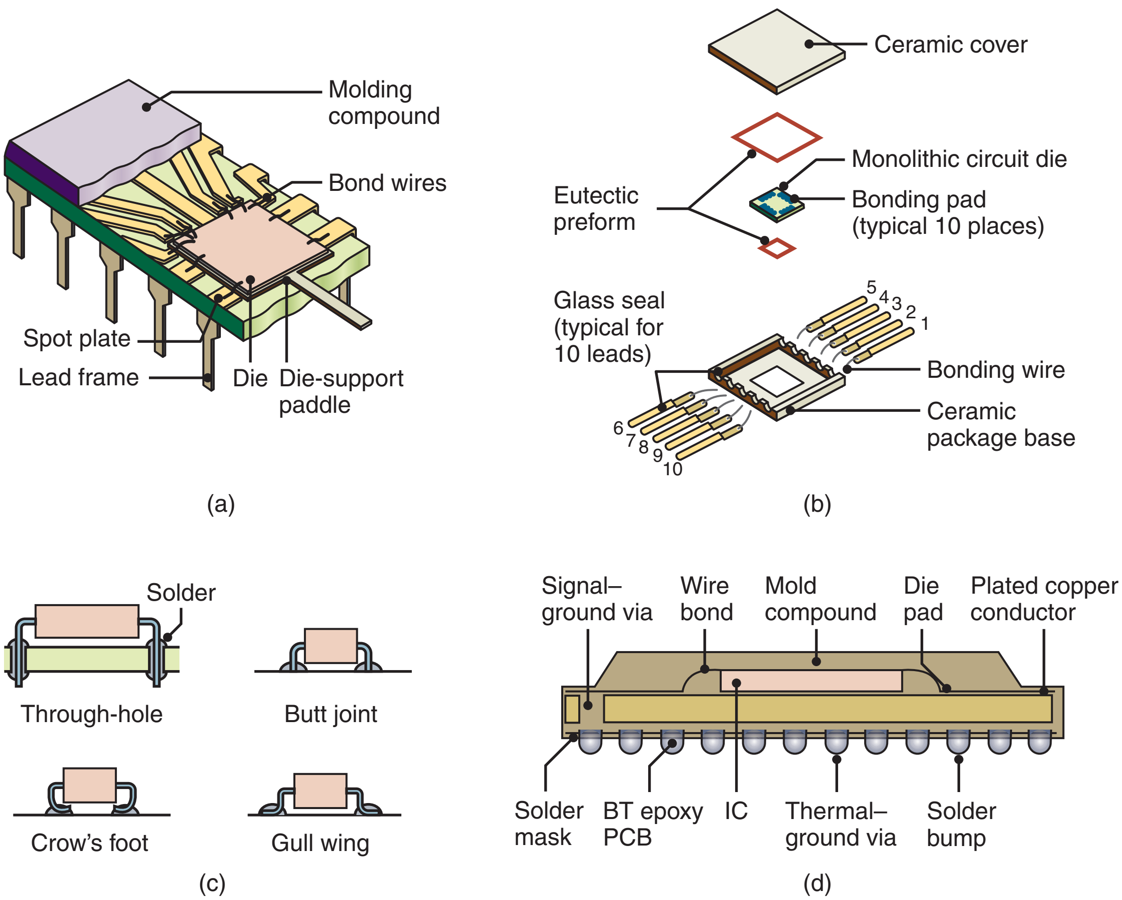


Figure 28.32: Schematic illustrations of various IC packages: (a) dual-in-line package (DIP); (b) flat ceramic package; (c) common surface-mount configurations; (d) ball-grid arrays.

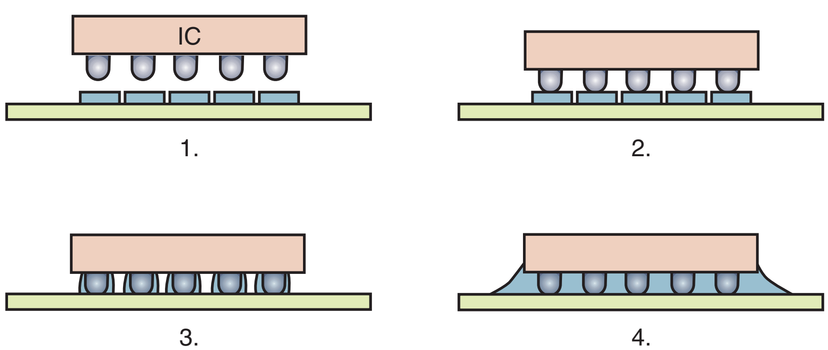


Figure 28.33: Illustration of flip-chip technology. Flip-chip package with 1. solder-plated metal balls and pads on the printed circuit board; 2. flux application and placement; 3. reflow soldering; 4. encapsulation.

Although these packages can be integrated horizontally, vertical integration through stacked or embedded structures (Fig. 28.34b and c) has the advantage of achieving performance increases over conventional packages.

These benefits have been described as “more than Moore” (see Example 28.1), although SiPs also have other advantages, such as (a) they present reduced size and less noise, (b) cross talk between chips can be better isolated, and (c) individual chips can be upgraded more easily. On the other hand, these packages are more complex, require higher power density and associated heat extraction, and are more expensive than conventional packages.

SiP packages can, however, be made very simple by incorporating more than one chip inside a single package (Fig. 28.34a). To preserve area on a circuit board, chips and/or flip chips can be stacked and bonded to a circuit board, to produce three-dimensional integrated circuits (Fig. 28.34b). Here, an interposing layer, commonly an adhesive, separates the chips and electrically isolates adjacent layers.

An alternative is to employ a so-called *interposerless* structure, using **through-silicon vias** (TSVs) instead of wire bonding, to provide electrical connections to all layers. TSVs are sometimes considered a packaging feature; however, it has been noted that this is perhaps a case of 3D integration of a wafer, as shown in Fig. 28.34c.

28.12 Yield and Reliability

Yield is defined as the ratio of functional chips to the total number of chips produced. The overall yield of the total IC manufacturing process is the product of the wafer yield, the bonding yield, the packaging yield, and the test yield. This quantity can range from only a few percent for new processes to more than 90% for mature manufacturing lines. Most yield loss occurs during wafer processing, due to its more complex nature. Wafers are commonly separated into regions of good and bad chips. Failures at this stage can arise from point defects (such as oxide pinholes), film contamination, metal particles, and area defects (such as uneven film deposition or nonuniformity of an etch process).

A major concern about completed ICs is their **reliability** and **failure rate**. Since no device has an infinite lifetime, statistical methods are used to characterize the expected lifetimes and failure rates of microelectronic devices. The unit of failure rate is the FIT, defined as *one failure per 1 billion device-hours*. However, complete systems may have millions of devices, hence the overall failure rate in entire systems is correspondingly higher.

Equally important in failure analysis is determining the *failure mechanism*, the actual process that causes the device to fail. Common failures due to processing involve:

- Diffusion regions: nonuniform current flow and junction breakdown
- Oxide layers: dielectric breakdown and accumulation of surface charge
- Lithography: uneven definition of features and mask misalignment
- Metal layers: poor contact and electromigration resulting from high current densities; shorts or filaments due to poor planarization
- Other failures: originating in improper chip mounting, poorly formed wire bonds, or loss of the package’s hermetic seal.

Because device lifetimes are on the order of years, it is impractical to study device failures under normal operating conditions. One method of studying failures efficiently is **accelerated life testing**; it involves accelerating the conditions whose effects cause device breakdown. Cyclic variations in temperature, humidity, voltage, and current are used to stress the components. Chip mounting and packaging are strained by cyclical temperature variations. The statistical data taken from these tests are then used to predict device-failure modes and device life under normal operating conditions.

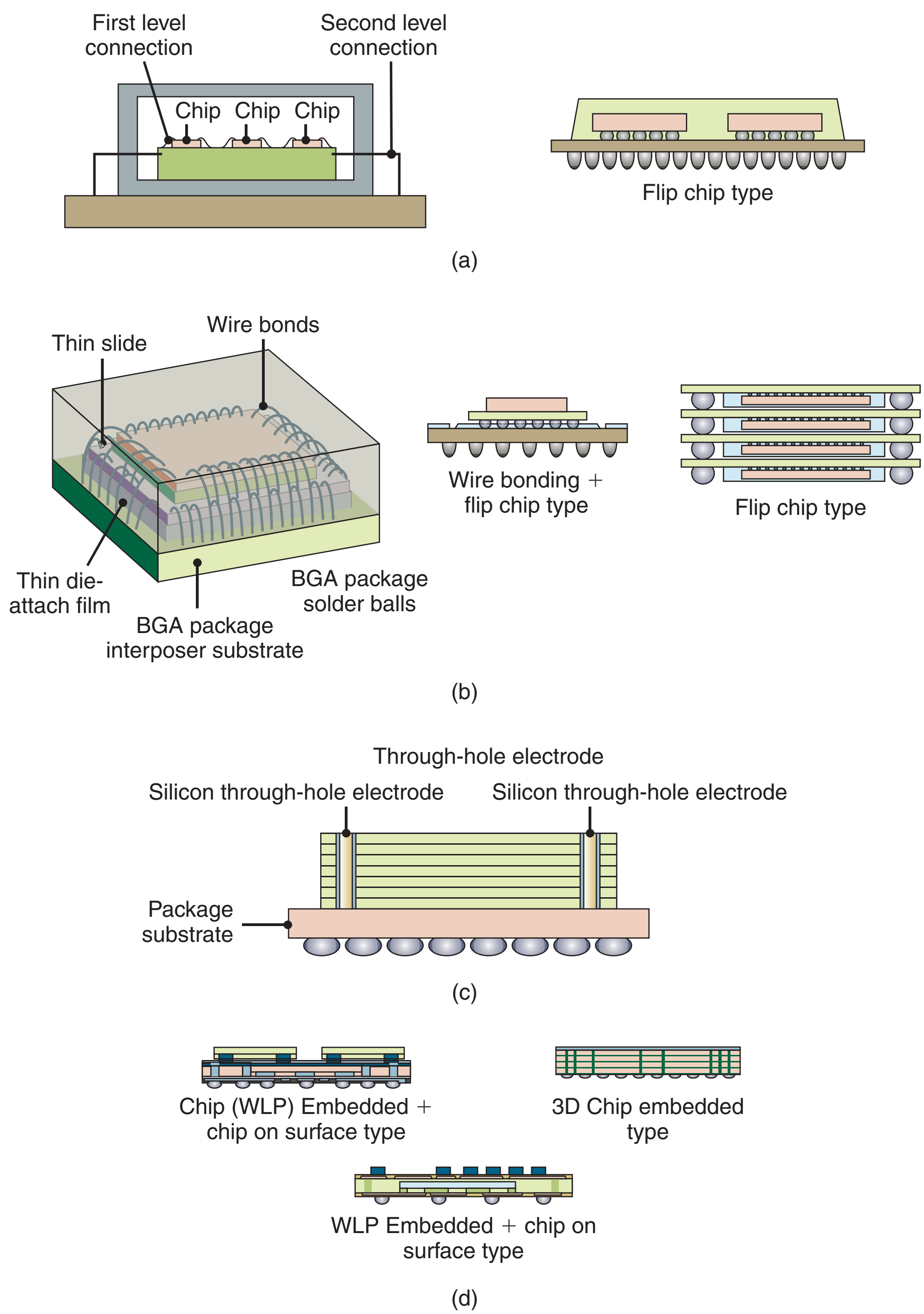


Figure 28.34: Major categories of system-in-package designs. (a) Horizontal placement, or multichip modules (MCMs); (b) interposer-type stacked structure; (c) interposerless stacked structure with through-silicon vias; and (d) embedded structure. WLP = wafer level package.

28.13 Printed Circuit Boards

Packaged ICs seldom are used alone; they usually are combined with other ICs to serve as building blocks of a yet larger system. A **printed circuit board** (PCB) is the substrate for the final interconnections among all of the completed chips. It serves as the communication link between the outside world and the microelectronic circuitry within each packaged IC.

In addition to possessing ICs, circuit boards usually contain discrete circuit components, such as resistors and capacitors, which take up too much "real estate" on the limited silicon surface, have special power-dissipation requirements, or cannot be implemented on a chip. Other common discrete components are inductors (that cannot be integrated onto the silicon surface), high-performance transistors, large capacitors, precision resistors, and crystals (for frequency control).

A PCB is basically a plastic (resin) material, containing several layers of copper foil (Fig. 28.35). *Single-sided PCBs* have copper tracks on only one side of an insulating substrate; *double-sided boards* have copper tracks on both sides. *Multilayered boards* also can be constructed from alternating layers of copper and insulator, but single-sided boards are the simplest form of circuit board.

Double-sided boards usually have locations where electrical connectivity is established between the features on both sides of the board; this is accomplished with **vias** (Fig. 28.35). Multilayered boards can have partial, buried, or through-hole vias to allow for extremely flexible PCBs. Double-sided and multilayered boards are beneficial, because IC packages can be bonded to both sides of the board, thus allowing for more compact designs.

The insulating material is usually an epoxy resin 0.25 to 3 mm thick, reinforced with an epoxy-glass fiber, and is referred to as E-glass (Section 9.2.1 and Fig. 9.3). The assembly is produced by impregnating sheets of glass fiber with epoxy, and pressing the layers together between hot plates or rolls. The heat and pressure cure the board, resulting in a stiff and strong basis for printed circuit boards.

The boards are first sheared to a desired size, and about 3-mm-diameter locating holes are then drilled or punched into the board's corners, to permit alignment and proper location of the board within the chip-insertion machines. Holes for vias and connections are punched or produced through CNC drilling (Section 37.3); stacks of boards can thus be drilled simultaneously to increase production rates.

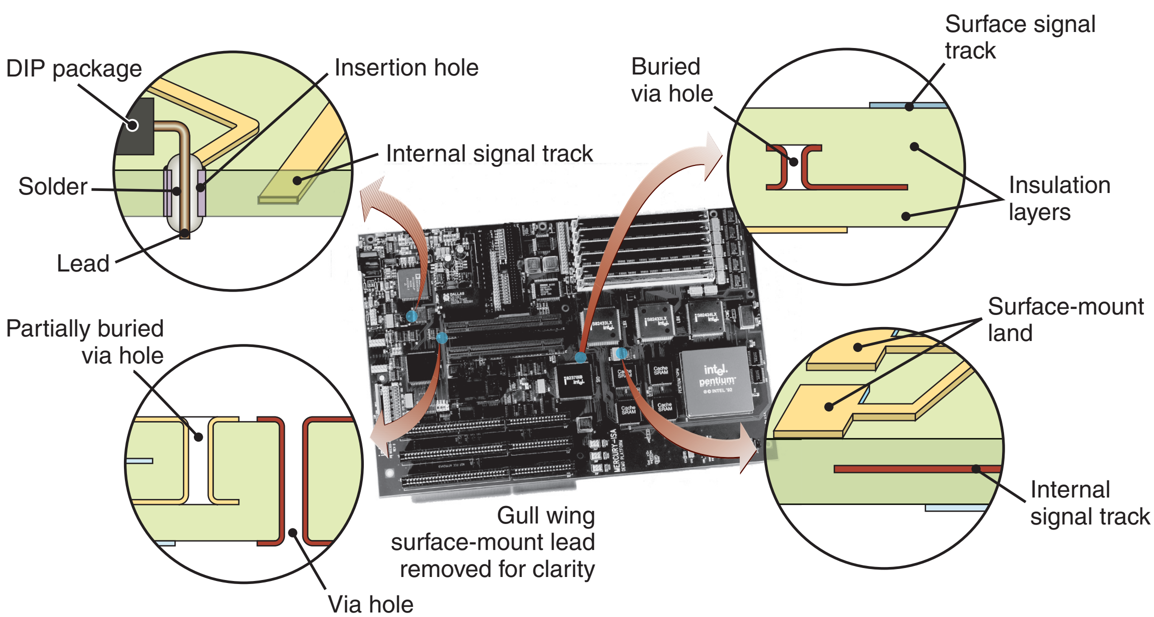


Figure 28.35: Printed circuit board structures and design features.

The conductive patterns on circuit boards are defined by lithography, although originally they were produced through screen-printing technologies, hence the term *printed* circuit board or *printed wiring board* (PWB). In the *subtractive method*, a copper foil is bonded to the circuit board. The desired pattern on the board is defined by a positive mask developed through photolithography, and the remaining copper is removed through wet etching. In the *additive method*, a negative mask is placed directly onto an insulator substrate, to define the desired shape. Electroless plating and electroplating of copper then serve to define the connections, tracks, and lands on the circuit board.

The ICs and other discrete components are then soldered to the board. This is the final step in making both the ICs, and the microelectronic devices they contain; into larger systems, through connections on PCBs. *Wave soldering* and *reflow paste soldering* (Section 32.3.3 and Example 32.1) are the preferred methods of soldering ICs onto circuit boards.

Basic *design considerations* in laying out PCBs are:

1. Wave soldering should be used only on one side of the board, thus all through-hole mounted components should be inserted from the same side of the board.

2. Surface-mount devices placed on the insertion side of the board must be reflow soldered in place; surface-mount devices on the lead side can be wave soldered.

3. To allow good solder flow in wave soldering, IC packages should be laid out carefully on the PCB. Inserting the packages in the same direction is advantageous for automated placing, because random orientations can cause problems in the flow of solder across all of the connections.

4. The spacing of ICs is determined mainly by the requirement to remove heat during their operation. Sufficient clearance between packages and adjacent boards is thus required to allow forced airflow and heat convection.

5. There should be sufficient space also around each IC package to allow for reworking and repairing without disturbing adjacent devices.

28.14 Roll-to-Roll Printing of Flexible Electronics

A relatively new development for electronic device manufacturing is the use of **roll-to-roll** printing, also known as *R2R processing*. This approach uses various rotary printing processes to transfer functional ink to a sheet or continuous feed of a flexible substrate (usually metal, paper, or plastic), which is then wound into a roll. While the resolution that can be achieved in photolithography is much better than that for R2R, the main advantage of R2R is that large areas can be printed quickly and relatively inexpensively. For example, printing speeds of 10 m^2/s are now routine with R2R, while soft lithography generally reaches speeds around 1×10^{-4} m^2/s. R2R has been widely applied in flexible solar panels, *organic light emitting diodes* (OLEDs) used in curved displays, and radio frequency identification (RFID) circuits, where large aerials can be printed at the same time as the circuit. There are numerous applications under development, including fuel cells, advanced batteries, multilayer capacitors, X-ray and other radiation detectors, and membranes for chemical processing industries.

Roll-to-roll processing requires the use of special inks, which can be either dissolved solutions of ink in a carrier fluid, an emulsion of one liquid suspended in the other in the form of micro- or nanoscale droplets, or suspensions of particles in a carrier fluid. Several different inks have been developed, of which the following are particularly important:

- **Silver nanoparticles** can be suspended in water or other carriers. Deposited silver is useful as an electrode material. Nanoparticle suspensions of gold can also be inkjetted and used as a conductor.

- **Indium tin oxide** (ITO) is a conductor that is transparent in thin layers.

- **Organic polymers** can be produced with a wide variety of attractive properties, including semiconducting, conducting, photovoltaic, and electroluminescent formulations. Conductive polymers, for example, are based on poly(3,4-ethylene dioxitiophene) or PEDOT, while semiconductors are based on this polymer doped with styrene sulfonate or PEDOT:PSS. Special formulations are available for each of these functions.

- *Inorganic semiconductors*, such as *copper indium gallium selenide (CIGS)*, can be deposited through roll-to-roll techniques, and are common materials used in solar cells.

- *Electroluminescent* materials can be produced from inorganic materials, such as copper-doped phosphor on a plastic film.

Substrates can be glass or silicon (for inkjet printing), but they can also be polymer films, often poly(ethylene terephthalate) (PET) for roll-to-roll processing, or paper. A flexible substrate allows continuous printing, significantly increasing speed and lowering costs. Furthermore, flexible electronics enable such designs as flexible displays, solar panels that conform to an automobile roof, and wearable electronics. However, since the resolution is not as good as photolithography, R2R is limited to applications that are not as demanding, such as *radio frequency identification* (RFID) or *smart labels*, solar panels, and flexible displays.

Commonly used printing techniques in R2R are shown in Fig. 28.36, and include the following:

- **Inkjet printing** allows the deposition of a wide variety of inks in the form of small droplets whose position can be precisely controlled.

- **Gravure**, or *rotogravure*, printing, is commonly used in printing magazines and catalogs. The technique requires the transfer of ink from cavities in the engraved or gravure cylinder to the web. Ink is continuously fed into the rotating gravure cylinder, using a doctor blade (Fig. 28.36a) or squeegee to remove excess ink. The gravure cylinder can have varying depths to control the amount of deposited ink over the printed image.

- **Flexographic printing** is similar to gravure printing, except that the ink is transferred to an intermediate anilox roller. This method allows uniform distribution of ink across the roller, which can be transferred to the printing cylinder which then transfers its pattern to the web.

- **Screen printing** or *screening* is available in two forms. *Flat bed screen printing* uses a moving squeegee to push an ink paste through a screen with a specific pattern. This technique allows the deposition of very thick layers, and is a preferred technique for producing electrodes where high conductivity is needed.

- **Rotary screen printing** is a continuous form of screen printing, where the ink is inside the rotating screen and is pushed outwards by a stationary squeegee.

Another important process in R2R is soft lithography, referred to as **self-aligned imprint lithography** (SAIL). Flexible stamps are used to deposit multiple layers of photopolymers, which are cured with UV light, allowing the stamp to be peeled off easily. SAIL consists of deposition of base films, followed by multiple stamp layers, and completed by wet and dry etching to produce high-resolution patterns.

R2R involves printing layers one at a time, as in the approach for integrated circuits (Fig. 28.2), but the layers are printed on a continuous web at very high rates. Each layer requires its own printing steps, just as a color magazine requires layers of red, blue, yellow, and black ink. Once a multiple-layer device has been printed, it is laminated to protect the circuit and to provide operational stability. *Cold lamination* involves application of a polymer film with a pressure-sensitive adhesive; *hot melt lamination* uses a material that becomes adhesive when heated; other forms of lamination are also in use.

Roll-to-roll printing is not yet in common use, but is being developed rapidly. It is expected that R2R will allow mass production of inexpensive circuits, enabling such innovations as the Internet of Things (Section 39.8.1).

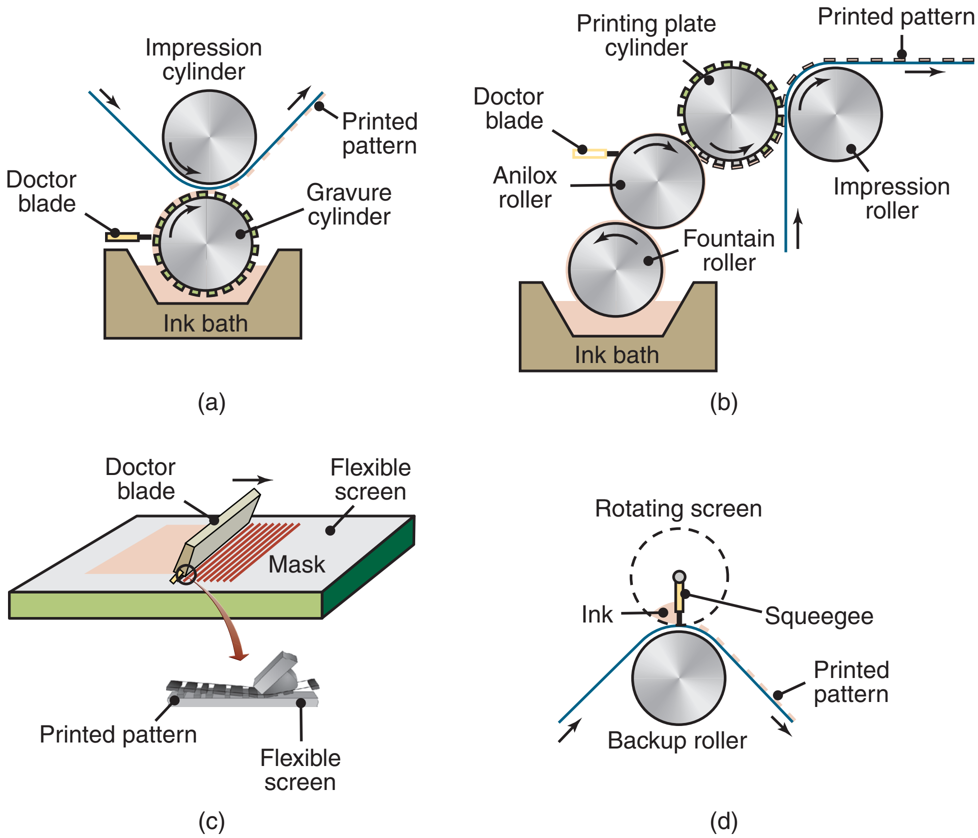
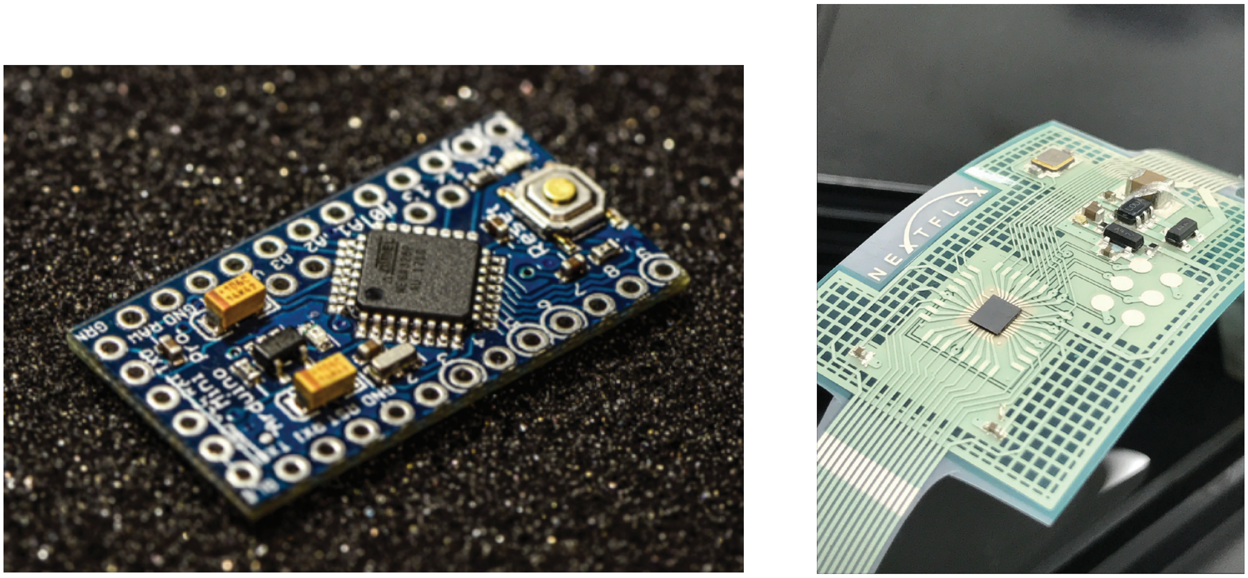


Figure 28.36: Schematic illustration of roll-to-roll processing approaches. (a) Gravure or Rotogravure, using an engraved roller or rotograve with a rubber doctor blade to transfer a desired ink pattern to the flexible web. (b) Flexography, where ink is transferred from an *anilox roller* (a hard cylinder, coated with a ceramic with dimples or micro cavities embedded into its surface) to raised portions of a printing plate which is then transferred to a flexible substrate. (c) Flat bed screen printing (see also Fig. 32.7). (d) Rotary screen printing.

28.15 Flexible Hybrid Electronics

Flexible hybrid electronic (FHE) devices combine the low cost and functionality of printed electronics (Section 28.14) with the high performance of standard silicon devices discussed throughout this chapter. Instead of packaging and bonding as in Section 28.11, the silicon die is directly attached to a printed substrate, which also allows encapsulation of the system. If the die has been made sufficiently thin (through back grinding), and is placed on a flexible (polymer or metal foil) substrate, then the system can bend and stretch, making them useful for a large number of applications. For example, a flexible device can be produced on an adhesive-backed film and then applied like a sticker to a machine component, clothing, skin, etc.

Substrate materials depend on the operating temperature, but PET is common for applications below 120°C, and polyimide for higher temperatures, although other options, such as flexible glass or polyurethane, are also used. The substrates are coated with copper that is then patterned and etched using photolithography. Direct printing of substrates using different conductive and insulating inks, as appropriate, has also been performed. The materials selected depend on the application, and often, on the expected life. Smart medical patches placed on the skin may have an intended life of a few hours to a few days,



(a) (b)

Figure 28.37: (a) A conventional Arduino Mini package, using a rigid circuit board and conventionally encapsulated die. (b) A flexible hybrid electronic Arduino, produced on a polymer film, that can be easily attached to any flat or curved surface.

and therefore may not need protective encapsulation. On the other hand, flexible hybrid electronics incorporating sensors and placed on machines for monitoring purposes may be used for years and will need protective coatings.

Printing is commonly done with roll-to-roll approaches (Section 28.14) using inkjet, gravure offset printing, and screening using silver inks. Through-substrate vias can be produced through laser machining followed by filling or coating with the silver inks.

The silicon die must be reduced in thickness to around 50 μm, although 25 μm has also been achieved in order to achieve desired flexibility. There are no real restrictions on the type of functionality in the die, but the die generally has some microprocessing and communication capability. After attachment to the substrate, the backside of the die can be used to print devices such as strain gages or other sensors.

Case Study 28.1 A Flexible Arduino®

Arduino is an open-source, microcontroller-based electronics prototyping platform that utilizes versatile, easy-to-use hardware and software. Arduino has achieved a high degree of popularity with developers ranging from hobbyist/novices to seasoned experts because it is open source, with publicly available design files and low cost. Up until now, however, Arduino products have been built with traditionally packaged die microcontrollers which deliver high performance and functionality, but because of their rigid packages, they are difficult to integrate into newer sensor devices that may be flexible or curved in design.

The NextFlex® Manufacturing Institute developed a process flow for manufacturing a flexible Arduino, and at the same time reduced the number of traditional process steps by more than 60% compared to that of a rigid board. NextFlex replaced the traditional circuit board with a thin, flexible plastic sheet and used digital printing processes for circuit elements. A thin bare die attached to the substrate eliminated traditional microcontroller packaging while further enabling flexibility of the product. The new process translates to an anticipated savings in manufacturing time and cost, as well as a significant reduction in the end-product weight; the flexible Arduino is only a third of the weight of the rigid Arduino Mini board.

At the project's conclusion, NextFlex proved the robustness of the FHE manufacturing process by producing multiple functional samples of a flexible Arduino system while integrating the degree of complexity required for Internet of Things (IoT) and sensor applications. Most importantly, the project's success is serving as a springboard for other FHE innovations.

Source: Courtesy NextFlex

Summary

- The microelectronics industry continues to develop rapidly, and possibilities for new device concepts and circuit designs appear to be endless. Fabrication of microelectronic devices and integrated circuits involves several different types of processes, many of which have been adapted from those of other fields in manufacturing.
- A rough shape of single-crystal silicon is first obtained by the Czochralski process. This shape is ground to a cylinder of well-controlled dimensions, and a notch or flat is machined into the cylinder. The cylinder is then sliced into wafers, which are ground on their edges and subjected to chemical–mechanical polishing to complete the wafer.
- After bare wafers have been prepared, they undergo repeated oxidation or film deposition, and lithographic or etching steps to open windows in the oxide layer in order to access the silicon substrate.
- Wet etching is isotropic and relatively fast. Dry etching, using gas plasmas, is anisotropic and allows for more accurate lithography and large-scale integration of integrated circuits.
- After each of the processing cycles is completed, dopants are introduced into various regions of the silicon structure, through diffusion and ion implantation.
- The devices are then interconnected by multiple metal layers, and the completed circuit is packaged and made accessible through electrical connections.
- The packaged circuit and other discrete devices are then soldered to a printed circuit board for final installation.
- New approaches have been developed to allow the production of flexible electronic devices and flexible hybrid electronic devices, using advanced printing techniques. These approaches are referred to as roll-to-roll printing.

Key Terms

Accelerated-life testing

Bonding

Chemical–mechanical polishing

Chemical-vapor deposition

Chip

Chip on board

Contacts
Critical dimension
Czochralski process
Die
Diffusion
Dopants
Dry etching
Dry oxidation
Dual-in-line package
Electromigration
Epitaxy
Etching
Evaporation
Failure rate
Film deposition
Flexible hybrid electronics
Flip-chip on board
Gallium arsenide
Gravure printing
Immersion Lithography
Integrated circuit
Ion implantation
LELE process
Line width
Lithography
Masking
Metal-oxide-semiconductor field-effect transistor
Metallization
Microcontact printing
Micromolding in capillaries
Microtransfer molding
Oxidation
Packaging
Photoresist
Pitch splitting lithography
Planarization
Postbaking
Prebaking
Printed circuit board
Registration
Reliability
Reticle
Roll-to-roll printing
Rotary screen printing
SCALPEL
Screen printing
Selective oxidation
Selectivity
Self-aligned imprint lithography
Semiconductor
Silicon
Soft lithography
Sputtering
Surface-mount package
System in package
Very large scale integration
Vias
Wafer
Wet etching
Wet oxidation
Wire bonding
Yield

Bibliography

Anderson, B.L., and Anderson, R.L., **Fundamentals of Semiconductor Devices**, 2nd ed., McGraw-Hill, 2017.

Campbell, S.A., **Fabrication Engineering at the Micro- and Nanoscale**, Oxford, 2012.

Coombs, C.F., Jr., and Holden, H., **Printed Circuits Handbook**. 7th ed., McGraw-Hill, 2017.

Geng, H., **Semiconductor Manufacturing Handbook**, 2nd ed., McGraw-Hill, 2017.
Jackson, K.A., and Schroter, W. (eds.), **Processing of Semiconductors**, Wiley-VCH, 2011.
Jaeger, R.C., **Introduction to Microelectronic Fabrication**, 2nd ed., Pearson, 2001.
Khandpur, R.S., **Printed Circuit Boards: Design, Fabrication, and Assembly**, McGraw-Hill, 2005.
May, G.S., and Spanos, C.J., **Fundamentals of Semiconductor Manufacturing and Process Control**, Wiley, 2006.
Nishi, Y., and Doering, R., **Handbook of Semiconductor Manufacturing Technology**, 2nd ed., CRC Press, 2007.
Schroder, D.K., **Semiconductor Material and Device Characterization**, 3rd ed., Wiley-Interscience, 2015.
Sze, S.M., and Lee, M.-K., **Semiconductor Devices: Physics and Technology**, 3rd ed., Wiley, 2012.
Van Zant, P., **Microchip Fabrication**, 6th ed., McGraw-Hill, 2014.

Review Questions

28.1. Define the terms wafer, chip, die, device, integrated circuit, line width, registration, surface mount, accelerated-life testing, and yield.

28.2. Why is silicon the semiconductor most used in IC technology?

28.3. What do the abbreviations BJT, MOSFET, VLSI, IC, CVD, CMP, LELE, and DIP stand for?

28.4. Explain the differences between wet and dry oxidation.

28.5. Explain the differences between wet and dry etching.

28.6. What are the purposes of prebaking and postbaking in lithography?

28.7. Define selectivity and isotropy and their importance in relation to etching.

28.8. Compare the diffusion and ion-implantation processes.

28.9. Explain the difference between evaporation and sputtering.

28.10. What are the levels of interconnection?

28.11. Which is cleaner, a Class-10 or a Class-1 clean room?

28.12. Review Fig. 28.2 and describe the fabrication sequence for integrated circuits.

28.13. What is a via? Why is it important?

28.14. Describe how electrical connections are established between a die and a package.

28.15. What is a flip chip?

28.16. Describe the procedures of image splitting lithography and immersion lithography.

Qualitative Problems

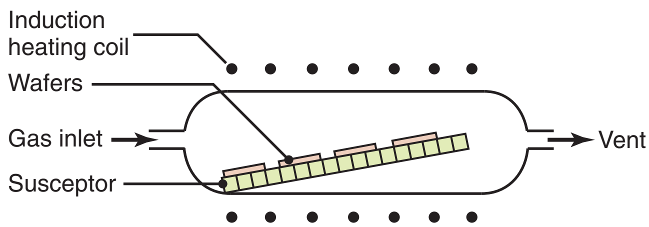
28.17. Comment on your observations regarding the contents of Fig. V.1.

28.18. Describe how n-type and p-type dopants differ.

28.19. How is silicon nitride used in oxidation?

28.20. How is epitaxy different from other techniques used for deposition? Explain.

28.21. Note that, in a horizontal epitaxial reactor (see the figure below), the wafers are placed on a stage (susceptor) that is tilted by a small amount, usually 1° to 3°. Explain why this is done.



28.22. The table that follows describes three wafer-manufacturing changes: increasing the wafer diameter, reducing the chip size, and increasing process complexity. Complete the table by filling in "increase," "decrease," or "no change," and indicate the effect that each change would have on the wafer yield and on the overall number of functional chips.

Change	Wafer yield	Number of functional chips
Increase wafer diameter		
Reduce chip size		
Increase process complexity		

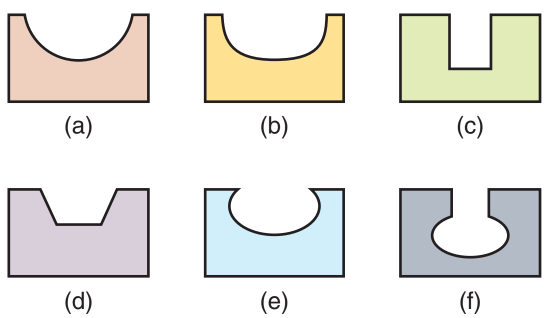
28.23. The speed of a transistor is directly proportional to the width of its polysilicon gate; thus, a narrower gate results in a faster transistor and a wider gate in a slower transistor. Knowing that the manufacturing process has a certain variation for the gate width, how would a designer modify the gate size of a critical circuit in order to minimize its variation in speed? Are there any negative effects of this change?

28.24. What is accelerated life testing? Why is it practiced?

28.25. Explain the difference between a die, a chip, and a wafer.

28.26. A common problem in ion implantation is channeling, in which the high-velocity ions travel deep into the material via channels along the crystallographic planes before finally being stopped. How could this effect be avoided? Explain.

28.27. Examine the hole profiles shown below and explain how they might be produced.



28.28. Referring to Fig. 28.24, sketch the shape of the holes generated from a circular mask.

Quantitative Problems

28.29. A certain wafer manufacturer produces two equal-sized wafers, one containing 500 chips and the other containing 300 chips. After testing, it is observed that 50 chips on each wafer are defective. What are the yields of the two wafers? Can any relationship be drawn between chip size and yield?

28.30. A chlorine-based polysilicon etch process displays a polysilicon:resist selectivity of 4:1 and a polysilicon:oxide selectivity of 50:1 How much resist and exposed oxide will be consumed in etching 400 nm of polysilicon? What would the polysilicon:oxide selectivity have to be in order to lose only 5 nm of exposed oxide?

28.31. During a processing sequence, four silicon-dioxide layers are grown by oxidation: 500 nm, 200 nm, 50 nm, and 25 nm. How much of the silicon substrate is consumed?

28.32. A certain design rule calls for metal lines to be no less than 2 μm wide. If a 1 μm-thick metal layer is to be wet etched, what is the minimum photoresist width allowed (assuming that the wet etch is perfectly isotropic)? What would be the minimum photoresist width if a perfectly anisotropic dry-etch process were used?

28.33. Using Fig. 28.20, obtain mathematical expressions for the etch rate as a function of temperature.

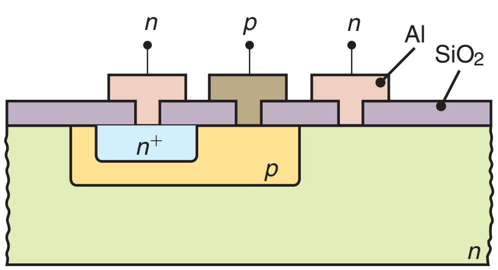
28.34. If a square mask of side length 100 μm is placed on a $\{100\}$ plane and oriented with a side in the $\langle 110 \rangle$ direction, how long will it take to etch a hole 4 μm deep at 80°C using thylene-diamine/pyrocatechol? Sketch the resulting profile.

28.35. Obtain an expression for the width of the trench bottom as a function of time for the mask shown in Fig. 28.24b.

28.36. A polyimide photoresist needs 125 mJ/cm^2 per micrometer of thickness in order to develop properly. How long does a 175 μm film need to develop when exposed by a 1100 W/m^2 light source?

Synthesis, Design, and Projects

28.37. The accompanying figure shows the cross section of a simple npn bipolar transistor. Develop a process flow chart to fabricate this device.



28.38. Describe products that would not exist today without the knowledge and techniques described in this chapter. Explain.

28.39. Inspect various electronic and computer equipment, take them apart as much as you can, and identify components that may have been manufactured by the techniques described in this chapter.

28.40. Describe your understanding of the important features of clean rooms and how they are maintained.

28.41. Make a survey of the necessity for clean rooms in various industries, including the medical, pharmacological, and aerospace industries, and what their requirements are.

28.42. Review the technical literature, and give further details regarding the type and shape of the abrasive wheel used in the wafer-cutting process shown in Step 2 in Fig. 28.2 (see also Chapter 26).

28.43. List and discuss the technologies that have enabled manufacturing of the products described in this chapter.

28.44. Estimate the time required to etch a spur gear blank from a 75-mm-thick slug of silicon.

28.45. Microelectronic devices may be subjected to hostile environments, such as high temperature, humidity, and vibration, as well as physical abuse, such as being dropped onto a hard surface. Describe your thoughts on how you would go about testing these devices for their endurance under these conditions. Are there any industry standards regarding such tests? Explain.

28.46. Review the specific devices, shown in Fig. V.2. Choose any one of these devices, and investigate what they are, what their characteristics are, how they are manufactured, and what their costs are.

Chapter 29

Fabrication of Microelectromechanical Devices and Systems and Nanoscale Manufacturing

- Many of the processes and materials used for manufacturing microelectronic devices are also used for manufacturing micromechanical devices and microelectromechanical systems; this chapter investigates topics in the production of very small mechanical and electromechanical products. The chapter begins with considerations of micromachining and surface machining of mechanical structures from silicon.
- The LIGA process and its variations are then described, along with micromolding, EFAB, and various other techniques for replicating small-scale mechanical devices.
- Solid free-form fabrication processes are sometimes suitable for the production of MEMS and MEMS devices.
- The chapter ends with a discussion of the emerging area of nanoscale manufacturing.

Typical parts made: Sensors, actuators, accelerometers, optical switches, ink-jet printing mechanisms, micromirrors, micromachines, and microdevices.

Alternative methods: Fine blanking, small-scale machining, microforming.

29.1 Introduction

The preceding chapter dealt with the manufacture of integrated circuits and products that operate purely on electrical or electronic principles called **microelectronic devices**. These semiconductor-based devices often have the common characteristic of extreme miniaturization. A large number of devices exist that are mechanical in nature and are of a similar size as microelectronic devices. A **micromechanical device** is a product that is purely mechanical in nature and has dimensions between a few mm and atomic length scales, such as some very small gears and hinges.

A **microelectromechanical device** is a product that combines mechanical and electrical or electronic elements at these very small length scales. A **microelectromechanical system** (MEMS) is a microelectromechanical device that also incorporates an integrated electrical system into one product. Common examples of micromechanical devices are sensors of all types (Fig. 29.1). Microelectromechanical systems include accelerometers in mobile phones, gyroscopes and GPS and air-bag sensors in automobiles, and digital micromirror devices. Parts made by **nanoscale manufacturing** generally have dimensions that are between 10^{-6} and 10^{-9} m, as described in Section 29.6.

Many of the materials and manufacturing methods and systems described in Chapter 28 also apply to the manufacture of microelectromechanical devices and systems. However, microelectronic devices are semiconductor-based, whereas microelectromechanical devices and portions of MEMS do not have this restriction. Thus, many more materials and processes are suitable for these applications. Regardless, silicon often is used because several highly advanced and reliable manufacturing processes using silicon have been developed for microelectronic applications. This chapter emphasizes the manufacturing processes

Figure 29.1: Exploded view of the camera in a smart phone. Perhaps the most widespread use of MEMS devices is in cameras and sensors of all kinds. *Source:* Shutterstock/Who is Danny.

that are applicable specifically to microelectromechanical devices and systems, but it should be realized that processes and concepts such as lithography, metallization, etching, coating, and packaging discussed in Chapter 28 still apply.

MEMS and MEMS devices are rapidly advancing, and new processes or variations on existing processes are continually being developed. A significant leap in the numbers of commercial MEMS devices occurred in the past few years as mobile phones and tablet computers integrated accelerometers and gyroscopes into their products. However, it should be recognized that many of the processes described in this chapter have not yet become widespread, but are of interest to researchers and practitioners in MEMS and hold great potential for future applications.

29.2 Micromachining of MEMS Devices

The topics described in the preceding chapter dealt with the manufacture of integrated circuits and products that operate based purely on electrical or electronic principles. These processes also are suitable for manufacturing devices that incorporate mechanical elements or features as well. The following four types of devices can be made through the approach described in Fig. 28.2:

1. **Microelectronic devices** are semiconductor-based devices that often have the common characteristics associated with extreme miniaturization and use electrical principles in their design.

2. **Micromechanical devices** are products that are purely mechanical in nature and have dimensions between atomic length scales and a few mm. Very small gears and hinges are examples.

3. **Microelectromechanical devices** are products that combine mechanical and electrical or electronic elements at very small length scales. Most sensors are examples of microelectromechanical devices.

4. **Microelectromechanical systems** are microelectromechanical devices that also incorporate an integrated electrical system in one product. Microelectromechanical systems are rare compared with microelectronic, micromechanical, or microelectromechanical devices, typical examples being air-bag sensors and digital micromirror devices.

The production of features from μm to mm in size is called *micromachining*. MEMS devices have been constructed from **polycrystalline silicon** (*polysilicon*) and **single-crystal silicon** because the technologies for integrated-circuit manufacture, described in Chapter 28, are well developed and exploited for these devices; other, new processes also have been developed that are compatible with the existing processing steps. The use of anisotropic etching techniques (Section 28.8.1) allows the fabrication of devices with well-defined walls and high aspect ratios; for this reason, some MEMS devices have been fabricated from single-crystal silicon.

One of the difficulties associated with the use of silicon for MEMS devices is the high adhesion between components encountered at small length scales and the associated rapid wear (Section 33.5). Most commercial devices are designed to avoid friction by, for example, using flexing springs instead of bearings. However, this approach complicates designs and makes some MEMS devices not feasible. Consequently, significant research is being conducted to identify materials and lubricants that provide reasonable life and performance and that would allow sliding on the microscale without excessive wear.

Silicon carbide, diamond, and metals (such as aluminum, tungsten, and nickel) have been investigated as potential MEMS materials; various lubricants also have been investigated. It is known, for example, that surrounding the MEMS device in a silicone oil practically eliminates adhesive wear, but it also limits the

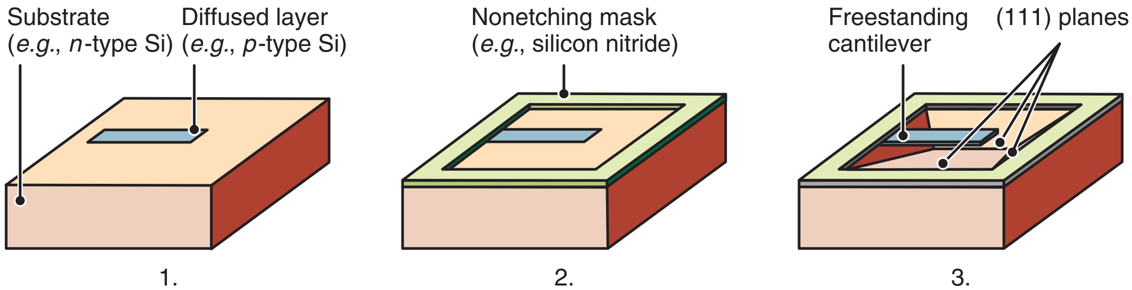


Figure 29.2: Schematic illustration of bulk micromachining. 1. Diffuse dopant in desired pattern. 2. Deposit a pattern and masking. 3. Orientation-dependent etching (ODE) leaves behind a freestanding structure. *Source:* Courtesy of K.R. Williams, Agilent Technologies.

performance of the device. Self-assembling layers of polymers also are being investigated, as well as novel and new materials with self-lubricating characteristics. However, the tribology of MEMS devices remains a main technological barrier to any further expansion of their already widespread use.

29.2.1 Bulk Micromachining

Until the early 1980s, *bulk micromachining* was the most common method of machining at micrometer scales. This process uses orientation-dependent etches on single-crystal silicon (see Fig. 28.15b), an approach that depends on wet etching (Section 28.8) into a surface and stopping on certain crystal faces, doped regions, and etchable films to form a desired structure. As an example of this process, consider the fabrication of the silicon cantilever shown in Fig. 29.2. Using the masking techniques described in Section 28.7, the process changes a rectangular patch of the $n-$type silicon substrate to $p-$type silicon through boron doping. Etchants such as potassium hydroxide will not be able to remove heavily boron doped silicon; hence, this patch will not be etched.

A mask is then produced—for example, with silicon nitride on silicon. When etched with potassium hydroxide, the undoped silicon will be removed rapidly, while the mask and the doped patch will essentially be unaffected. Etching progresses until the (111) planes are exposed in the $n-$type silicon substrate; they undercut the patch, leaving a suspended cantilever (as shown in Fig. 29.2).

29.2.2 Surface Micromachining

Although bulk micromachining is useful for producing very simple shapes, it is restricted to single-crystal materials because polycrystalline materials will not wet etch at different rates in different directions. Many MEMS applications require the use of other materials or material combinations; hence, alternatives to bulk micromachining are needed. One such method is *surface micromachining*, the basic steps of which are illustrated for silicon devices in Fig. 29.3.

In surface micromachining, a spacer or **sacrificial layer** is deposited onto a silicon substrate coated with a thin dielectric layer, called an *isolation*, or *buffer*, *layer*. Phosphosilicate glass deposited by chemical-vapor deposition is the most common material for a spacer layer, because it etches very rapidly in hydrofluoric acid, a property that is useful in step 5. Step 2 in Fig. 29.3 shows the spacer layer after the application of masking and etching. At this stage, a structural thin film is deposited onto the spacer layer; the film can be polysilicon, metal, metal alloy, or a dielectric (step 3 in Fig. 29.3). The structural film is then patterned, usually through dry etching, in order to maintain vertical walls and tight dimensional tolerances. Finally, wet etching of the sacrificial layer leaves a freestanding, three-dimensional structure, as shown in step 5

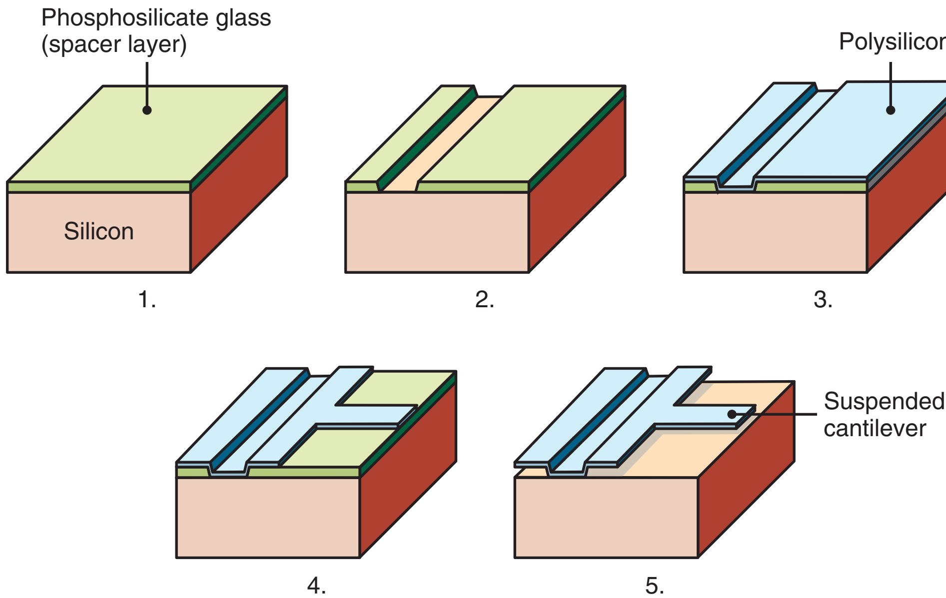


Figure 29.3: Schematic illustration of the steps in surface micromachining: 1. Deposition of a phosphosilicate glass (PSG) spacer layer; 2. Lithography and etching of spacer layer; 3. Deposition of polysilicon; 4. Lithography and etching of polysilicon; 5. Selective wet etching of PSG, leaving the silicon substrate and deposited polysilicon unaffected.

of Fig. 29.3. Note that the wafer must be annealed to remove the residual stresses in the deposited metal before it is patterned; otherwise the structural film will severely warp once the spacer layer is removed (see also Section 2.11).

Figure 29.4 shows a microlamp that emits a white light when current passes through it; it has been produced through a combination of surface and bulk micromachining. The top patterned layer is a 2.2-μm

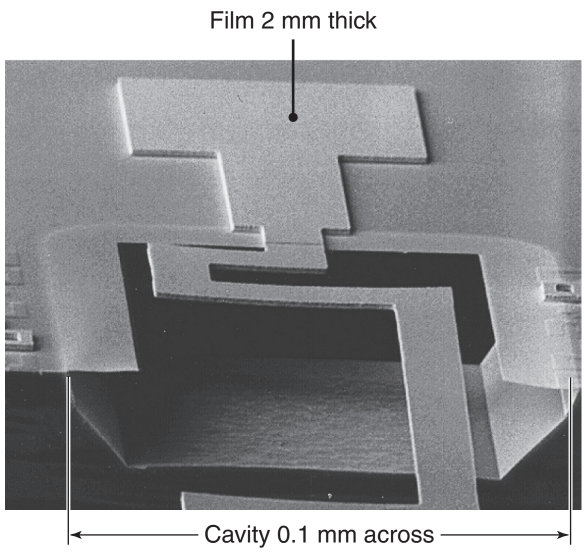


Figure 29.4: A microlamp produced from a combination of bulk and surface micromachining processes. *Source:* Courtesy of K.R. Williams, Agilent Technologies.

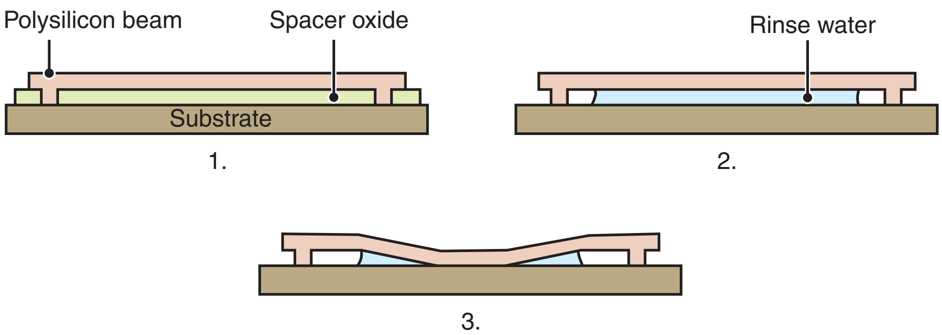


Figure 29.5: Stiction after wet etching: 1. Unreleased beam. 2. Released beam before drying. 3. Released beam pulled to the surface by capillary forces during drying. Once contact is made, adhesive forces prevent the beam from returning to its original shape. *Source:* After B. Bhushan.

layer of plasma-etched tungsten, forming a meandering filament and bond pad. The rectangular overhang is dry-etched silicon nitride. The steeply sloped layer is wet-hydrofluoric acid-etched phosphosilicate glass; the substrate is silicon, which is orientation-dependent etched.

The etchant used to remove the spacer layer must be selected carefully, as it must preferentially dissolve the spacer layer while leaving the dielectric, silicon, and structural film as intact as possible. With large features and narrow spacer layers, this task becomes very difficult, and etching can take many hours. To reduce the etching time, additional etched holes can be designed into the microstructures to increase access of the etchant to the spacer layer.

Another difficulty that must be overcome is **stiction** after wet etching, which can be described by considering the situation illustrated in Fig. 29.5. After the spacer layer has been removed, the liquid etchant is dried from the wafer surface. A meniscus forms between the layers, resulting in capillary forces that can deform the film and cause contraction of the substrate as the liquid evaporates. Since adhesive forces are more significant at small length scales, it is possible that the film may *stick* permanently to the surface; thus, the desired three-dimensional features will not be produced.

Example 29.1 Surface Micromachining of a Hinge

Surface micromachining is a widespread technology for the production of MEMS, with applications that include accelerometers, pressure sensors, micropumps, micromotors, actuators, and microscopic locking mechanisms. Often, these devices require very large vertical walls, which cannot be manufactured directly because the high vertical structure is difficult to deposit. This obstacle is overcome by machining large, flat structures horizontally, and then rotating or folding them into an upright position, as shown in Fig. 29.6.

Figure 29.6a shows a micromirror that has been inclined with respect to the surface on which it was manufactured. Such systems can be used for reflecting light (that is oblique to a surface) onto detectors or towards other sensors. It is apparent that a device which has such depth and has the aspect ratio of the deployed mirror is very difficult to machine directly. Instead, it is easier to surface micromachine the mirror along with a linear actuator, and then to fold the mirror into a deployed position. In order to do so, special hinges (as shown in Fig. 29.6b) are integrated into the design.

Figure 29.7 shows the cross-section of a hinge during its manufacture. The following steps are involved in the production of the hinges:

1. A 2-μm-thick layer of phosphosilicate glass is first deposited onto the substrate material.
2. A 2-μm-thick layer of polysilicon (Poly1 in step 1 in Fig. 29.7) is deposited onto the PSG, patterned by photolithography, and dry etched to form the desired structural elements, including the hinge pins.
3. A second layer of sacrificial PSG with a thickness of 0.5 μm is deposited (step 2 in Fig. 29.7).
4. The connection locations are etched through both layers of PSG (step 3 in Fig. 29.7).
5. A second layer of polysilicon (Poly2 in step 4 in Fig. 29.7) is deposited, patterned, and etched.
6. The sacrificial layers of PSG are then removed by wet etching.

Hinges such as these have very high friction. Thus, if mirrors (as shown in Fig. 29.6) are manipulated manually and carefully with probe needles, they will remain in position. Often, such mirrors will be combined with linear actuators to precisely control their deployment.

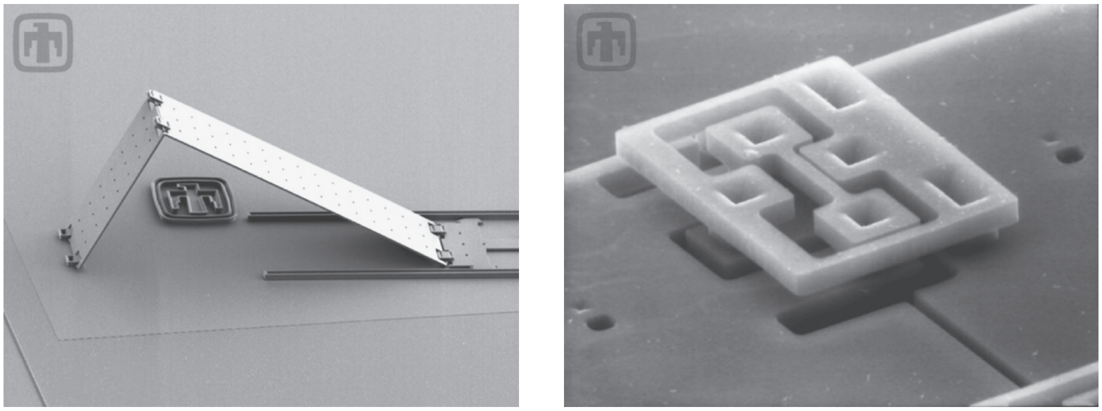
Case Study 29.1 Digital Micromirror Device

An example of a commercial MEMS-based product is the *digital pixel technology* (DPT$^{\text{TM}}$) device, illustrated in Fig. 29.8. This device uses an array of *digital micromirror devices* (DMD) to project a digital image, as in movie theater projection systems, nano projectors (Fig. 29.9), or in the CLIP process (Fig. 20.9). The aluminum mirrors can be tilted so that light is directed into or away from the optics that focus light onto a screen. That way, each mirror can represent a pixel of an image's resolution. The mirror allows light or dark pixels to be projected, but levels of gray also can be accommodated. Since the switching time is about 15 μs (which is much faster than the human eye can respond), the mirror will switch between the on and off states in order to reflect the proper dose of light to the optics.

The fabrication steps for producing the DMD device are shown in Fig. 29.10. This sequence is similar to that of other surface micromachining operations, but has the following important differences:

- All micromachining steps take place at temperatures below 400°C, which is sufficiently low to ensure that no damage occurs to the electronic circuit.
- A thick silicon dioxide layer is deposited and is chemical–mechanical polished (Section 26.7) to provide an adequate foundation for the MEMS device.
- The landing pads and electrodes are produced from aluminum, which is deposited by sputtering.
- High reliability requires low stresses and high strength in the torsional hinge, which is produced from a proprietary aluminum alloy.
- The MEMS portion of the DMD is very delicate, and special care must be taken in separating the dies. When completed, a wafer saw (see Fig. 28.6c) cuts a trench along the edges of the DMD, which allows the individual dice to be broken apart at a later stage.
- A special step deposits a layer that prevents adhesion between the yoke and landing pads.
- The DMD is placed in a hermetically sealed ceramic package (Fig. 29.11) with an optical window.

An array of such mirrors represents a grayscale screen. Using three mirrors (one each for red, green, and blue light) for each pixel results in a color image with millions of discrete colors. Digital pixel technology is widely applied in digital projection systems, high-definition television, and other optical equipment. However, to produce the device shown in Fig. 29.8 requires much more than two-and-one-half-dimensional features, so full three-dimensional, multipart assemblies have to be manufactured.



(a) (b)

Figure 29.6: (a) SEM image of a deployed micromirror, incorporating six hinges. (b) Detail of a micromirror hinge. *Source:* Courtesy of Sandia National Laboratories.

SCREAM. Another method for making very deep MEMS structures is the SCREAM (*single-crystal silicon reactive etching and metallization*) process, depicted in Fig. 29.12. In this technique, standard lithography and etching processes produce trenches 10–50 μm deep, which are then protected by a layer of chemically vapor deposited silicon oxide. An anisotropic-etching step removes the oxide only at the bottom of the trench, and the trench is then extended through dry etching. An isotropic etching step (using sulfur hexafluoride, SF_6) laterally etches the exposed sidewalls at the bottom of the trench. This undercut (when it overlaps adjacent undercuts) releases the machined structures.

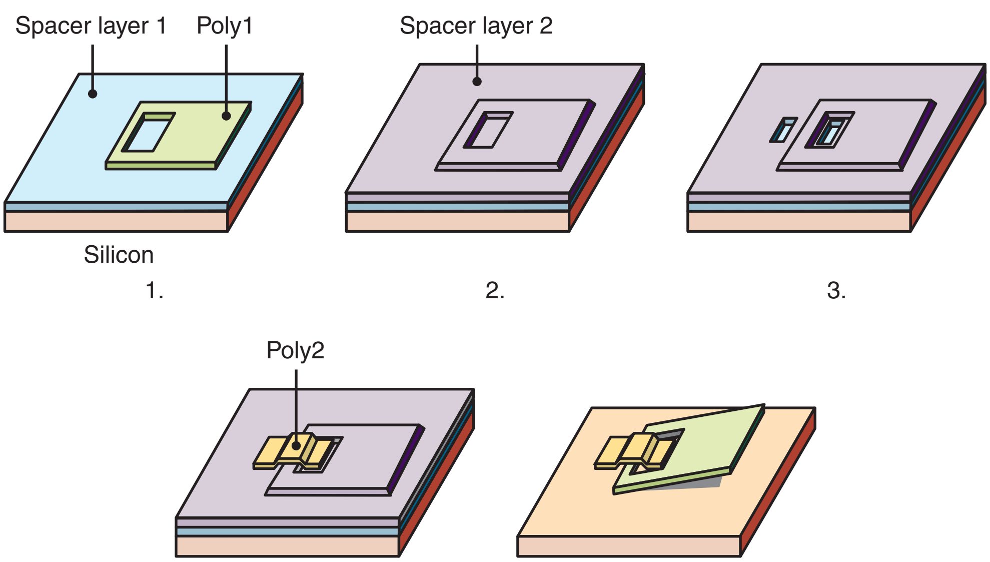


Figure 29.7: Schematic illustration of the steps required to manufacture a hinge. 1. Deposition of a phosphosilicate glass (PSG) spacer layer and polysilicon layer (see Fig. 29.3). 2. Deposition of a second spacer layer. 3. Selective etching of the PSG. 4. Deposition of polysilicon to form a staple for the hinge. 5. After selective wet etching of the PSG, the hinge can rotate.

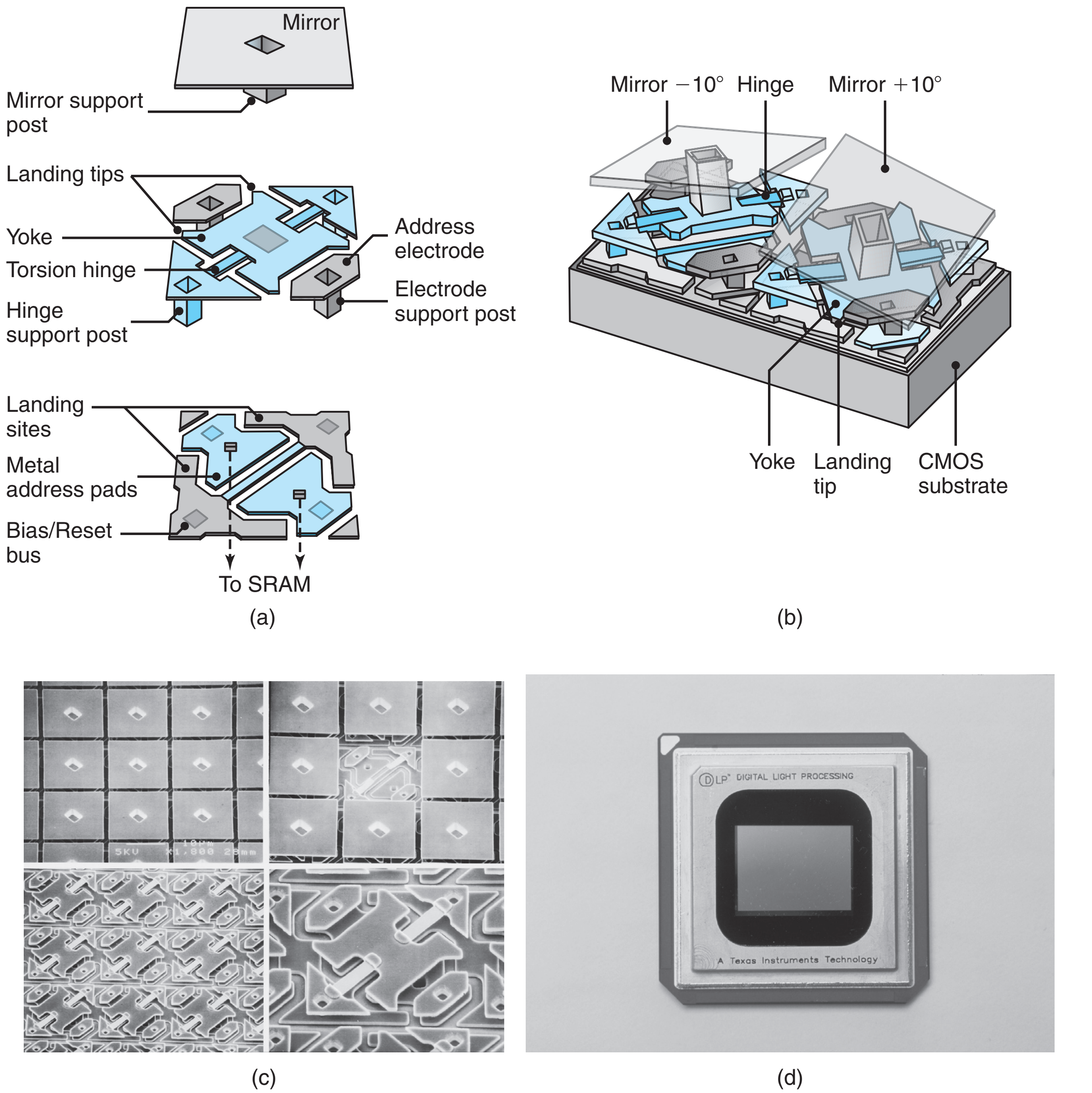


Figure 29.8: The Texas Instruments digital pixel technology (DPT) device. (a) Exploded view of a single digital micromirror device (DMD). (b) View of two adjacent DMD pixels. (c) Images of DMD arrays with some mirrors removed for clarity; each mirror measures approximately 17 μm on a side. (d) A typical DPT device used for digital projection systems, high-definition televisions, and other image display systems. The device shown contains 1,310,720 micromirrors and measures less than 50 mm per side. *Source:* Courtesy of Texas Instruments Corp.

Figure 29.9: A prototype pico projector based on DPT. *Source:* Courtesy of Texas Instruments Corp.

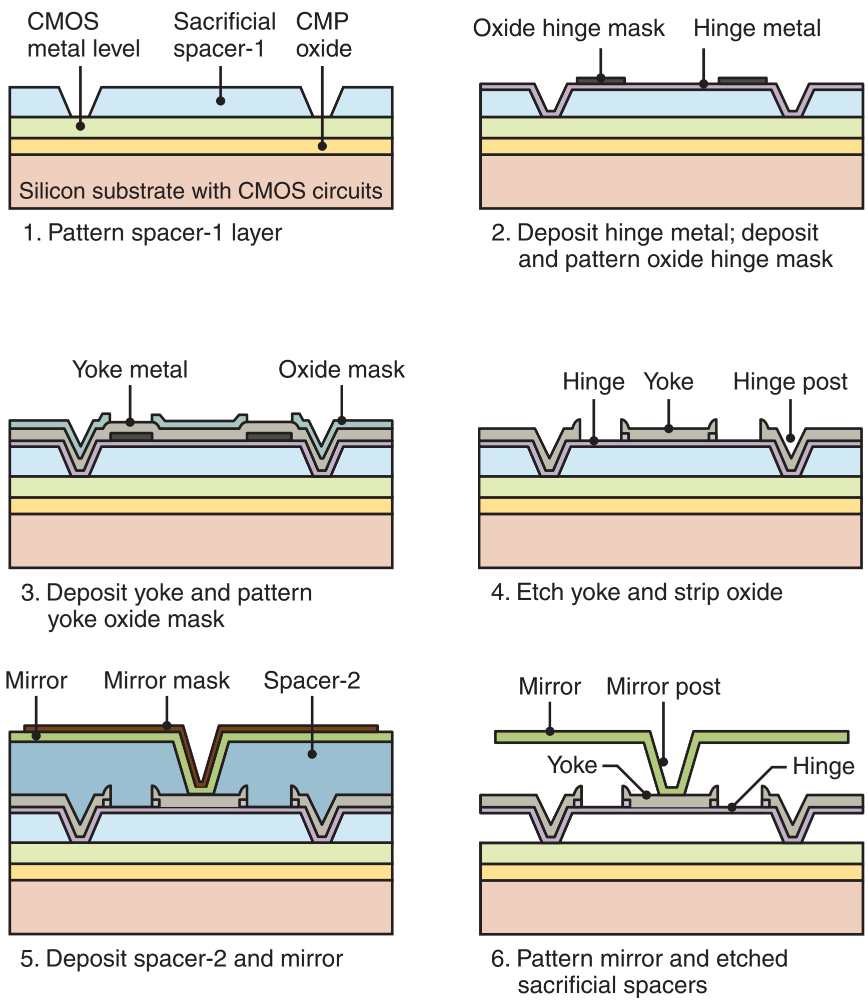


Figure 29.10: Manufacturing sequence for the Texas Instruments DMD device.

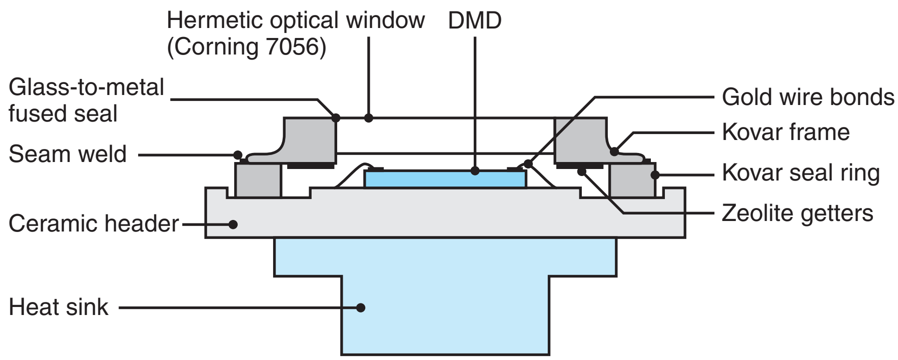


Figure 29.11: Ceramic flat-package construction used for the DMD device.

SIMPLE. An alternative to SCREAM is SIMPLE (*silicon micromachining by single-step plasma etching*), as depicted in Fig. 29.13. This technique uses a chlorine-gas-based plasma-etching process that machines *p*-doped or lightly doped silicon anisotropically, but heavily *n*-doped silicon isotropically. A suspended MEMS device can thus be produced in one plasma-etching device, as shown in the figure.

Some of the concerns with the SIMPLE process are:

- The oxide mask is machined, although at a slower rate, by the chlorine-gas plasma; therefore, relatively thick oxide masks are required.
- The isotropic etch rate is low, typically 50 nm/min; consequently, this is a very slow process.
- The layer beneath the structures will have developed deep trenches, which may affect the motion of free-hanging structures.

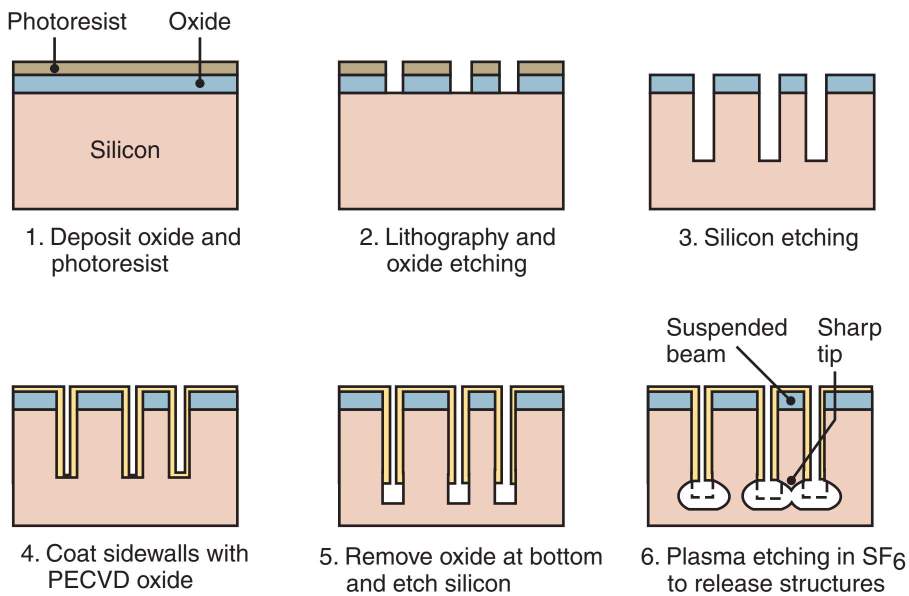


Figure 29.12: The SCREAM process. *Source:* After N. Maluf.

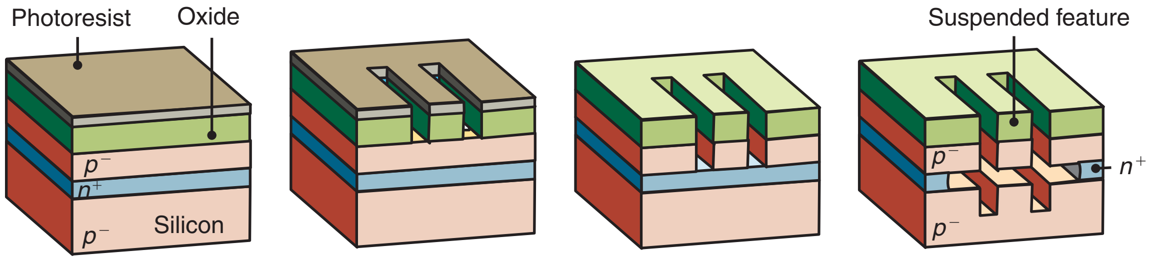


Figure 29.13: Schematic illustration of silicon micromachining by single-step plasma etching (SIMPLE) process.

Etching Combined with Diffusion Bonding. Very tall structures can be produced in crystalline silicon through a combination of *silicon-diffusion bonding and deep reactive-ion etching* (SFB–DRIE), as illustrated in Fig. 29.14. First, a silicon wafer is prepared with an insulating oxide layer, with the deep trench areas defined by a standard lithography procedure. This step is followed by conventional wet or dry etching to form a large cavity. A second layer of silicon is then fusion bonded to the oxide layer; the second silicon layer can be ground and lapped to the desired thickness if necessary. At this stage, integrated circuitry is manufactured through the steps outlined in Fig. 28.2. A protective resist is applied and exposed, and the desired trenches are then etched by deep reactive-ion etching to the cavity in the first layer of silicon.

Example 29.2 Operation and Fabrication Sequence for a Thermal Ink-jet Printer

Thermal ink-jet printers are among the most successful applications of MEMS to date. These printers operate by ejecting nano- or picoliters (10^{-12} liter) of ink from a nozzle towards the paper. Ink-jet printers use a variety of designs, but silicon-machining technology is most applicable to high-resolution printers. Note that a resolution of 1200 dpi requires a nozzle spacing of approximately 20 μm.

The mode of operation of an ink-jet printer is shown in Fig. 29.15. When an ink droplet is to be generated and expelled, a tantalum resistor (placed below a nozzle) is heated, which makes a thin film of ink form a bubble within 5 μs, with internal pressures reaching 1.4 MPa. The bubble then expands rapidly, and as a result, the fluid is forced rapidly out of the nozzle. Within 24 μs, the tail of the ink-jet droplet separates because of surface tension, the heat source is turned off, and the bubble collapses inside the nozzle. Within 50 μs, sufficient ink has been drawn into the nozzle from a reservoir to form the desired meniscus for the next droplet.

Traditional ink-jet printer heads have been made with electroformed nickel nozzles, produced separately from the integrated circuitry, thus requiring a bonding operation to attach these two components. With increasing printer resolution, it is more difficult to bond the components with a tolerance of less than a few micrometers. For this reason, single-component, or monolithic, fabrication is of interest.

The fabrication sequence for a monolithic ink-jet printer head is shown in Fig. 29.16. A silicon wafer is first prepared and coated with a phosphosilicate-glass (PSG) pattern and a low-stress silicon-nitride coating. The ink reservoir is obtained by isotropically etching the back side of the wafer, followed by PSG removal and enlargement of the reservoir. The required CMOS (complementary metal-oxide semiconductor) controlling circuitry is then produced, and a tantalum heater pad is deposited. The aluminum interconnection between the tantalum pad and the CMOS circuit is formed, and the nozzle is produced through laser ablation. An array of such nozzles can be placed inside an ink-jet printing head, and resolutions of 2400 dpi or higher can be achieved.

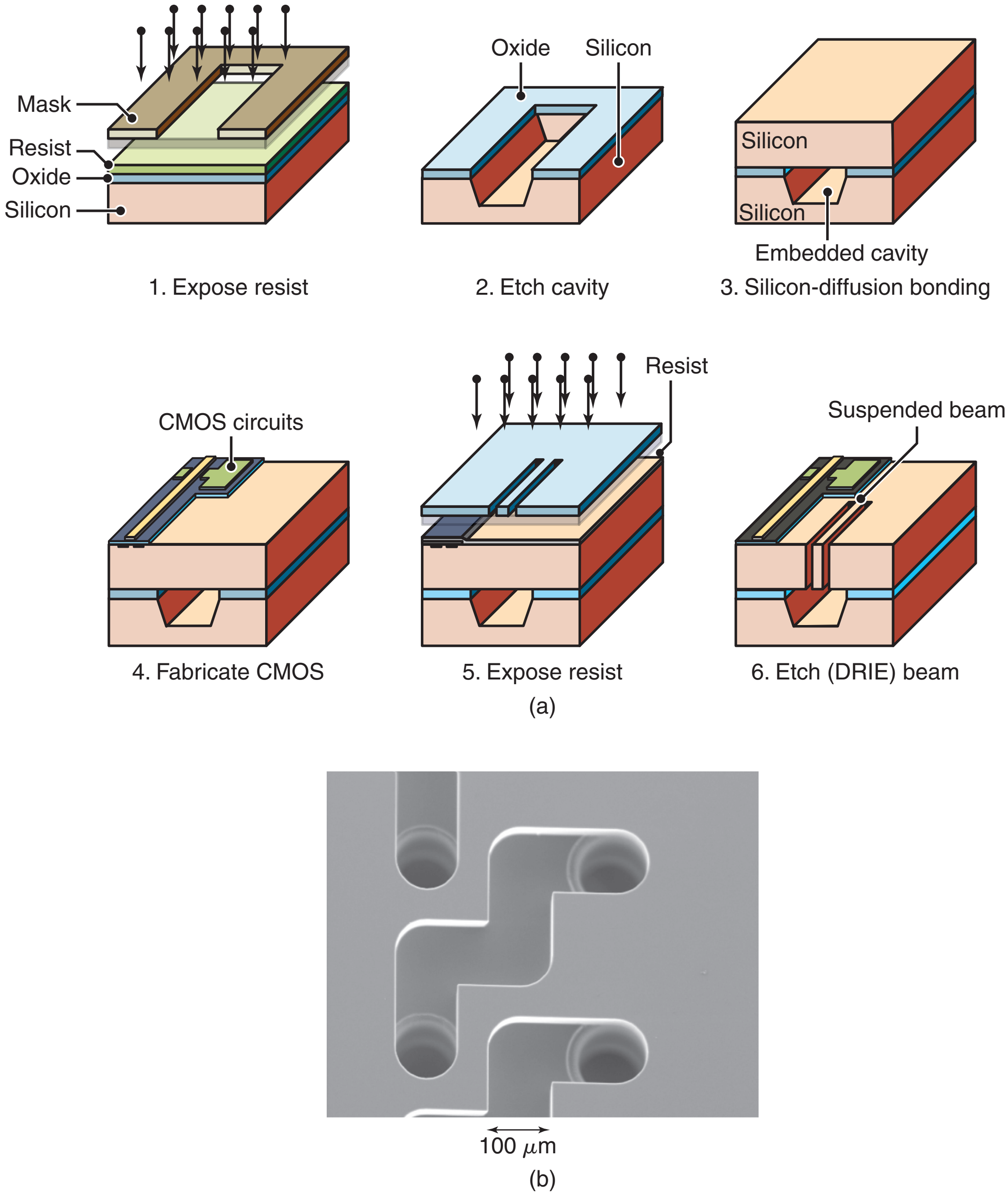


Figure 29.14: (a) Schematic illustration of silicon-diffusion bonding combined with deep reactive-ion etching to produce large, suspended cantilevers. (b) A microfluid-flow device manufactured by DRIE etching two separate wafers and then aligning and silicon-fusion bonding them together. Afterward, a Pyrex layer (not shown) is anodically bonded over the top to provide a window to observe fluid flow. *Source:* (a) After N. Maluf. (b) Courtesy of K.R. Williams, Agilent Technologies.

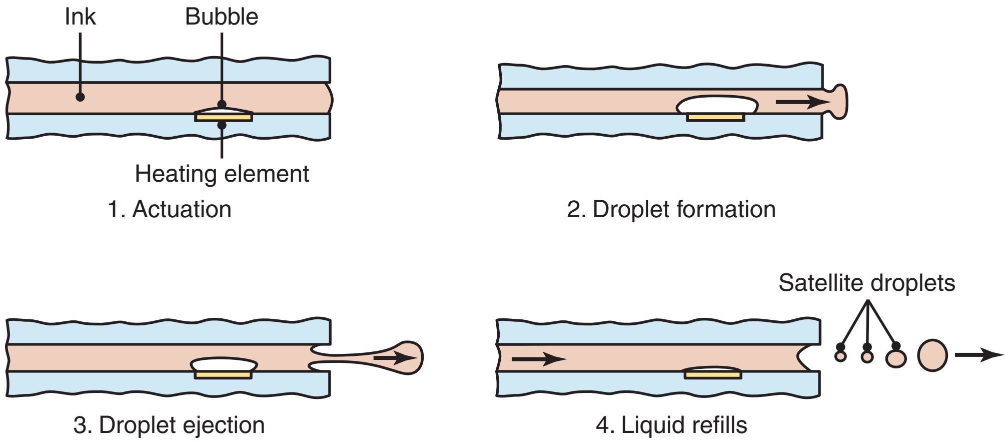


Figure 29.15: Sequence of operation of a thermal ink-jet printer. 1. Resistive heating element is turned on, rapidly vaporizing ink and forming a bubble. 2. Within 5 μs, the bubble has expanded and displaced liquid ink from the nozzle. 3. Surface tension breaks the ink stream into a bubble, which is discharged at high velocity. The heating element is turned off at this time, so that the bubble collapses as heat is transferred to the surrounding ink. 4. Within 24 μs an ink droplet (and undesirable satellite droplets) is ejected, and surface tension of the ink draws in more liquid from the reservoir. *Source:* After F.-G. Tseng.

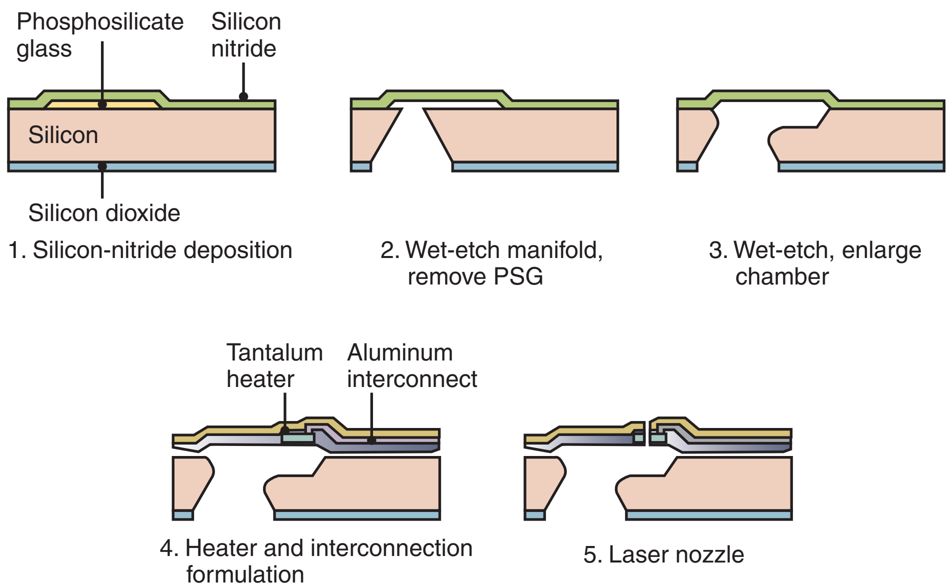


Figure 29.16: The manufacturing sequence for producing thermal ink-jet printer heads. *Source:* After F.-G. Tseng

29.3 Electroforming-based Processes

29.3.1 LIGA

This is a German acronym for the combined process of X-ray lithography, electrodeposition, and molding (in German, X-ray *li*thographie, *g*alvanoformung, und *a*bformung). A schematic illustration of this process is given in Fig. 29.17.

The LIGA process involves the following steps:

1. A very thick (up to hundreds of micrometers) resist layer of polymethylmethacrylate (PMMA) is deposited onto a primary substrate.

2. The PMMA is exposed to columnated X-rays and is developed.

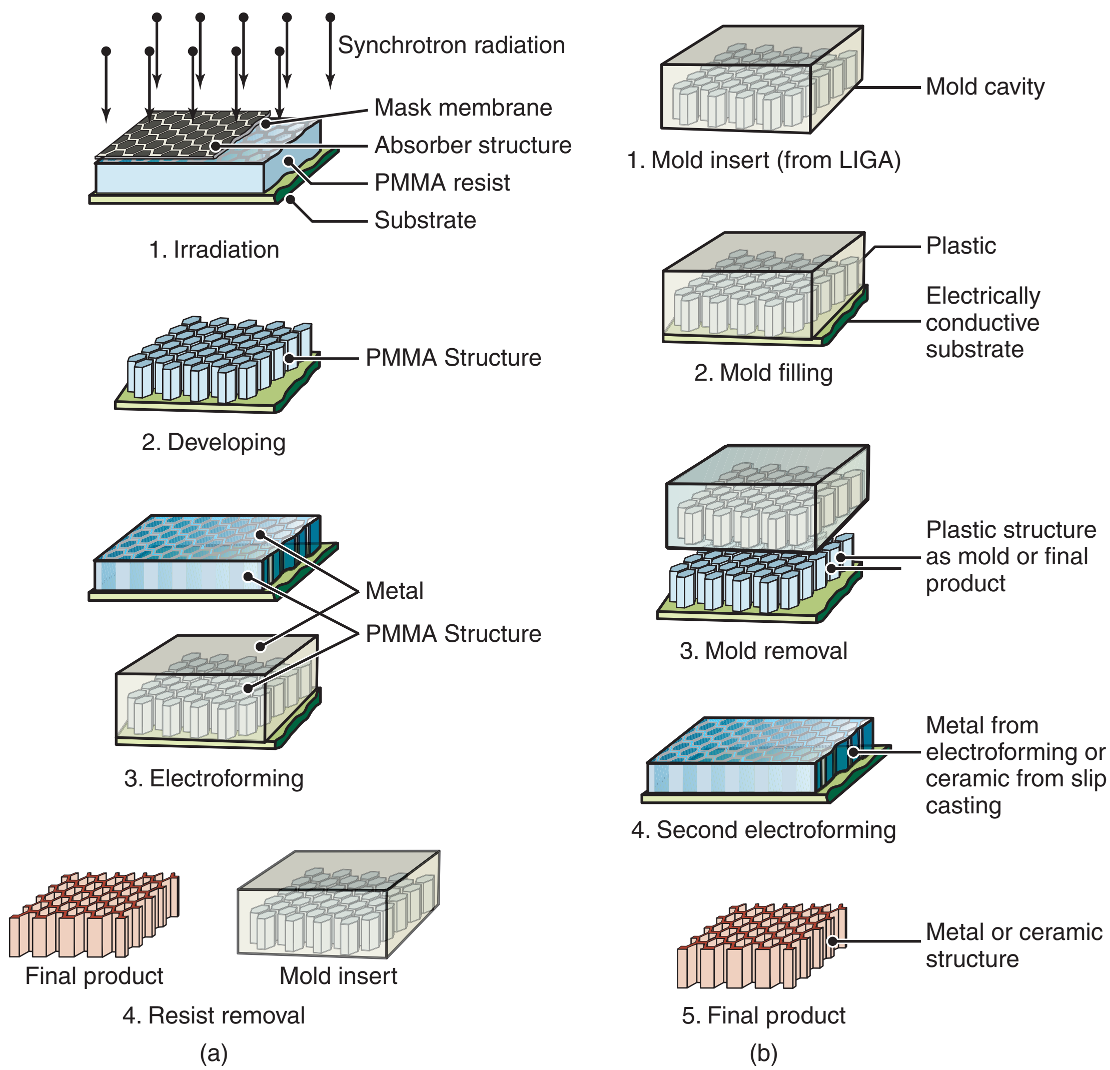


Figure 29.17: The LIGA (lithography, electrodeposition and molding) technique. (a) Primary production of a metal final product or mold insert. (b) Use of the primary part for secondary operations or replication. *Source:* Courtesy of IMM Institut für Mikrotechnik, Mainz, Germany.

3. Metal is electrodeposited onto the primary substrate.
4. The PMMA is removed or stripped, resulting in a freestanding metal structure.
5. Plastic is injection molding into the metal structure.

Depending on the application, the final product from a LIGA process may consist of one of the following:

- A freestanding metal structure, resulting from the electrodeposition process.
- A plastic injection-molded structure.
- An investment-cast metal part, using the injection-molded structure as a blank.
- A slip-cast ceramic part, produced with the injection-molded parts as the molds.

The substrate used in LIGA is a conductor or a conductor-coated insulator. Examples of primary substrate materials include austenitic steel plate, silicon wafers with a titanium layer, and copper plated with gold, titanium, or nickel. Metal-plated ceramic and glass also have been used. The surface may be roughened by grit blasting to encourage good adhesion of the resist material.

Resist materials must have high X-ray sensitivity, dry- and wet-etching resistance when unexposed, and thermal stability. The most common resist material is polymethylmethacrylate, which has a very high molecular weight (more than 10^6 per mole; Section 7.2). The X-rays break the chemical bonds, leading to the production of free radicals and to a significantly reduced molecular weight in the exposed region. Organic solvents then preferentially dissolve the exposed PMMA in a wet-etching process. After development, the remaining three-dimensional structure is rinsed and dried, or it is spun and blasted with dry nitrogen.

Two newer forms of LIGA are **UV-LIGA** and **Silicon-LIGA**. In *UV-LIGA*, special photoresists are used, instead of PMMA, and they are exposed through ultraviolet lithography (Section 28.7). *Silicon-LIGA* uses deep reactive-ion-etched silicon (Section 28.8.2) as a preform for further operations. These processes, like the traditional X-ray-based LIGA, are used to replicate MEMS devices, but, unlike LIGA, they do not require the expensive columnated X-ray source for developing their patterns.

The electrodeposition of metal usually involves the electroplating of nickel (Section 34.9). The nickel is deposited onto exposed areas of the substrate; it fills the PMMA structure and can even coat the resist (Fig. 29.17a). Nickel is the preferred material because of the relative ease in electroplating with well-controlled deposition rates. Electroless plating of nickel also is possible, and the nickel can be deposited directly onto electrically insulating substrates. However, because nickel displays high wear rates in MEMS, significant research is being directed towards the use of other materials or coatings.

After the metal structure has been deposited, precision grinding removes either the substrate material or a layer of the deposited nickel. The process is referred to as *planarization* (Section 28.10). The need for planarization is obvious when it is recognized that three-dimensional MEMS devices require micrometer tolerances on layers many hundreds of micrometers thick. Planarization is difficult to achieve, because conventional lapping leads to preferential removal of the soft PMMA and smearing of the metal. Planarization usually is accomplished with a diamond-lapping procedure (Section 26.7) referred to as *nanogrinding*. Here, a diamond-slurry-loaded, soft-metal plate is used to remove material in order to maintain flatness within 1 μm over a 75-mm diameter substrate.

If cross-linked, the PMMA resist is then exposed to synchrotron X-ray radiation and removed by exposure to an oxygen plasma or through solvent extraction. The result is a metal structure, which may be processed further. Examples of freestanding metal structures produced through the electrodeposition of nickel are shown in Fig. 29.18.

The processing steps used to make freestanding metal structures are time-consuming and expensive. The main advantage of LIGA is that these structures serve as molds for the rapid replication of submicron features through molding operations. The processes that can be used for producing micromolds are shown and compared in Table 29.1, where it can be seen that LIGA provides some clear advantages. Reaction injection molding, injection molding, and compression molding (described in Chapter 19) also have been used to make the micromolds.

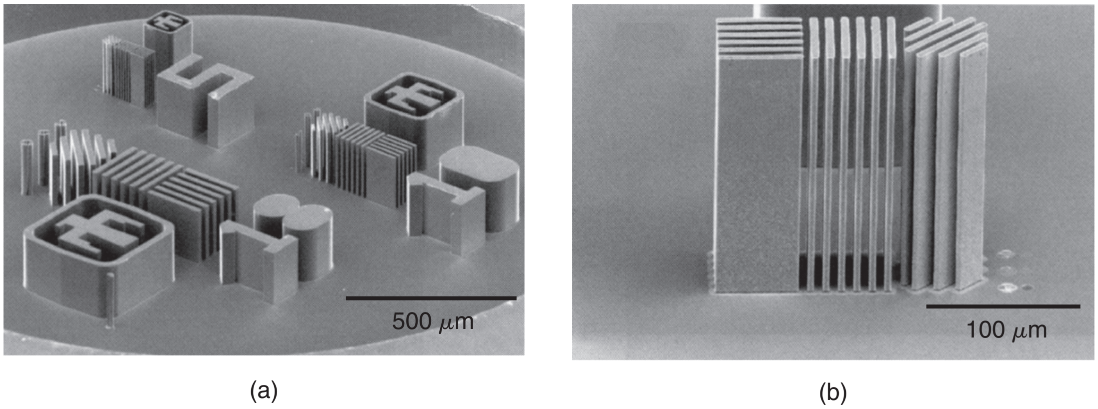


Figure 29.18: (a) Electroformed nickel structures, (b) Detail of nickel lines and spaces. *Source:* After T. Christenson, Sandia National Laboratories.

Table 29.1: Comparison of Micromold Manufacturing Techniques.

	Production technique		
Characteristic	**LIGA**	**Laser machining**	**EDM**
Aspect ratio	10–50	10	up to 100
Surface roughness	$<$ 50 nm	100 nm	0.3-1 μm
Accuracy	$<$ 1 μm	1-3 μm	1-5 μm
Mask required	Yes	No	No
Maximum height	1-500 μm	200-500 μm	μm to mm

Source: After L. Weber, W. Ehrfeld, H. Freimuth, M. Lacher, M. Lehr, P. Pech, and K.R. Williams.

Example 29.3 Production of Rare-earth Magnets

A number of scaling issues in electromagnetic devices indicate that there is an advantage in using rare-earth magnets from the samarium cobalt (SmCo) and neodymium iron boron (NdFeB) families. These materials are available in powder form and are of interest, because they can produce magnets that are an order of magnitude more powerful than conventional ones (Table 29.2). Thus, these materials can be used when effective miniature electromagnetic transducers are to be produced, as are commonly encountered with in-ear headphones or ear buds.

The processing steps involved in manufacturing these magnets are shown in Fig. 29.19. The PMMA mold is produced by exposure to X-ray radiation and solvent extraction. The rare-earth powders are mixed with a binder of epoxy and applied to the mold through a combination of calendering (see Fig. 19.20) and pressing. After curing in a press at a pressure around 70 MPa, the substrate is planarized. The substrate is then subjected to a magnetizing field, of at least 35 kilo-oersteds (kOe), in the desired orientation. Once the material has been magnetized, the PMMA substrate is dissolved, leaving behind the rare-earth magnets, as shown in Fig. 29.20.

29.3.2 Multilayer X-ray Lithography

The LIGA technique is very powerful for producing MEMS devices with large aspect ratios and reproducible shapes. It is, however, often useful to obtain a multilayer stepped structure that cannot be made directly through LIGA. For nonoverhanging part geometries, direct plating can be applied. In a technique

Table 29.2: Comparison of Properties of Permanent-magnet Materials.

Material	Energy product (Gauss-Oersted $\times 10^{-6}$)
Carbon steel	0.20
36% Cobalt steel	0.65
Alnico I	1.4
Vicalloy I	1.0
Platinum–cobalt	6.5
$Nd_2Fe_{14}B$, fully dense	40
$Nd_2Fe_{14}B$, bonded	9

Source: Courtesy of T. Christenson, Sandia National Laboratories.

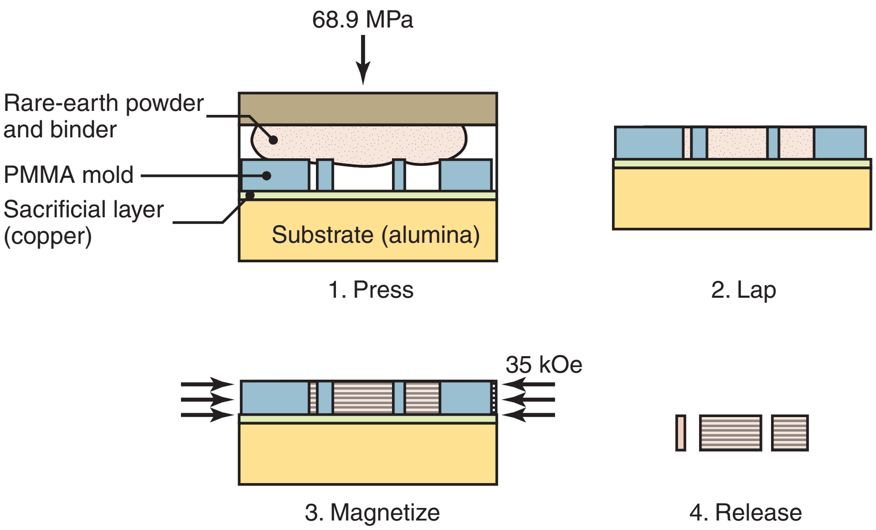


Figure 29.19: Fabrication process used to produce rare-earth magnets for microsensors. *Source:* Courtesy of T. Christenson, Sandia National Laboratories.

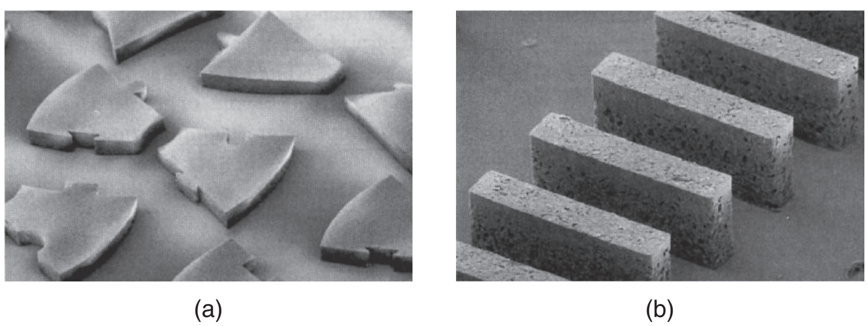


Figure 29.20: SEM images of $Nd_2Fe_{14}B$ permanent magnets. The powder particle size ranges from 1 to 5 μm, and the binder is a methylene chloride-resistant epoxy. Mild distortion is present in the images due to magnetic perturbation of the imaging electrons. Maximum energy products of 9 MGOe have been obtained with this process. *Source:* Courtesy of T. Christenson, Sandia National Laboratories.

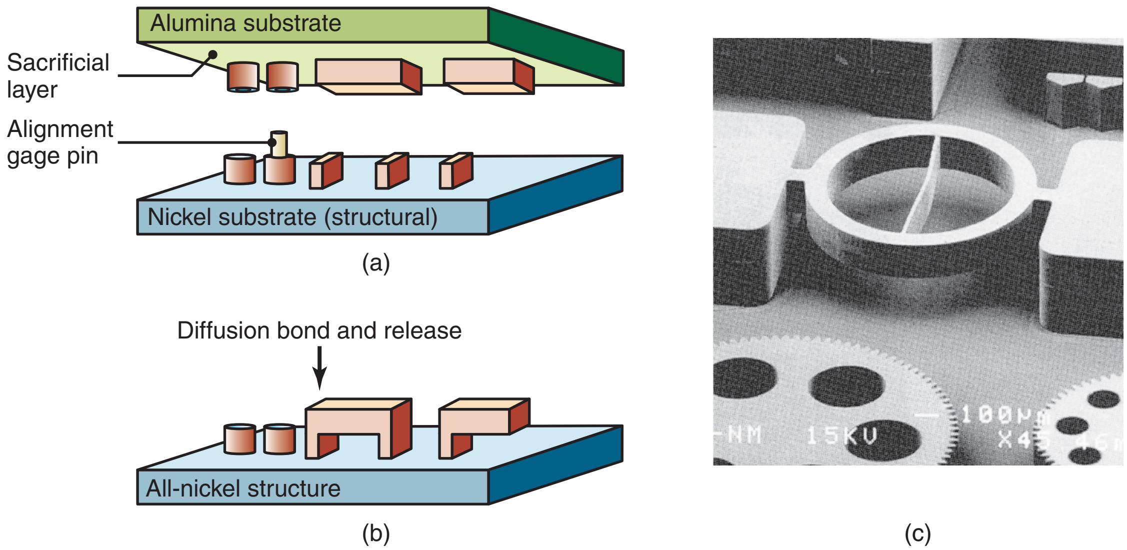


Figure 29.21: Multilevel MEMS fabrication through wafer-scale diffusion bonding. (a) Two wafers are aligned and assembled. (b) Resultant structure after diffusion bonding and removal of alumina substrate. (c) A suspended ring structure for measurement of tensile strain, formed by two-layer wafer-scale diffusion bonding. *Source:* (c) Courtesy of T. Christenson, Sandia National Laboratories.

whereby a layer of electrodeposited metal with surrounding PMMA is produced, as previously described. A second layer of PMMA resist is then bonded to this structure and X-ray exposed using an aligned X-ray mask.

Often, it is useful to have overhanging geometries within complex MEMS devices. A batch diffusion-bonding and release procedure has been developed for this purpose, as is schematically illustrated in Fig. 29.21a. This process involves the preparation of two PMMA patterned and electroformed layers, with the PMMA subsequently removed. The wafers are then aligned face to face with guide pins that press-fit into complementary structures on the opposite surface. Finally, the substrates are joined in a hot press, and a sacrificial layer on one substrate is etched away, leaving behind one layer bonded to the other. An example of such a structure is shown in Fig. 29.21b.

29.3.3 HEXSIL

This process, illustrated in Fig. 29.22, combines *hex*agonal honeycomb structures, *sil*icon micromachining, and thin-film deposition to produce high-aspect-ratio, freestanding structures. HEXSIL can produce tall structures with a shape definition that rivals that of structures produced by LIGA.

In HEXSIL, a deep trench is first produced in single-crystal silicon by dry etching, followed by shallow wet etching to make the trench walls smoother. The depth of the trench matches the desired structure height and is limited practically to around 100 μm. An oxide layer is then grown or deposited onto the silicon, followed by an undoped-polycrystalline silicon layer, which results in good mold filling and shape definition. A doped-silicon layer then follows, providing a resistive portion of the microdevice. Electroplated or electroless nickel plating is then deposited. Figure 29.22 shows various trench widths to demonstrate the different structures that can be produced in HEXSIL.

Microscale tweezers can also be produced through the HEXSIL process. A thermally activated bar activates the tweezers, which have been used for microassembly and microsurgery applications.

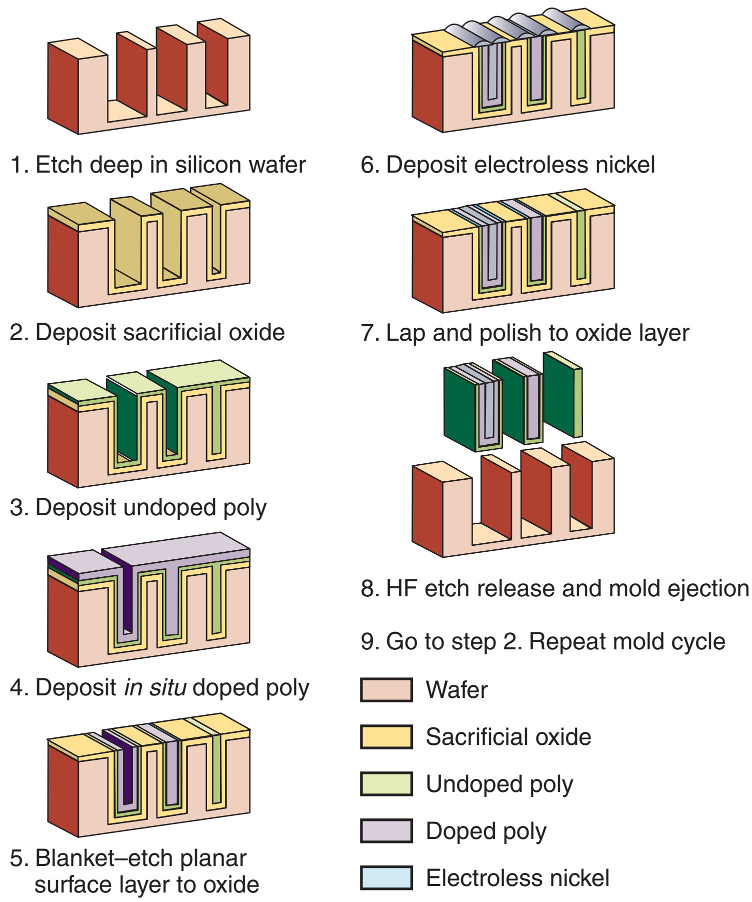


Figure 29.22: Illustration of the hexagonal honeycomb structure, silicon micromachining, and thin-film deposition (HEXSIL).

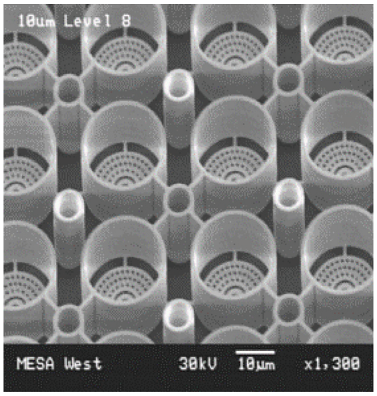
29.3.4 MolTun

This process (short for *mol*ding of *tun*gsten) was developed in order to utilize the higher mass of tungsten in micromechanical devices and systems. In MolTun, a sacrificial oxide is patterned through lithography and then etched, but instead of electroforming, a layer of tungsten is deposited through chemical vapor deposition. Excess tungsten is then removed by chemical mechanical polishing, which also ensures good control over layer thickness. Multiple layers of tungsten can be deposited to develop intricate geometries (Fig. 29.23).

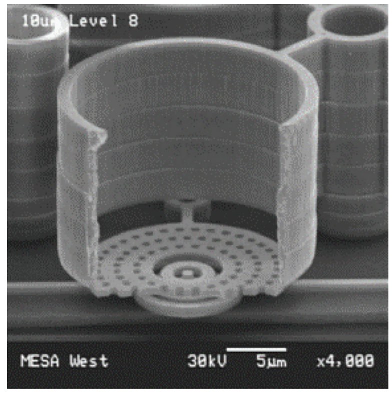
MolTun has been used for micro mass-analysis systems and a large number of micro-scale latching relays, which take advantage of tungsten's higher strength compared to other typical MEMS materials. The depth of MolTun structures can be significantly larger than those produced through silicon micromachining; the mass analysis array in Fig. 29.23, for example, has a total thickness of around 25 μm.

29.4 Solid Free-form Fabrication of Devices

Solid free-form fabrication is another term for *rapid prototyping*, as described in Chapter 20. This method is unique in that complex three-dimensional structures are produced through additive manufacturing, as



(a)



(b)

Figure 29.23: (a) An array of micro-mass analysis systems consisting of cylindrical ion traps, constructed of 14 layers of molded tungsten, including eight layers for the ring electrode. (b) Detail view of the tungsten structures. *Source:* Courtesy of Sandia National Laboratories.

opposed to material removal. Many of the advances in rapid prototyping also are applicable to MEMS manufacture for processes with sufficiently high resolution. *Stereolithography* (Section 20.4) involves curing a liquid thermosetting polymer, using a photoinitiator and a highly focused light source. Conventional stereolithography uses layers between 75 and 500 μm in thickness, with a laser dot focused to a diameter of 0.05 to 0.25 mm.

Microstereolithography. *Microstereolithography* uses the same basic approach as stereolithography; however, there are some important differences between the two processes, including the following:

- The laser is more highly focused, to a diameter as small as 1 μm, as compared with 10 to over 100 μm in stereolithography.
- Layer thicknesses are around 10 μm, which is an order of magnitude smaller than in stereolithography.
- The photopolymers used must have much lower viscosities, to ensure the formation of uniform layers.
- Support structures are not required in microstereolithography, since the smaller structures can be supported by the fluid.
- Parts with significant metal or ceramic content can be produced, by suspending nanoparticles in the liquid photopolymer.

The microstereolithography technique has several cost advantages, but the MEMS devices made by this method are difficult to integrate with the controlling circuitry.

Electrochemical Fabrication. The solid free-form fabrication of MEMS devices using instant masking is known as *electrochemical fabrication* (EFAB). Instant masking is a technique for producing MEMS devices (Fig. 29.24). A mask of elastomeric material is first produced through conventional photolithography techniques, described in Section 28.7. The mask is pressed against the substrate in an electrodeposition bath, so that the elastomer conforms to the substrate and excludes the plating solution in contact areas. Electrodeposition takes place in areas that are not masked, eventually producing a mirror image of the mask. By using a sacrificial filler, made of a second material, instant-masking technology can produce complex three-dimensional shapes complete with overhangs, arches, and other features.

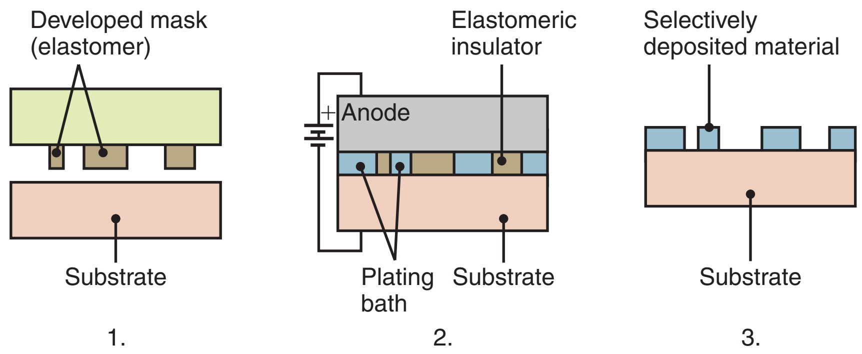


Figure 29.24: The instant-masking process: 1. Bare substrate; 2. During deposition, with the substrate and instant mask in contact; 3. The resulting pattern deposited. *Source:* Courtesy of A. Cohen, MEMGen Corporation.

Case Study 29.2 Photonic Integrated Circuits

There is much debate about whether conventional electronic circuitry has reached the limit of miniaturization, also known as "Moore's law." Traditional scaling in nano-electronics assumes that all relevant dimensions of a circuit can be reduced by 0.7× from generation to generation, while keeping the power consumption of the chip identical by area. However, to produce ever smaller features in a nano-electronics, such as transistors or wires connecting the transistors, requires that even more stringent manufacturing tolerances be achieved. Some transistor dimensions have reached their physical limit: scaling the insulator in the gate can only be done down to about 15–20 A, without incurring substantial leakage current due to electron tunneling through the insulator. In addition, the current that each individual transistor can provide scales with the dimension of the transistor, and the resistance that wires exhibit scales with the square root of the cross-section dimensions.

The nano-electronics industry has dealt with the transistor-based challenges with materials innovation, such as replacing the SiO_2 gate insulator with improved dielectrics, and architectural innovation. However, the gate-switching energy of transistors has not been reduced at the same rate as their size; the result is increased power consumption in integrated circuits. The metal wire resistance problem has offered fewer options for radical innovation and hence contributes substantially to power losses.

These effects have serious consequences, in that the power consumption per unit area of chip in modern microprocessors can reach the energy density of a clothing iron. Obviously, it is very difficult to dissipate the heat produced and the power lost to heat contributes to the power budget for large server farms used to power the Internet.

A potential solution is to use photonics-based circuitry, such as the advanced device shown in Fig. 29.25. Light does not have the same scaling issues that exist with electric current flow. Light can be modulated at very high frequencies and be transmitted instantaneously with little to no power loss. Also, since light can carry multiple signals simultaneously, it has a much higher bandwith.

The manufacture of photonic integrated circuits is very similar to conventional electronics, in that the process involves a cycle as shown in Fig. 28.2. The main departure from semiconductor electronics involves using components and features that modulate light instead of electric current.

The essential elements of a photonic integrated circuit are:

1. **Light sources.** Usually, these are light-emitting diodes (LEDs) or lasers, although a number of other light-emitting devices are under development for various applications. For long-distance optical communications, Indium Phosphate (InP)-based lasers are part of InP-based Photonic Integrated

Circuits (PICs), whereas data-center communication uses mainly silicon-based PICs with external indium-gallium-arsenide (InGaAs) lasers.

2. **Transmission media.** Light can be transmitted through any transparent media, including air which is the medium of choice for sensor applications. Glass fiber can be used to transmit data along a desired path; modern fiber-optic cables allow transmission distances of over 100 km without amplification. Glass fiber can be miniaturized on a wafer where it becomes a wave guide.

3. **Detectors.** Based on photodiodes, detectors recognize the light signals and then transmit an associated electrical current. Detectors in data-center applications can be made of Ge and integrated into a nano-electronics flow.

4. **Modulation.** Information is encoded in a light signal by modulation. A simple form of modulation is to turn a light on or off to transmit a message through the Morse code. A similar approach is used to modulate light signals going through a fiber or wave guide, although more advanced approaches are also being used.

5. **Amplifiers.** When needed, optical amplifiers will increase the light signal in an optical fiber. Being active elements, like lasers, amplifiers are more difficult to integrate in a Si-based semiconductor process.

Improving Internet speed and performance is ultimately a problem of increasing bandwith, or the rate at which information can be transferred, and reducing the power per bit transferred. Bandwidths up to 2 terabytes/s are achievable with conventional electronics. However, the capabilities of photonics circuits are dramatically superior: more than 23 terabytes/s can be achieved in a single optical fiber, while reducing size and power consumption per gigabyte by around 33%. Photonic circuits are intended to drive communications traffic at 400 gigabytes/s, a significant improvement over conventional electronic circuits.

Photonics is able to achieve much higher transmission rates because of Wavelength Division Multiplexing (WDM), allowing multiple channels of light, each with a different frequency, to be transmitted over a single optical fiber. This approach is shown in Fig. 29.26, where four light sources with different frequencies are combined, transmitted over a fiber-optic cable that can extend up to thousands of km (or just the distance between blades or racks in a server farm), and then separated back into its source signals. The light is then detected and the resulting current is amplified as needed.

Photonic circuits are already widely used for long-haul data transmission. Currently, data centers exploit fiber technology and photonic transceivers. The use of PICs and photonic switches have the potential to revolutionize data-center architecture in the next few years, and provide a high-volume application for PICs, further reducing manufacturing costs; it is often said that “photonics will power the Internet.” It is expected that photonics will find increased applications in other domains, such as the optical equivalent of radar and the sensors that will power the Internet of Things (see Section 39.8.1).

Source: Courtesy Michael Liehr, Integrated Photonics Institute for Manufacturing Innovation.

29.5 Mesoscale Manufacturing

Conventional manufacturing processes, as those described in Chapters 5 through 12, typically produce parts that can be described as visible to the naked eye. Such parts are generally referred to as *macroscale*, the word macro being derived from *makros* in Greek, meaning long. The processing of such parts is known as **macromanufacturing**, and is the most developed and best understood size range from a design and manufacturing perspective.

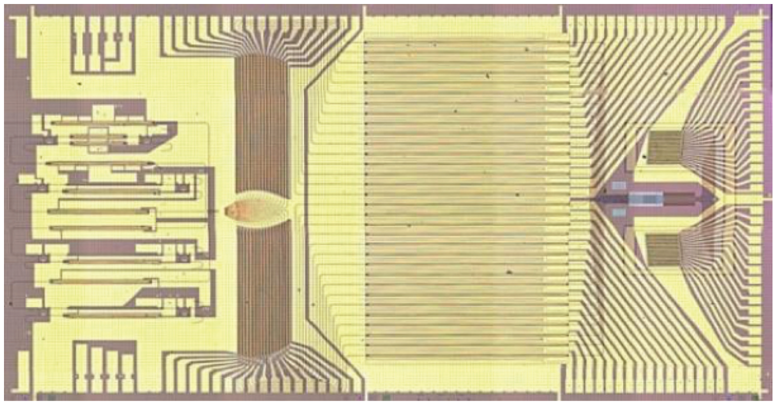


Figure 29.25: Two-dimensional scanning chip containing four tunable lasers, 32 phase shifters, 32 optical amplifiers, and 32 photodetectors. It uses heterogeneous integration to combine lasers, amplifiers, and phase shifters in a single chip that emits a beam that can be scanned in two dimensions. *Source:* J.C. Hulme, J.K. Doylend, M.J.R. Heck, J.D. Peters, M.L. Davenport, J.T. Bovington, L.A. Coldren, and J.E. Bowers, *Optics Express*, Vol. 23, No. 5, pp. 5861–5874, 2015. Used with permission of J.E. Bowers.

In contrast, examples of **mesomanufacturing** are components for miniature devices, such as hearing aids, medical devices such as stents and valves, mechanical watches, and extremely small motors and bearings. Note that mesomanufacturing overlaps both macro- and micro-manufacturing in Fig. V.1.

Two general approaches have been used for mesomanufacturing: scaling macromanufacturing processes *down*, and scaling micromanufacturing processes *up*. Examples of the former are a lathe with a 1.5-W motor that measures 32 mm $\times$ 25 mm $\times$ 30.5 mm that weighs 100 g. Such a lathe can machine brass to a diameter as small as 60 μm and with a surface roughness of 1.5 μm. Similarly miniaturized versions are milling machines, mechanical presses, and various other machine tools.

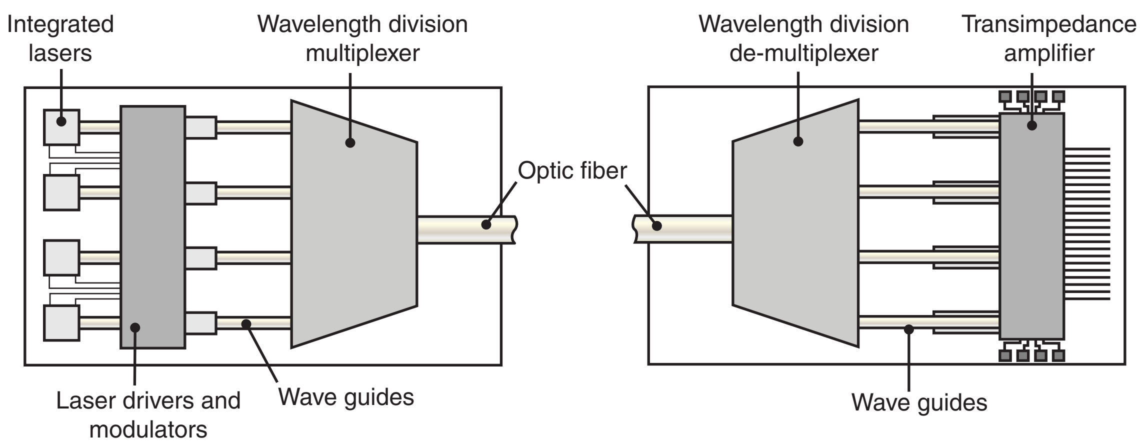


Figure 29.26: Schematic illustration of a photonic integrated circuit for data transmission. Numerous lasers, each producing a different wavelength of light, have their output signal modulated and combined into one fiber-optic cable by the wavelength division multiplexer (WDM Mux). The cable can extend thousands of kilometers to the receiving circuit, where the signal is separated into its components by a demultiplexer (WDM Demux), and then detected and amplified.

Scaling up of micromanufacturing processes is similarly performed; LIGA can produce micro-scale devices, but the largest parts that can be made are meso-scale. For this reason, mesomanufacturing processes are often indistinguishable from their micromanufacturing implementations.

29.6 Nanoscale Manufacturing

In *nanomanufacturing*, parts are produced at nanometer length scales; the term usually refers to manufacturing strategies below the micrometer scale, or between 10^{-6} and 10^{-9} m in length. Many of the features in integrated circuits are at this length scale, but very little else has significant relevance to manufacturing, Molecularly engineered medicines and other forms of biomanufacturing are the only commercial applications at present. However, it has been recognized that many physical and biological processes act at this length scale; consequently, the approach holds much promise for future innovations.

Nanomanufacturing takes two basic approaches: top down and bottom up. **Top-down** approaches use *large building blocks* (such as a silicon wafer; see Fig. 28.2) and various manufacturing processes (such as lithography, and wet and plasma etching) to construct ever smaller features and products (microprocessors, sensors, and probes). On the other hand, **bottom-up** approaches use *small building blocks* (such as atoms, molecules, or clusters of atoms and molecules) to build up a structure. In theory, bottom-up approaches are similar to the additive manufacturing technologies described in Section 20.2. When placed in the context of nanomanufacturing, however, bottom-up approaches suggest the manipulation and construction of products are on an atomic or molecular scale.

Bottom-up approaches are widely used in nature (for example, building cells is a fundamentally bottom-up approach), whereas conventional manufacturing has, for the most part, consisted of top-down approaches. In fact, there are presently no nanomanufactured products (excluding medicines and drugs "manufactured" by bacteria) that have demonstrated commercial viability.

Bottom-up approaches in various research applications can use **atomic-force microscopy** (AFM) for the manipulation of materials on the nanoscale. Figure 29.27 is an illustration of an atomic-force microscope. A probe is mounted into the microscope, and a laser is reflected from a mirror on the back side of the probe so that it reflects onto a set of photosensors. Any vertical or torsional deflection of the cantilever is registered as a change in voltage on the photosensors. Atomic-force microscopes can have true atomic resolution of $< 1 \times 10^{-10}$ m.

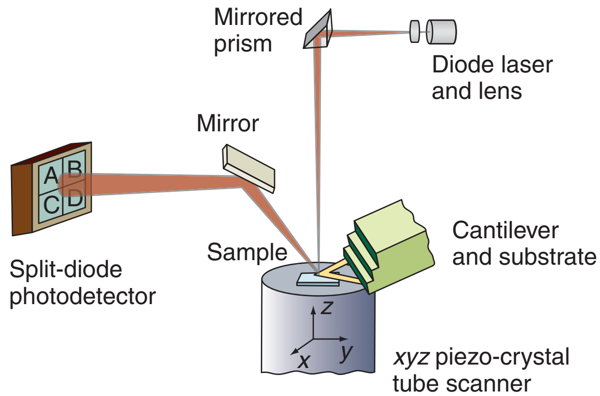


Figure 29.27: Schematic illustration of an atomic-force microscope. A probe is mounted on a cylinder containing piezoelectric material; this arrangement allows translation of the probe in three dimensions. A laser reflected from a mirror on the back of the probe onto a set of photosensors allows measurement of the probe's location and monitoring of interactions with a sample surface.

Atomic-force microscopes are widely used to measure the surface profile of very smooth surfaces (Section 33.3). Several approaches have been developed to allow nanoscale manufacturing processes to be performed on these microscopes. Some top-down approaches are:

- Photolithography, electron-beam lithography, and **nanoimprint (soft) lithography**, are capable of top-down manufacture of structures, with resolution under 100 nm, as discussed in Section 28.7.

- **Nanolithography.** The probes used in atomic-force microscopy vary greatly in size, materials, and capabilities. The diamond-tipped stainless-steel cantilever shown in Fig. 29.27 has a tip radius of around 10 nm. By contacting and plowing across a surface, it can produce grooves up to a few μm thick. The spacing between lines depends on the groove depth needed.

- **Dip Pen Nanolithography.** This approach is used on an atomic-force microscope to transfer chemicals onto substrates. The process can produce lines as narrow as 10 nm. Dip pen nanolithography can be used with many parallel pens, typically made of silicon nitride and containing as many as 55,000 pens in a 1 cm^2 area. In a top-down approach, dip pen nanolithography is used to produce a mask suitable for lithography.

Bottom-up approaches include the following:

- Dip pen nanolithography also can be a bottom-up approach, wherein the ink contains the material used to build the structure.

- *Microcontact printing* uses soft-lithography approaches, to deposit material on surfaces from which nanoscale structures can be produced.

- *Scanning tunneling microscopy* can be used to manipulate an atom on an atomically smooth surface (usually cleaved mica or quartz).

Summary

- MEMS is relatively new and developing rapidly. Although most successful commercial MEMS applications are in the optics, printing, and sensor industries, the possibilities for new device concepts and circuit designs appear to be endless.
- MEMS devices are manufactured through techniques and with materials that, for the most part, have been pioneered in the microelectronics industry. Bulk and surface micromachining are processes that are well developed for single-crystal silicon.
- Specialized processes for MEMS include variations of machining, such as DRIE, SIMPLE, and SCREAM. These processes produce freestanding mechanical structures in silicon.
- Polymer MEMS can be manufactured through LIGA or microstereolithography. LIGA combines X-ray lithography and electroforming to produce three-dimensional structures. Related processes include multilayer X-ray lithography and HEXSIL.
- Nanoscale manufacturing is a relatively new area that has significant potential. Nanoscale manufacturing processes are typically bottom up, whereas conventional manufacturing is top down. Some lithography processes extend to the nanoscale, as does dip pen lithography. Materials such as carbon nanotubes have great potential for nanoscale devices.

Key Terms

Atomic-force microscope
Bulk micromachining
Diffusion bonding
Dip pen nanolithography
EFAB
Electrochemical fabrication
HEXSIL
LIGA
MEMS
Micromachining
Microstereolithography
MolTun
Multilayer X-ray lithography
Photonic
Planarization
Sacrificial layer
SCREAM
Silicon-LIGA
SIMPLE
Stiction
Surface micromachining
UV-LIGA

Bibliography

Adams, T.M., and Layton, R.W., **Introductory MEMS: Fabrication and Applications**, Springer, 2009.
Allen, J.J., **Microelectromechanical System Design**, CRC Press, 2006.
Campbell, S.A., **The Science and Engineering of Microelectronic Fabrication**, 2nd ed., Oxford, 2001.
Elwenspoek, M., and Jansen, H., **Silicon Micromachining**, 2nd ed., Cambridge University Press, 2004.
Elwenspoek, M., and Wiegerink, R., **Mechanical Microsensors**, Springer, 2001.
Fraden, J., **Handbook of Modern Sensors: Physics, Designs, and Applications**, 4th ed., Springer, 2010.
Ghodssi, R., and L.P. (eds.), **MEMS Materials and Processes Handbook**, Springer, 2011.

Harper, C.A. (ed.), **Electronic Packaging and Interconnection Handbook**, 4th ed., McGraw-Hill, 2004.
Harper, C.A. (ed.), **High-Performance Printed Circuit Boards**, McGraw-Hill, 2000.
Hsu, T.-R., **MEMS & Microsystems: Design, Manufacture, and Nanoscale Engineering**, 2nd ed., Wiley, 2008.
Jha, A.R., **MEMS and Nanotechnology-Based Sensors and Devices for Communications, Medical and Aerospace Applications**, CRC Press, 2008.
Kempe, V., **Inertial MEMS: Principles and Practice**, Cambridge, 2011.
Korvink, J., and Oliver, P. (eds.), **MEMS: A Practical Guide to Design, Analysis, and Applications**, William Andrew, 2005.
Leondes, C.T., **MEMS/NEMS Handbook** (5 vols.), Springer, 2007.
Liu, C., **Foundations of MEMS**, 2d ed., PrenticeHall, 2011.
Madou, M.J., **Manufacturing Techniques for Microfabrication and Nanotechnology**, CRC Press, 2009.
Madou, M.J., **Applications of Microfabrication and Nanotechnology**, CRC Press, 2009.
Nishi, Y., and Doering, R. (eds.), **Handbook of Semiconductor Manufacturing Technologies**, 2nd ed., CRC Press, 2007.
Quirk, M., and Serda, J., **Semiconductor Manufacturing Technology**, Prentice Hall, 2000.
Rockett, A., **The Materials Science of Semiconductors**, Springer, 2008.
Schaeffer, R., **Fundamentals of Laser Micromachining**, Taylor & Francis, 2012.
Sze, S.M., **Semiconductor Devices: Physics and Technology**, 3rd ed., Wiley, 2012.
Varadan, V.K., Jiang, X., and Varadan, V., **Microstereo-lithography and other Fabrication Techniques for 3D MEMS**, Wiley, 2001.

Review Questions

29.1. Define MEMS, SIMPLE, SCREAM, and HEXSIL.

29.2. Give three examples of common micro-electromechanical systems.

29.3. Why is silicon often used with MEMS devices?

29.4. Describe bulk and surface micromachining.

29.5. What is the purpose of a spacer layer in surface micromachining?

29.6. What is the main limitation to successful application of MEMS?

29.7. What are common applications for MEMS and MEMS devices?

29.8. What is LIGA? What are its advantages?

29.9. What is a sacrificial layer?

29.10. Explain the differences between stereolithography and microstereolithography. What is MolTun? What are its main advantages?

29.11. What is HEXSIL?

29.12. What do SIMPLE and SCREAM stand for?

Qualitative Problems

29.13. Describe the difference between isotropic etching and anisotropic etching.

29.14. Lithography produces projected shapes, so true three-dimensional shapes are more difficult to produce. What lithography processes are best able to produce three-dimensional shapes, such as lenses? Explain.

29.15. Which process or processes in this chapter allow the fabrication of products from polymers?

29.16. Describe the difference between a microelectronic device, a micromechanical device and MEMS.

29.17. What is the difference between chemically reactive ion etching and dry-plasma etching?

29.18. The MEMS devices discussed in this chapter are applicable to macroscale machine elements, such as spur gears, hinges, and beams. Which of the following machine elements can or cannot be applied to MEMS, and why? (a) ball bearings, (b) bevel gears, (c) worm gears, (d) cams, (e) helical springs, (f) rivets, and (g) bolts.

29.19. Explain how you would produce a spur gear if its thickness was one-tenth of its diameter and its diameter was (a) (b) (c) 1 mm, (d) 10 mm, and (e) 100 mm.

29.20. What is a mask? Of what materials is it composed?

29.21. List the advantages and disadvantages of surface micromachining compared with bulk micromachining.

29.22. What are the main limitations to the LIGA process? Explain.

29.23. Other than HEXSIL, what process can be used to make microtweezers? Explain.

29.24. Is there an advantage to using the MolTun process for other materials? Explain.

Quantitative Problems

29.25. The atomic-force microscope probe shown in Fig. 29.27 has a stainless steel cantilever that is $450 \times 40 \times 2$ μm. Using equations from solid mechanics, estimate the stiffness of the cantilever, and the force required to deflect the end of the cantilever by 1.5 μm.

29.26. Tapping-mode probes for the atomic-force microscope are produced from etched silicon and have typical dimensions of 125 μm in length, 30 μm in width, and 3 μm in thickness. Estimate the stiffness and natural frequency of such probes.

29.27. Using data from Chapter 28, derive the time needed to etch the hinge shown in Fig. 29.7 as a function of the hinge thickness.

29.28. It is desired to produce a 500 μm by 500 μm diaphragm, 25 μm thick, in a silicon wafer 250 μm thick. Given that you will use a wet etching technique with KOH in water with an etch rate of 1 μm/min, calculate the etching time and the dimensions of the mask opening that you would use on a (100) silicon wafer.

29.29. If the Reynolds number for water flow through a pipe is 2000, calculate the water velocity if the pipe diameter is (a) 12 mm; (b) 120 μm. Do you expect flow in MEMS devices to be turbulent or laminar? Explain.

Synthesis, Design, and Projects

29.30. List similarities and differences between IC technologies described in Chapter 28 and miniaturization technologies presented in this chapter.

29.31. Figure I.7 in the General Introduction shows a mirror that is suspended on a torsional beam and can be inclined through electrostatic attraction by applying a voltage on either side of the micromirror at the bottom of the trench. Make a flowchart of the manufacturing operations required to produce this device.

29.32. Referring to Fig. 29.5, design an experiment to find the critical dimensions of an overhanging cantilever that will not stick to the substrate.

29.33. Design an accelerometer by using (a) the SCREAM process and (b) the HEXSIL process.

29.34. Design a micromachine or device that allows the direct measurement of the mechanical properties of a thin film.

29.35. Conduct a literature search and determine the smallest diameter hole that can be produced by (a) drilling, (b) punching, (c) water-jet cutting, (d) laser machining, (e) chemical etching; and (f) EDM.

29.36. Perform a literature search and write a one-page summary of applications in bio-MEMS.

PART VI
Joining Processes and Equipment

Some products, such as paper clips, nails, steel balls for bearings, screws, and bolts, are made of only one component; however, almost all products are assembled from components that have been made as individual parts. Even relatively simple products consist of at least two components, joined by various means. Note, for example: (a) the eraser at the end of an ordinary pencil is attached with a metal sleeve; (b) knives have wooden or plastic handles that are attached to the metal blade with fasteners; and (c) cooking pots and pans have metal, plastic, or wooden handles and knobs, attached to the pot by various methods.

On a much larger scale, observe power tools, washing machines, motorcycles, ships, and airplanes, and how their numerous components are assembled and joined so that they not only can function reliably, but also are economical to produce. As shown in Table I.1 in the General Introduction, a rotary lawn mower has about 300 parts, a typical automobile has 15,000 components, and a Boeing 747-400 aircraft has more than 6 million parts. In contrast, a Boeing 787 Dreamliner has fewer parts because its composite fuselage eliminates a large number of fasteners.

Joining is an all-inclusive term covering processes such as welding, brazing, soldering, adhesive bonding, and mechanical fastening. These processes are an essential and important aspect of manufacturing and **assembly** for one or more of the following reasons:

1. Even a relatively simple product may be impossible to manufacture as a single piece. Consider, for example, the tubular construction shown in Fig. VI.2a. Assume that each of the arms of this product is 5 m long, the tubes are 100 mm in diameter, and their wall thickness is 1 mm. After reviewing all of the manufacturing processes described in the preceding chapters, one would conclude that manufacturing this product in one piece would be impossible or uneconomical.

2. A product such as a cooking pot, with a handle, is *easier and more economical* to manufacture as assembly of individual components.

3. Products such as appliances, automobile engines, and hair dryers must be designed so as to be able to be easily taken apart for maintenance or for replacement of their worn or broken parts.

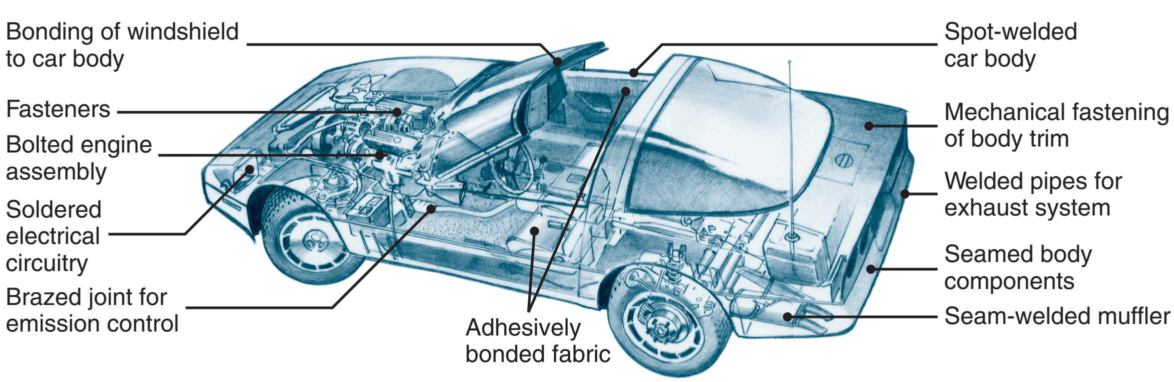


Figure VI.1: Various parts in a typical automobile that are assembled by the processes described in Part VI.

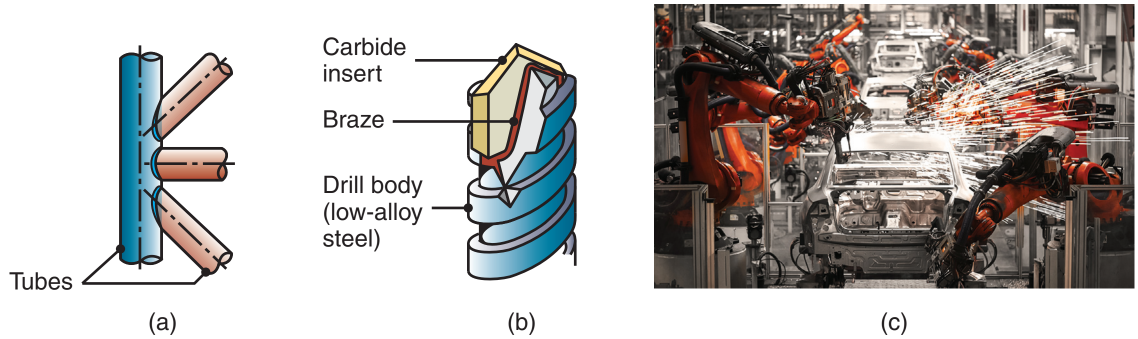


Figure VI.2: Examples of parts utilizing joining processes. (a) A tubular part fabricated by joining individual components. This product cannot be manufactured in one piece by any of the methods described in the previous chapters if it consists of thin-walled, large-diameter, tubular-shaped long arms. (b) A drill bit with a carbide cutting insert brazed to a steel shank—an example of a part in which two materials need to be joined for performance reasons. (c) Spot welding of automobile bodies. *Source:* (c) Shutterstock/Jensen.

4. *Different properties* are often desirable for functional purposes of a product. For example, surfaces subjected to friction, wear, corrosion, or environmental attack generally require characteristics that are significantly different from those of the component's bulk. Examples are: (a) masonry drills with carbide cutting tips, brazed to the shank of a drill (Fig. VI.2b); (b) automotive brake shoes attached to their support with rivets; and (c) grinding wheels bonded to a metal backing (Section 26.2).

5. *Transporting* the product in individual components and assembling them later may be easier and less costly than transporting the completed product. Note, for example, that metal or wood shelving, backyard grills, and large machinery are assembled after the components or subassemblies have been transported to their intended sites.

Although there are different ways of categorizing the wide variety of available joining processes, they basically fall into the following three major categories (Figs. VI.3 and I.7f):

- **Welding**
- **Adhesive bonding**
- **Mechanical fastening**

Table VI.3 lists the general characteristics of various joining processes, and welding processes, in turn, are generally classified into three basic categories:

- Fusion welding
- Solid-state welding
- Brazing and soldering

As will be shown later, some types of welding processes can be classified into both the fusion and the solid-state categories.

Fusion welding is defined as the *melting together and coalescing* of materials by means of *heat*, usually supplied by chemical or electrical means; filler metals may or may not be used. Fusion welding is composed of *consumable-* and *nonconsumable-electrode arc welding* and *high-energy-beam welding* processes. The welded

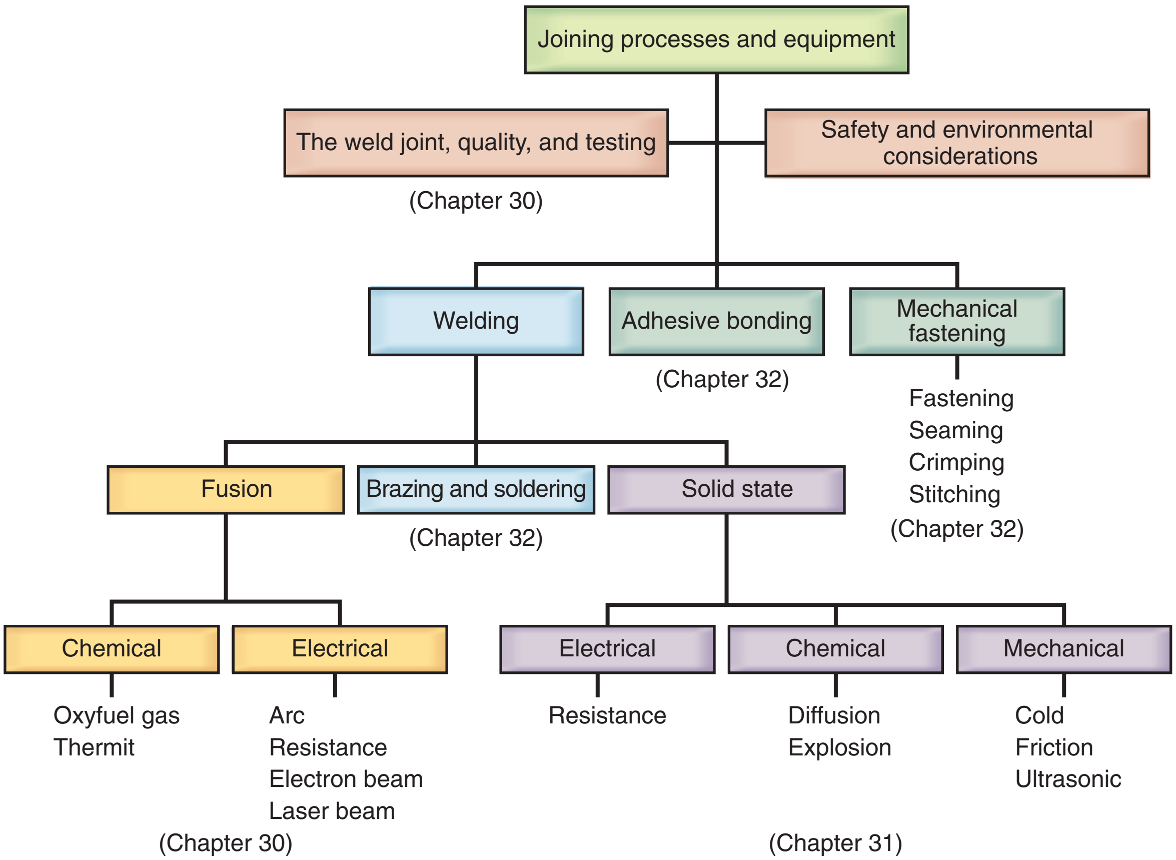


Figure VI.3: Outline of topics described in Part VI.

joint undergoes major metallurgical and physical changes, which, in turn, have a major influence on the properties and performance of the welded assembly. The terminology for some simple welded joints are illustrated in Fig. VI.4.

In **solid-state welding**, joining takes place without fusion; consequently, there is no liquid (molten) phase in the joint. The basic processes in this category are *diffusion bonding* and *cold, ultrasonic, friction, resistance, and explosion welding*. **Brazing** uses filler metals and involves lower temperatures than in welding. **Soldering** uses filler metals (*solders*) and involves even lower temperatures.

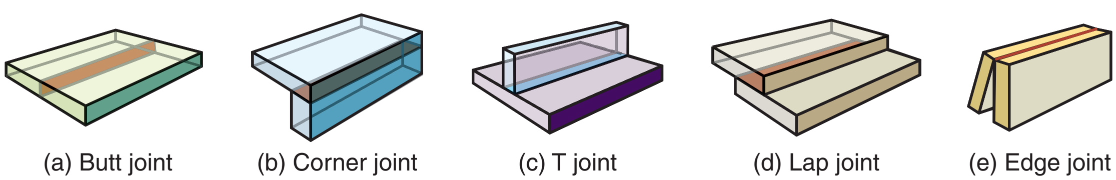


Figure VI.4: Examples of joints that can be made through the various joining processes described in Chapters 30 through 32.

Table VI.3: Comparison of Various Joining Methods.

	Characteristics								
Method	**Strength**	**Design**	**Small parts**	**Large parts**	**Tolerances**	**Reliability**	**Ease of manufacture**	**Ease of inspection**	**Cost**
Arc welding	1	2	3	1	3	1	2	2	2
Resistance welding	1	2	1	1	3	3	3	3	1
Brazing	1	1	1	1	3	1	3	2	3
Bolts and nuts	1	2	3	1	2	1	1	1	3
Rivets	1	2	3	1	1	1	3	1	2
Seaming and crimping	2	2	1	3	3	1	3	1	1
Adhesive bonding	3	1	1	2	3	2	3	3	2

Note: For cost, 1 is the lowest.

Adhesive bonding has unique applications requiring strength, sealing, thermal and electrical insulating, vibration damping, and resistance to corrosion between dissimilar metals. **Mechanical fastening** involves methods of using various fasteners, such as bolts, nuts, and rivets. The **joining of plastics** can be accomplished by adhesive bonding, fusion by various external or internal heat sources, and mechanical fastening.

Chapter 30

Fusion Welding Processes

- This chapter describes fusion-welding processes, in which two pieces are joined together by applying heat, which melts and fuses the interface; the operation is sometimes assisted with a filler metal.
- All fusion-welding processes are described in this chapter, beginning with oxyfuel–gas welding in which acetylene and oxygen provide the energy required for welding.
- Various arc-welding processes are then described, in which electrical energy and consumable or nonconsumable electrodes are used to produce the weld; specific processes reviewed include shielded metal arc welding, flux-cored arc welding, gas tungsten-arc welding, submerged arc welding, and gas metal-arc welding.
- Welding with high-energy beams is then described, in which electron beams or lasers provide highly focused heat sources.
- The chapter concludes with a description of the nature of the weld joint, including weld quality, inspection, and testing procedures, along with weld design practices and process selection.

30.1 Introduction

The welding processes described in this chapter involve the partial melting and fusion between two members to be joined. **Fusion welding** is defined as *melting together and coalescing* materials by means of heat. *Filler metals*, which are added to the weld area during welding, also may be used. Welds made without the use of filler metals are known as *autogenous welds*.

This chapter covers the basic principles of each welding process; the equipment used; the relative advantages, limitations, and capabilities of the process; and the economic considerations affecting process selection (Table 30.1). The chapter continues with a description of the weld zone and the variety of discontinuities and defects that can exist in joints. Weldability of ferrous and nonferrous metals and alloys are then reviewed. The chapter concludes with design guidelines for welding, with several examples of good weld-design practices, and the economics of welding.

30.2 Oxyfuel–Gas Welding

Oxyfuel–gas welding (OFW) is a general term used to describe any welding process that uses a **fuel gas** combined with *oxygen* to produce a flame, the source of the heat required to melt the metals at the joint. The most common gas-welding process uses *acetylene*, known as *oxyacetylene–gas welding* (OAW), typically used for structural metal fabrication and repair work.

Developed in the early 1900s, OAW utilizes the heat generated by the combustion of acetylene gas (C_2H_2) in a mixture with oxygen. The primary combustion process, which occurs in the inner core of the flame (Fig. 30.1), involves the following reaction:

$$C_2H_2 + O_2 \rightarrow 2CO + H_2 + \text{Heat}. \tag{30.1}$$

This reaction dissociates the acetylene into carbon monoxide and hydrogen, and produces about one-third of the total heat generated in the flame. The secondary combustion process is

$$2CO + H_2 + 1.5O_2 \rightarrow 2CO_2 + H_2O + \text{Heat}. \tag{30.2}$$

Table 30.1: General Characteristics of Fusion-welding Processes.

Joining process	Operation	Advantage	Skill level required	Welding position	Current type	Distortion*	Typical cost of equipment ($)
Shielded metal arc	Manual	Portable and flexible	High	All	AC, DC	1 to 2	Low (1500+)
Submerged arc	Automatic	High deposition	Low to medium	Flat and horizontal	AC, DC	1 to	Medium (5000+)
Gas metal arc	Semiautomatic or automatic	Most metals	Low to high	All	DC	2 to 3	Medium (5000+)
Gas tungsten arc	Manual or automatic	Most metals	Low to high	All	AC, DC	2 to	Medium (2000+)
Flux-cored arc	Semiautomatic or automatic	High deposition	Low to high	All	DC	1 to 3	Medium (2000+)
Oxyfuel	Manual	Portable and flexible	High	All	–	2 to 4	Low (500+)
Electron beam, laser beam	Semiautomatic or automatic	Most metals	Medium to high	All	–	3 to 5	High (100,000–1 million)
Thermit	Manual	Steels	Low	Flat and horizontal	–	2 to 4	Low (500+)

* 1 = highest; 5 = lowest

This reaction consists of the further burning of both the hydrogen and the carbon monoxide, producing about two-thirds of the total heat. The temperatures developed in the flame can reach 3300°C. Note from Eq. (30.2) that the reaction also produces water vapor.

Types of Flames. The proportion of acetylene and oxygen in the gas mixture is an important factor in oxyfuel–gas welding. At a ratio of 1:1 (*i.e.,* when there is no excess oxygen), the flame is considered to be **neutral** (Fig. 30.1a). With higher oxygen supply, the flame oxidizes the metal (especially steels), and hence known as an **oxidizing flame** (Fig. 30.1b). Only in the welding of copper and copper-based alloys is an oxidizing flame desirable, because in those cases, a thin protective layer of *slag* (compounds of oxides) forms over the molten metal.

If the oxygen is insufficient for full combustion, the flame is known as a **reducing** or **carburizing flame** (Fig. 30.1c). The temperature of a reducing flame is lower; hence, it is suitable for applications requiring low heat, such as in brazing and soldering (Chapter 32), and flame-hardening (Table 4.1).

Other fuel gases, such as hydrogen and methylacetylene propadiene, also can be used in oxyfuel–gas welding. However, the temperatures developed by these gases are lower than those produced by acetylene. Therefore, they are used for welding metals with low melting points, such as lead, and parts that are small and thin.

Filler Metals. *Filler metals* are used to supply additional metal to the weld zone, and are available as **filler rods** or **wire** (Fig. 30.1d) and may be bare or coated with **flux**. The purpose of the flux is to retard oxidation of the surfaces of the parts being welded by generating a gaseous shield around the weld zone. The flux also helps to dissolve and remove oxides and other substances from the weld zone, thus making the joint stronger. The *slag* developed (compounds of oxides, fluxes, and electrode-coating materials) protects the molten puddle of metal against oxidation as the weld cools.

Welding Practice and Equipment. Oxyfuel–gas welding can be used with most ferrous and nonferrous metals and for almost any workpiece thickness; however, the relatively low heat input limits this process

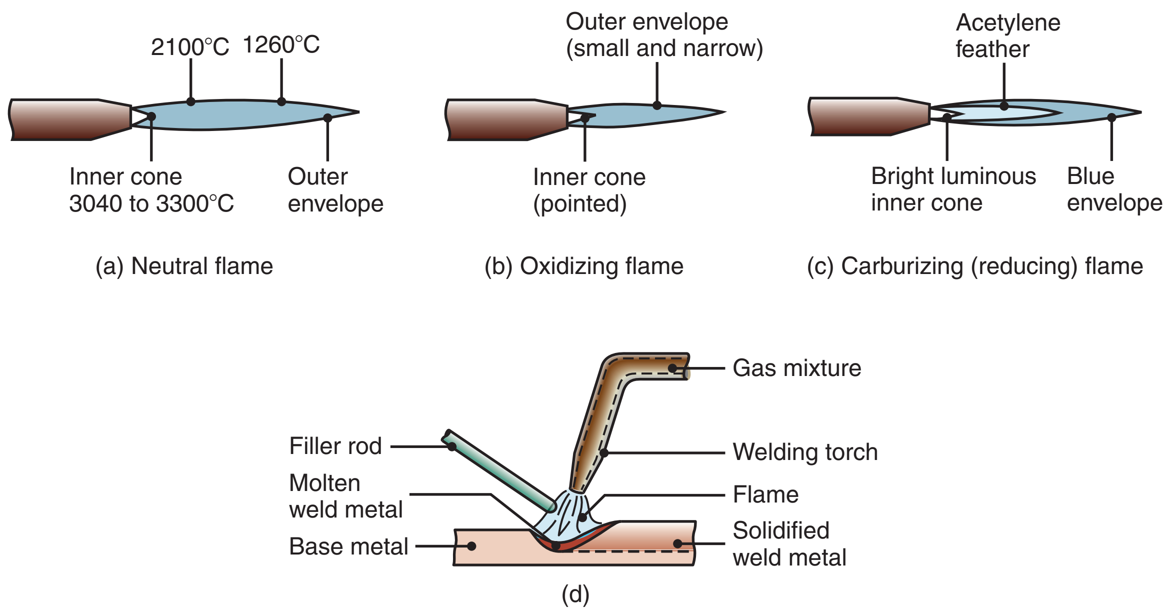


Figure 30.1: Three basic types of oxyacetylene flames used in oxyfuel–gas welding and cutting operations: (a) neutral flame; (b) oxidizing flame; (c) carburizing, or reducing, flame. The gas mixture in (a) is basically equal volumes of oxygen and acetylene. (d) The principle of the oxyfuel–gas welding process.

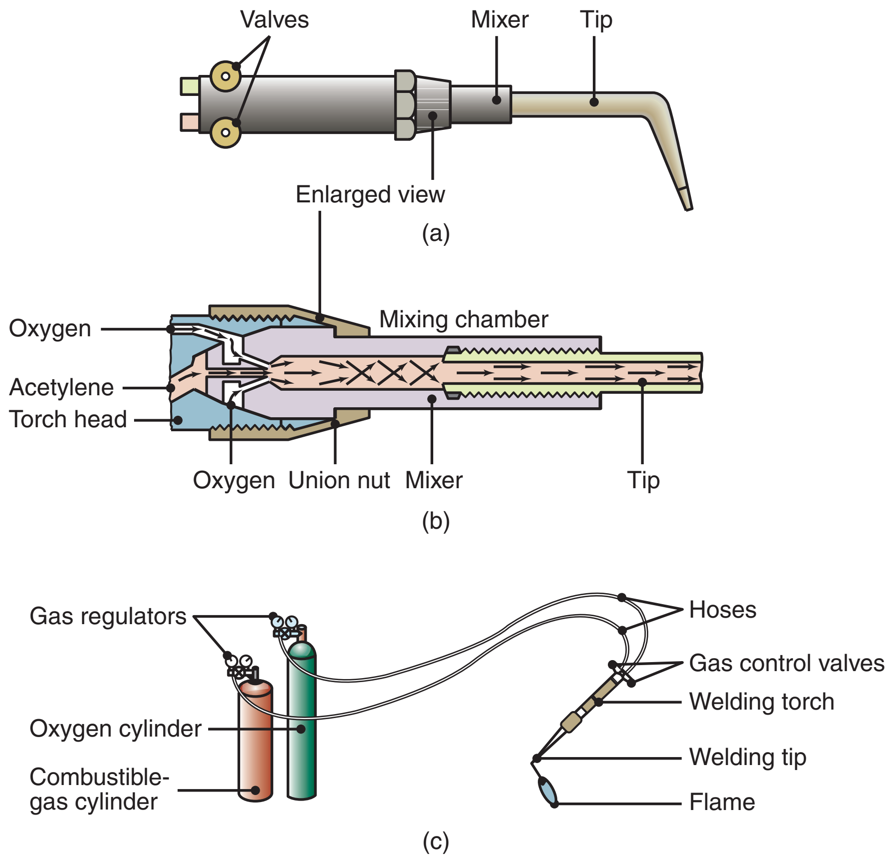


Figure 30.2: (a) General view of, and (b) cross section of, a torch used in oxyacetylene welding. The acetylene valve is opened first. The gas is lit with a spark lighter or a pilot light. Then the oxygen valve is opened and the flame is adjusted. (c) Basic equipment used in oxyfuel–gas welding. To ensure correct connections, all threads on acetylene fittings are left handed, whereas those for oxygen are right handed. Oxygen regulators usually are painted green and acetylene regulators red.

to thicknesses of less than 6 mm. Small joints made by this process may consist of a single-weld bead; deep-V groove joints are made in multiple passes. Cleaning the surface of each weld bead prior to depositing a second layer over it is important for joint strength and in avoiding defects (see Section 30.9). Wire brushes (hand or power) may be used for this purpose.

The equipment for oxyfuel–gas welding basically consists of a **welding torch**, connected by hoses to high-pressure gas cylinders, equipped with pressure gages and regulators (Fig. 30.2). The use of safety equipment, such as goggles with shaded lenses, face shields, gloves, and protective clothing, is essential. Proper connection of the hoses to the cylinders also is an important factor in safety. The oxygen and acetylene cylinders have different threads, so that the hoses cannot be connected to the wrong cylinders. Although it can be mechanized, the operation is essentially manual, and therefore slow. It has, however, the advantages of being portable, versatile, and economical for simple and low-quantity work.

Pressure-gas Welding. In this method, welding of two components starts with first heating the interface by means of a torch, using typically an oxyacetylene–gas mixture (Fig. 30.3a). After the interface begins to melt, the torch is withdrawn; a force is then applied to press the two components together and is maintained until the interface solidifies. Note in Fig. 30.3b the formation of a *flash* due to the upsetting of the joined ends of the two components.

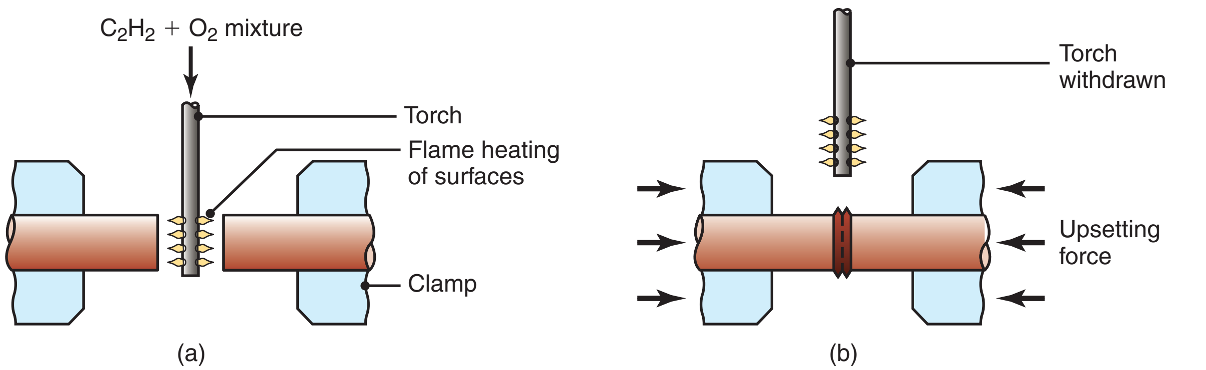


Figure 30.3: Schematic illustration of the pressure-gas welding process: (a) before and (b) after. Note the formation of a flash at the joint; later the flash can be trimmed off.

Thermit Welding. Also known as *thermite* or *exothermic welding*, and developed in 1895, *thermit welding* involves mixing a metal powder with a metal oxide, then using a high-temperature ignition source to cause an oxidation-reduction reaction (Fig. 30.4). A common arrangement in this process is to use iron oxide (rust) powder in combination with aluminum powder; upon ignition by a magnesium fuse, the resulting chemical reaction forms aluminum oxide (Al_2O_3) and iron.

Temperatures can reach up to 2500°C, melting the iron, which subsequently flows into a pouring basin and then into a mold placed around the parts to be welded. The aluminum oxide floats to the slag basin because of its lower density. The features of a thermit welding mold are very similar to a casting mold (see Fig. 11.3). Note from Fig. 30.4 the presence of a heating port, a feature that allows insertion of an oxyacetylene torch to preheat the workpieces and prevent weld cracks (Section 30.9.1).

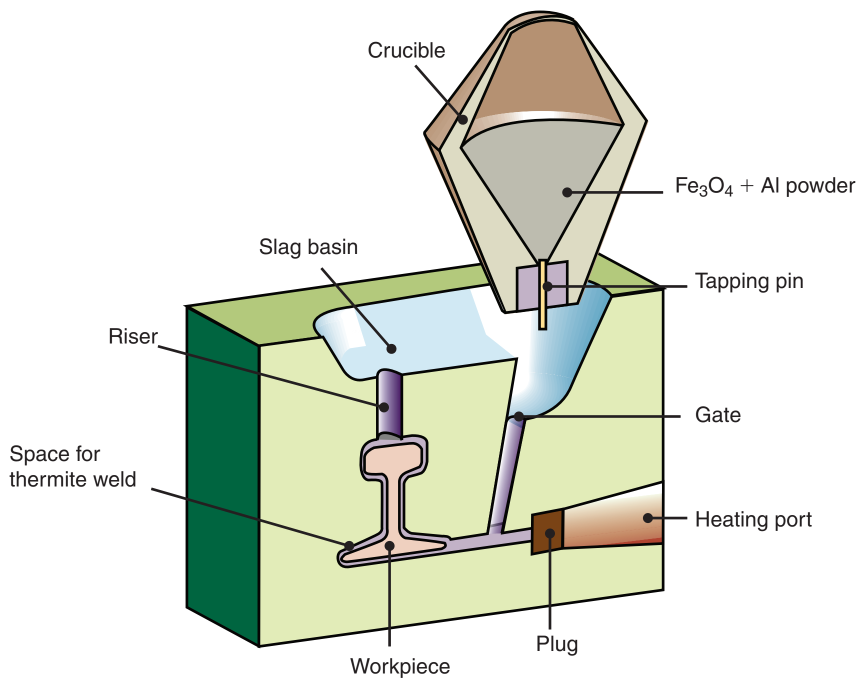


Figure 30.4: Schematic illustration of thermite welding.

Several combinations of powder and oxide can be used in thermit welding, although aluminum powder combined with iron oxide is the most common because of the widespread use of thermit welding for joining railroad rails. Some copper and magnesium oxides are often added to improve flammability. Other applications of thermit welding include welding of large-diameter copper conductors, using copper oxide, and field repair of large equipment, such as locomotive axle frames.

30.3 Arc-welding Processes: Nonconsumable Electrode

In *arc welding*, developed in the mid-1800s, the heat required is from electrical energy. The process involves either a *nonconsumable* or a *consumable electrode*. An AC or a DC power supply produces an arc between the tip of the electrode and the workpiece to be welded. The arc generates temperatures of about 30,000°C, higher than those developed in oxyfuel–gas welding.

Nonconsumable-electrode welding processes typically use a **tungsten electrode** (Fig. 30.5). Because of the high temperatures involved, an externally supplied shielding gas is necessary in order to prevent oxidation of the weld zone. Typically, *direct current* is used, and, as described below, its **polarity** (the direction of current flow) is important. The selection of current levels depends on such factors as the type of electrode, the metals to be welded, and the depth and width of the weld zone.

In **straight polarity**, also known as *direct-current electrode negative* (DCEN), the workpiece is positive (anode), and the electrode is negative (cathode). DCEN generally produces welds that are narrow and deep (Fig. 30.6a). In **reverse polarity**, also known as *direct-current electrode positive* (DCEP), the workpiece is negative and the electrode is positive. Weld penetration is less, and the weld zone is shallower and wider

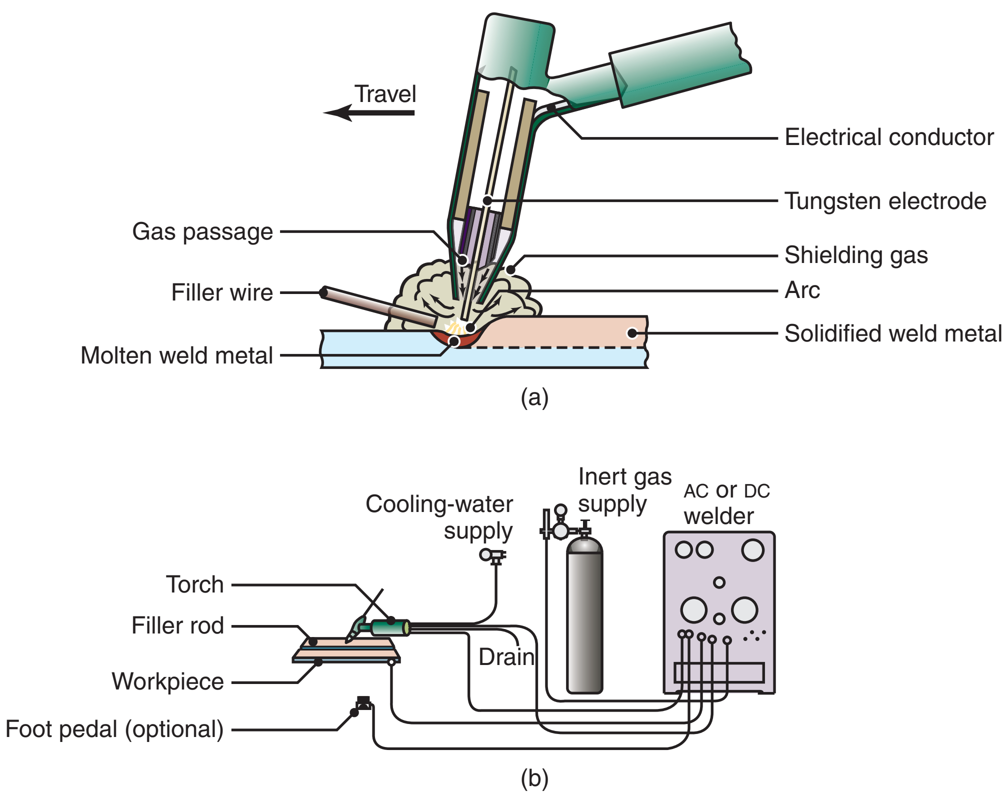


Figure 30.5: (a) The gas tungsten-arc welding process, formerly known as TIG (for tungsten–inert gas) welding. (b) Equipment for gas tungsten-arc welding operations.

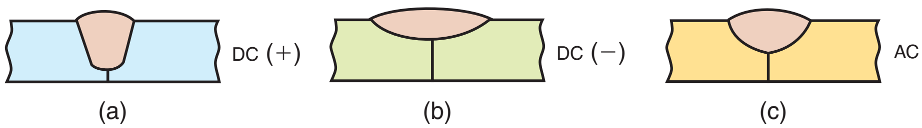


Figure 30.6: The effect of polarity and current type on weld beads: (a) DC current with straight polarity; (b) DC current with reverse polarity; (c) AC current.

(Fig. 30.6b); consequently, DCEP is preferred for sheet metals and for joints with wide gaps. In the **AC current** method, the arc pulsates rapidly. This method is suitable for welding thick sections and for using large-diameter electrodes at maximum currents (Fig. 30.6c).

Heat Transfer in Arc Welding. The heat input in arc welding is given by the equation

$$\frac{H}{l} = e\frac{VI}{v}, \tag{30.3}$$

where H is the heat input (J or BTU), l is the weld length, V is the voltage applied, I is the current (amperes), and v is the welding speed. The term e is the efficiency of the process, which varies from around 75% for shielded metal-arc welding to 90% for gas metal-arc and submerged-arc welding. The efficiency is an indication that not all of the available energy is beneficially used to melt the material; because the heat is conducted through the workpiece, some of it is dissipated by radiation and still more is lost by convection to the surrounding environment.

The heat input given by Eq. (30.3) melts a volume of material (usually the electrode or filler metal), and can also be expressed as

$$H = uV_m = uAl, \tag{30.4}$$

where u is the specific energy required for melting, V_m is the volume of metal melted, l is the length of the weld bead, and A is the cross-section of the weld. Some typical values of u are given in Table 30.2. Equations (30.3) and (30.4) allow an expression of the welding speed as

$$v = e\frac{VI}{uA}. \tag{30.5}$$

Although these equations have been developed for arc welding, similar expressions can be obtained for other fusion-welding operations, while taking into account differences in weld geometry and process efficiency.

Example 30.1 Welding Speed for Different Materials

Given: Consider a welding operation being performed with $V = 20$ volts, $I = 200$ A, and the cross-sectional area of the weld bead of $A = 30$ mm^2.

Find: Estimate the welding speed if the workpiece and electrode are made of (a) aluminum, (b) carbon steel, and (c) titanium. Assume an efficiency of 75%.

Solution: From Table 30.2, the specific energy required for aluminum is $u = 2.9$ J/mm^3. Therefore, from Eq. (30.5),

$$v = e\frac{VI}{uA} = (0.75)\frac{(20)(200)}{(2.9)(30)} = 34.5 \text{ mm/s}.$$

Similarly, for carbon steel, u is estimated as 9.7 J/mm^3 (average of extreme values in the table), and thus $v = 10.3$ mm/s. For titanium, $u = 14.3$ J/mm^3, and thus $v = 7.0$ mm/s.

Table 30.2: Approximate Specific Energies Required to Melt a Unit Volume of Commonly Welded Metals.

Material	Specific energy, *u* J/mm^3
Aluminum and its alloys	2.9
Cast irons	7.8
Copper	6.1
Bronze (90Cu–10Sn)	4.2
Magnesium	2.9
Nickel	9.8
Steels	9.1–10.3
Stainless steels	9.3–9.6
Titanium	14.3

Gas Tungsten-arc Welding. In *gas tungsten-arc welding* (GTAW), formerly known as *TIG* (for tungsten–inert gas) *welding*, the filler metal is supplied from a **filler wire** (Fig. 30.5a). Because the tungsten electrode is not consumed, a stable arc gap is maintained at a constant current level. The filler metals are similar to the metals to be welded, and flux is not used. The shielding gas is typically argon or helium, or a mixture of the two gases. Welding with GTAW may be done without filler metals, such as in welding close-fit joints.

Depending on the metals to be joined, the power supply is either DC at 200 A or AC at 500 A (Fig. 30.5b). In general, AC is preferred for aluminum and magnesium, because the cleaning action of AC removes oxides and improves weld quality. Thorium or zirconium may be used in tungsten electrodes to improve their electron emission characteristics. The power supply ranges from 8 to 20 kW. Contamination of the tungsten electrode by the molten metal can be a significant problem, particularly in critical applications, because it can cause discontinuities in the weld; contact of the electrode with the molten-metal pool should be avoided.

The GTAW process is used for a wide variety of metals and applications, particularly aluminum, magnesium, titanium, and refractory metals; it is especially suitable for thin metals. The cost of the inert gas makes this process more expensive than SMAW, but it provides welds of very high quality and good surface finish. GTAW is used in a variety of critical applications with a wide range of part thicknesses and shapes, and the equipment is portable.

Plasma-arc Welding. Plasma is an ionized hot gas consisting of nearly equal numbers of electrons and ions. In *plasma-arc welding* (PAW), developed in the 1960s, a concentrated plasma arc is produced and directed toward the weld area. The arc is stable and reaches temperatures as high as 33,000°C. The plasma is initiated between the tungsten electrode and the orifice by a low-current pilot arc. The plasma arc is concentrated, because it is forced through a small orifice. Operating currents usually are below 100 A, but they can be higher for special applications. When filler metal is used, it is fed into the arc, as is done in GTAW. Arc and weld-zone shielding is provided by means of an outer-shielding ring and the use of gases, such as argon, helium, or mixtures.

There are two methods of plasma-arc welding:

- In the **transferred-arc** method (Fig. 30.7a), the workpiece being welded is part of the electrical circuit. The arc transfers from the electrode to the workpiece, and hence the term *transferred*.

- In the **nontransferred** method (Fig. 30.7b), the arc is between the electrode and the nozzle, and the heat is carried to the workpiece by the plasma gas. This thermal-transfer mechanism is similar to that for an oxyfuel flame (see Section 30.2).

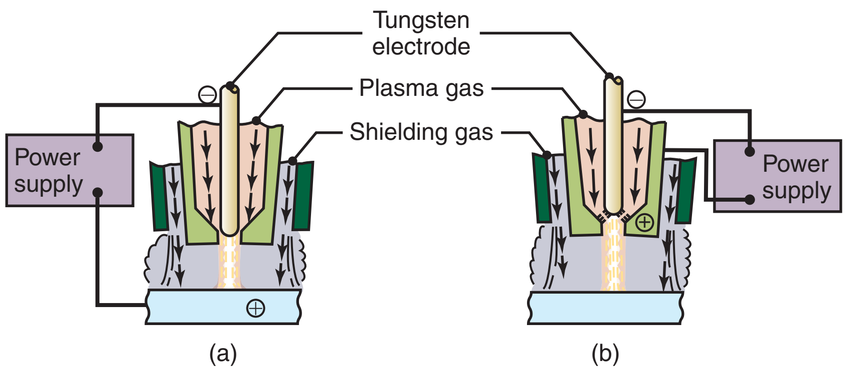


Figure 30.7: Two types of plasma-arc welding processes: (a) transferred and (b) nontransferred. Deep and narrow welds can be made by these processes at high welding speeds.

Compared with other arc-welding processes, plasma-arc welding has better arc stability, higher energy concentration, and less thermal distortion; the welds are deeper and narrower welds. Also, higher welding speeds, from 120 to 1000 mm/min, can be achieved. A variety of metals can be welded, with part thicknesses generally less than 6 mm.

The high heat concentration can completely penetrate through the joint, with thicknesses as much as 20 mm for some titanium and aluminum alloys. Known as the **keyhole technique**, the force of the plasma arc displaces the molten metal and produces a hole at the leading edge of the weld pool. Plasma-arc welding (rather than the GTAW process) is often used for butt and lap joints, because of its higher energy concentration, better arc stability, and higher welding speeds. Safety considerations include protection against glare, spatter, and noise from the plasma arc.

Atomic-hydrogen Welding. In *atomic-hydrogen welding* (AHW), an arc is generated between two tungsten electrodes within a shielding atmosphere of hydrogen gas. The gas normally is diatomic (H_2); however, where the temperatures are over 6000°C near the arc, the hydrogen breaks down into its atomic form, thus simultaneously absorbing a large amount of heat from the arc. When the gas strikes the relatively cold surface of the workpieces to be joined, it recombines into its diatomic form and rapidly releases the stored heat, reaching temperatures up to 4000°C. Thus, it is one of the few joining processes that can be used for welding tungsten. The energy in AHW can be easily varied by changing the distance between the arc stream and the workpiece surface.

30.4 Arc-welding Processes: Consumable Electrode

There are several consumable-electrode arc-welding processes, as described below.

30.4.1 Shielded Metal-arc Welding

Shielded metal-arc welding (SMAW) is one of the oldest, simplest, and most versatile joining processes; consequently, about 50% of all industrial and maintenance welding is done by this method. The electric arc is generated by touching the tip of a **coated electrode** against the workpiece, and withdrawing it quickly to a distance sufficient to maintain the arc (Fig. 30.8a). The electrodes are in the shapes of thin, long round rods that are held manually; hence, the process also is referred to as **stick welding**.

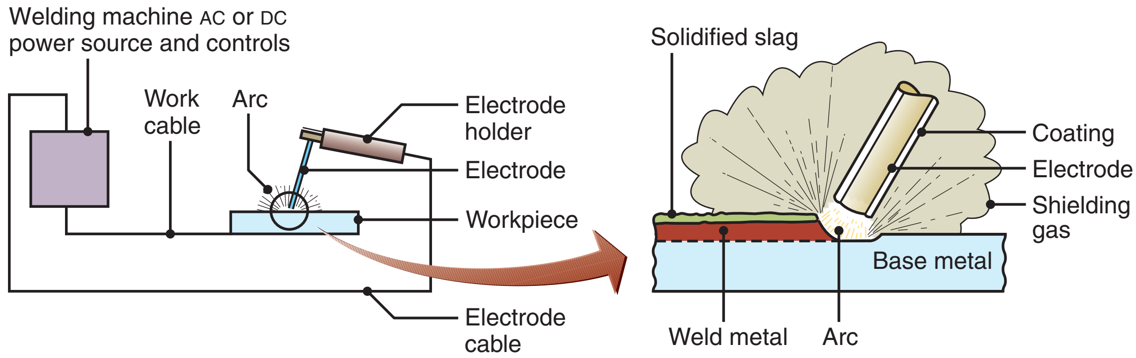


Figure 30.8: Schematic illustration of the shielded metal-arc welding process. About 50% of all large-scale industrial-welding operations use this process.

The heat generated melts a portion of the electrode tip, its coating, and the base metal in the immediate arc area. The molten metal consists of a mixture of base metal (the workpiece), electrode metal, and substances from the coating on the electrode; this mixture becomes the weld when it solidifies. The electrode coating deoxidizes the weld area and provides a shielding gas to protect it from oxygen in the environment.

A bare section at the end of the electrode is first clamped to one terminal of the power source, while the other terminal is connected to the workpiece being welded (Fig. 30.8b). The current, which may be DC or AC, usually ranges from 50 to 300 A. For sheet-metal welding, DC is preferred because of the steady arc it produces. Power requirements generally are less than 10 kW. The equipment consists of a power supply, cables, and an electrode holder.

The SMAW process is commonly used in general construction, shipbuilding, pipelines, and for maintenance. It is especially useful for work in remote areas where a portable fuel-powered generator can be used as the power supply. This process is best suited for workpiece thicknesses of 3 to 19 mm, although this range can easily be extended by skilled operators using *multiple-pass* techniques (Fig. 30.9).

The multiple-pass approach requires that slag be removed after each weld bead. Unless removed completely, the solidified slag can cause severe corrosion of the weld area, and thus lead to failure of the weld; also, it prevents fusion of weld layers, compromising weld strength. The slag can be removed by wire brushing or by chipping of the weld. Labor costs and material costs are both high.

30.4.2 Submerged-arc Welding

In *submerged-arc welding* (SAW), the weld arc is shielded by a *granular flux*, which consists of lime, silica, manganese oxide, calcium fluoride, and other compounds. The flux is fed into the weld zone from a hopper by gravity flow through a nozzle (Fig. 30.10). The thick layer of flux completely covers the molten metal, and prevents spatter and sparks, and suppresses the intense ultraviolet radiation and fumes characteristic of the SMAW process. The flux also acts as a thermal insulator, by promoting deep penetration of heat into the workpiece.

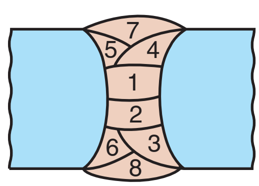


Figure 30.9: A deep weld showing the buildup sequence of eight individual weld beads.

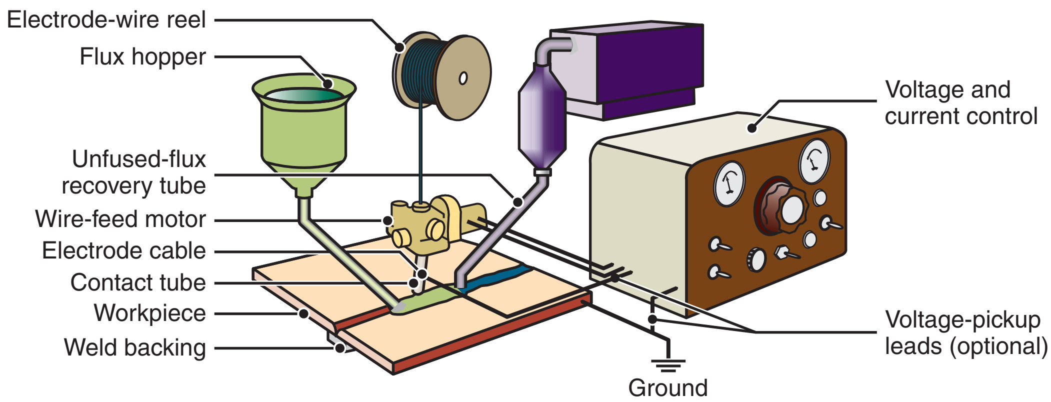


Figure 30.10: Schematic illustration of the submerged-arc welding process and equipment. The unfused flux is recovered and reused.

The consumable electrode is a coil of bare round wire, 1.5–10 mm in diameter, and is fed automatically through a tube (**welding gun**). Electric currents typically range from 300 to 2000 A, but multiple arc arrangements can be as high as 5000 A. The power supplies are usually connected to standard single- or three-phase power lines, with a primary rating up to 440 V.

Because the flux is gravity fed, the SAW process is limited mostly to welds in a flat or horizontal position, and having a backup piece. Circular welds can be made on pipes and cylinders, provided that they can be rotated during welding. As Fig. 30.10 illustrates, the unfused flux can be recovered, treated, and reused; typically, 50–90% of the flux is recovered. The process is automated and is used to weld a variety of carbon and alloy steels and stainless-steel sheets or plates, at speeds as high as 5 m/min; occasionally it is also used for nickel-based alloys. The quality of the weld is very high, with good toughness, ductility, and uniformity of properties. The SAW process provides very high welding productivity, depositing 4–10 times the amount of weld metal per hour as the SMAW process. Typical applications include welding of thick plates for shipbuilding and pressure vessels.

30.4.3 Gas Metal-arc Welding

In *gas metal-arc welding* (GMAW), developed in the 1950s and formerly called *metal inert-gas* (MIG) *welding*, the weld area is shielded by an inert atmosphere of argon, helium, carbon dioxide, or other gas mixtures (Fig. 30.11a). The consumable bare wire is automatically fed through a nozzle into the weld arc by a wire-feed drive motor (Fig. 30.11b). Multiple-weld layers also can be deposited at the joint. Deoxidizers usually are present in the electrode metal itself, in order to prevent oxidation of the molten-weld puddle.

Metal can be transferred by three methods:

1. In **spray transfer**, small, molten droplets from the electrode are transferred to the weld area, at a rate of several hundred droplets per second. The transfer is spatter free and very stable. High DC currents and voltages and large-diameter electrodes are used with argon or an argon-rich gas mixture as the shielding gas. The average current required can be reduced by using a **pulsed arc**, superimposing high-amplitude pulses onto a steady low current. The process can be used in all welding positions.

2. In **globular transfer**, carbon-dioxide-rich gases are utilized, and the globules are propelled by the forces of the electric-arc transfer of the metal, resulting in considerable spatter. Welding currents are high, making it possible for deeper weld penetration; welding speeds are higher than in spray transfer. Heavier sections are commonly welded by this method.

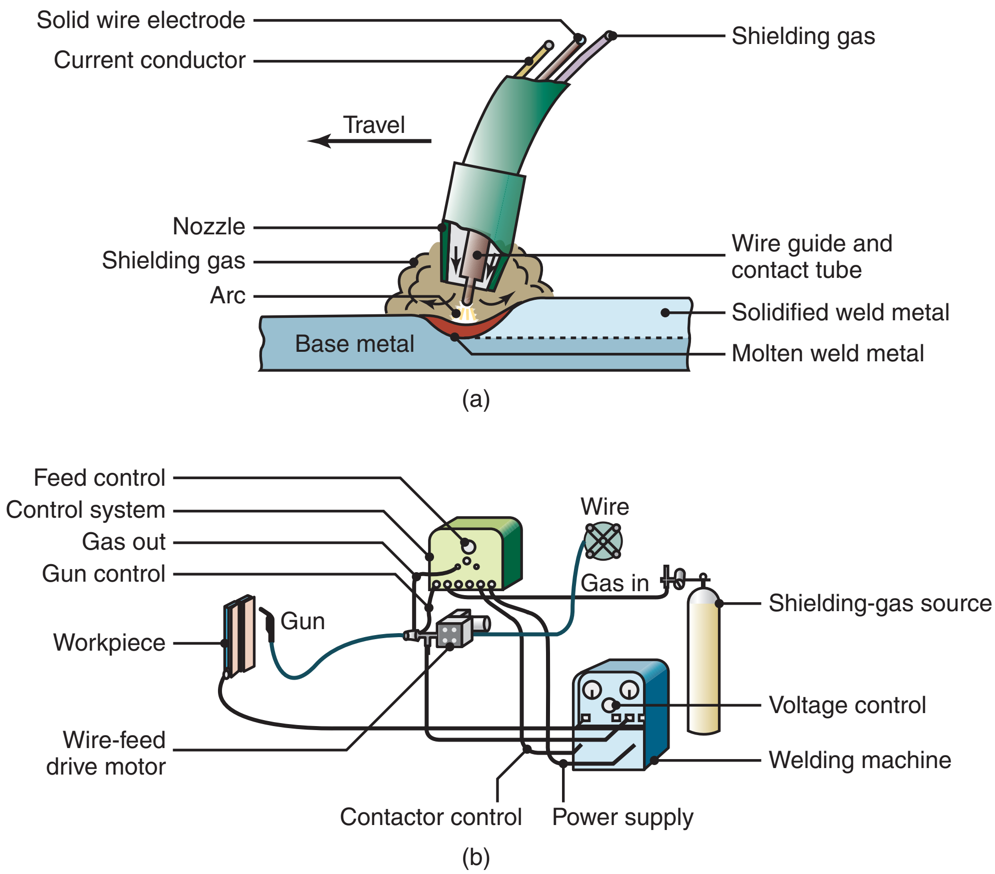


Figure 30.11: (a) Schematic illustration of the gas metal-arc welding process, formerly known as MIG (for metal inert-gas) welding. (b) Basic equipment used in gas metal-arc welding operations.

3. In **short circuiting**, the metal is transferred as individual droplets (at a rate more than 50/s), as the electrode tip touches the molten weld metal and short circuits. Low currents and voltages are utilized with carbon-dioxide-rich gases, and electrodes made of small-diameter wire. Power requirement is about 2 kW.

The temperatures generated in GMAW are relatively low; consequently, this method is suitable only for thin sheets and sections of less than 6 mm, as otherwise fusion may be incomplete. The operation, which is easy to perform, is commonly used for welding ferrous metals with thin sections. Pulsed-arc systems are used for thin ferrous and nonferrous metals.

The GMAW process is suitable for welding most ferrous and nonferrous metals and is used extensively in the metal-fabrication industry. Because of the relative simplicity of the process, training of operators is easy. The process is versatile, rapid, and economical, and welding productivity is double that of the SMAW process. The process can easily be automated and lends itself readily to robotics and to flexible manufacturing systems (Chapters 37 and 39).

30.4.4 Flux-cored Arc Welding

Flux-cored arc welding (FCAW), illustrated in Fig. 30.12, is similar to gas metal-arc welding, with the exception that the electrode is tubular and is filled with flux, and hence the term *flux-cored*. Cored electrodes

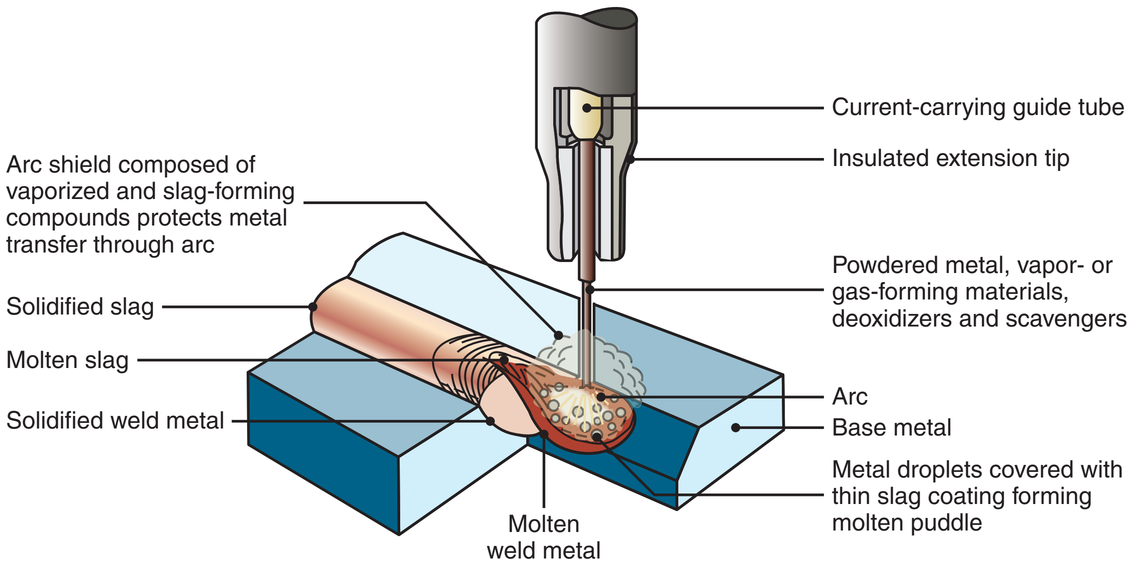


Figure 30.12: Schematic illustration of the flux-cored arc-welding process. This operation is similar to gas metal-arc welding, shown in Fig. 30.11.

produce a more stable arc, improve the weld contour, and produce better mechanical properties of the joint. The flux in the electrodes is much more flexible than the brittle coating used on SMAW electrodes, as the tubular electrode can be made in long coiled lengths.

The electrodes are usually 0.5–4 mm in diameter, and the power required is about 20 kW. *Self-shielded cored electrodes* also are available; they do not require any external shielding gas, because they contain emissive fluxes that shield the weld area against the surrounding atmosphere. Small-diameter electrodes have made welding of thinner materials not only possible but often preferable. Furthermore, small-diameter electrodes make it relatively easy to weld parts at various positions, and the flux chemistry permits the welding of many metals.

The FCAW process combines the versatility of SMAW with the continuous and automatic electrode-feeding feature of GMAW. The process is economical and versatile; thus it is used for welding various types of joints, mainly on steels, stainless steels, and nickel alloys. The higher weld-metal deposition rate of this process, as compared with that of GMAW, has led to its use in joining sections with various thicknesses. Using *tubular electrodes* with very small diameters has extended the use of this process to workpieces with small cross-sections.

A major advantage of FCAW is the ease with which specific weld-metal chemistries can be developed and used, by adding various alloying elements to the flux core. The process is easy to automate and is readily adaptable to flexible manufacturing systems and robotics.

30.4.5 Electrogas Welding

Electrogas welding (EGW) is primarily used for welding the edges of parts, vertically and in one pass, with the parts placed edge to edge (*butt joint*). The process is classified as *machine welding*, because it requires special equipment (see Fig. 30.13). The weld metal is deposited into a weld cavity between the two parts to be joined. The space in between is enclosed by two water-cooled copper *dams* (*shoes*) to prevent the molten slag from running off; mechanical drives move the shoes upward. Circumferential welds, such as those on pipes, also are possible, provided that the workpiece can be rotated.

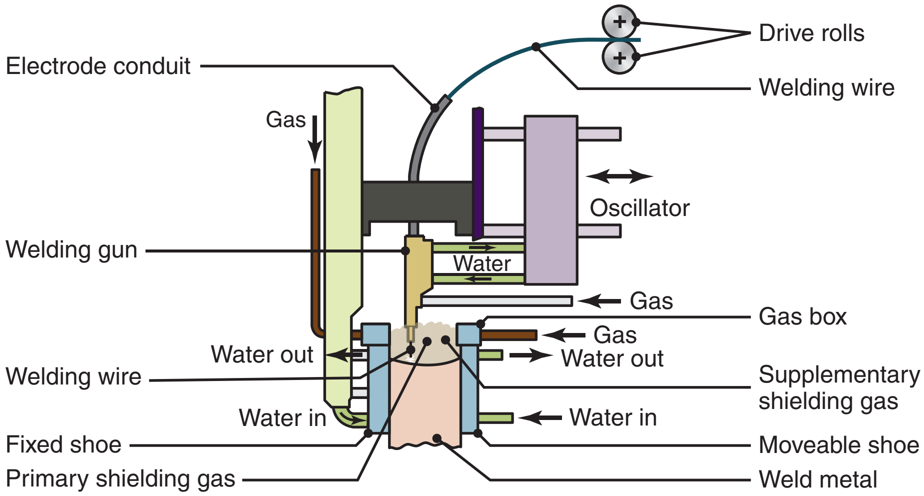


Figure 30.13: Schematic illustration of the electrogas-welding process.

Single or multiple electrodes are fed through a conduit, and a continuous arc is maintained, using flux-cored electrodes at up to 750 A or solid electrodes at 400 A. Power requirements are about 20 kW. Shielding is provided by means of an inert gas, such as carbon dioxide, argon, or helium, depending on the type of material being welded. The gas may be supplied either from an external source or from a flux-cored electrode, or from both.

Weld thickness ranges from 12 to 75 mm, on steels, titanium, and aluminum alloys. Typical applications include construction of bridges, pressure vessels, thick-walled and large-diameter pipes, storage tanks, and ships. The equipment is reliable and training for operators is relatively simple.

30.4.6 Electroslag Welding

Electroslag welding (ESW) and its applications are similar to electrogas welding (Fig. 30.14), the main difference being that the arc is initiated between the electrode tip and the bottom of the part to be welded. Flux is added, which then melts by the heat of the arc. After the molten slag reaches the tip of the electrode, the arc is extinguished. Heat is produced continuously by the electrical resistance of the molten slag. Because the arc is extinguished, ESW is not strictly an arc-welding process. Single or multiple solid or flux-cored electrodes may be used.

Electroslag welding is capable of welding plates with thicknesses ranging from 50 mm to more than 900 mm, and welding is done in one pass. The current required is on the order of 600 A at 40 to 50 V, although higher currents are used for thick plates; the travel speed of the weld is in the range of 12–36 mm/min. This process is used for large structural-steel sections, such as heavy machinery, bridges, oil rigs, ships, and nuclear-reactor vessels; weld quality is good.

30.5 Electrodes for Arc Welding

Electrodes for consumable arc-welding processes are classified according to the following properties:

- Strength of the deposited weld metal
- Current (AC or DC)
- Type of coating

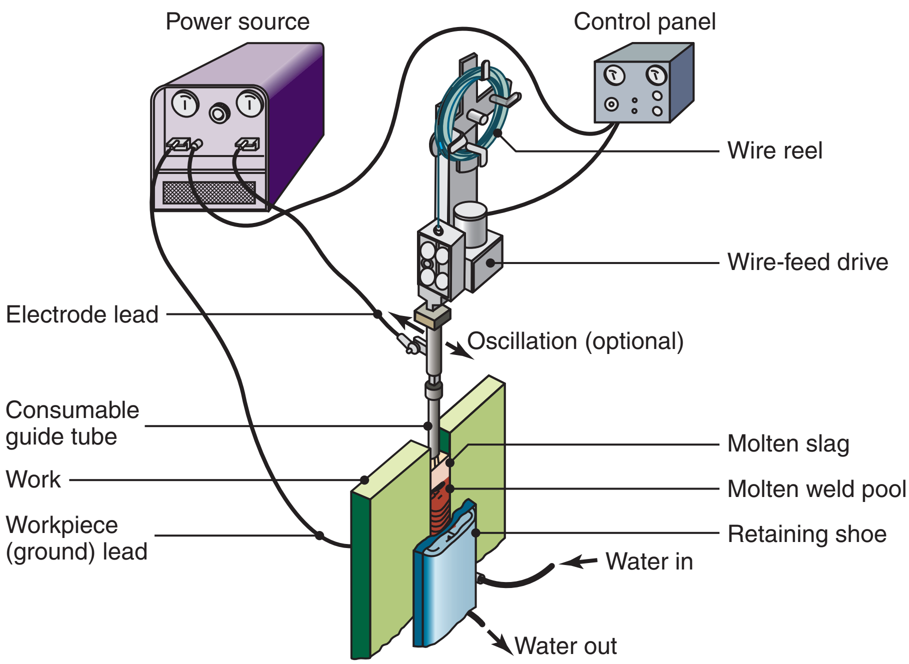


Figure 30.14: Equipment used for electroslag-welding operations.

Electrodes are identified by numbers and letters (Table 30.3), or by color code if the numbers and letters are too small to imprint. Typical coated-electrode dimensions are in the range of 150–460 mm in length and 1.5–8 mm in diameter.

Specifications for electrodes and filler metals, including dimensional tolerances, quality control procedures, and processes, are published by the American Welding Society (AWS) and the American National Standards Institute (ANSI). Some specifications are available in the Aerospace Materials Specifications (AMS) by the Society of Automotive Engineers (SAE). Electrodes are sold by weight and are available in a wide variety of sizes and specifications. Criteria for selection and recommendations for electrodes for a particular metal and its application can be found in suppliers' literature and in various handbooks and references, as in the Bibliography of this chapter.

Electrode Coatings. Electrodes are *coated* with claylike materials, which include silicate binders and powdered materials, including oxides, carbonates, fluorides, metal alloys, cotton cellulose, and wood flour. The coating is brittle and takes part in complex interactions during welding; it has the following basic functions:

- Stabilize the arc
- Generate gases to act as a shield against the surrounding atmosphere; the gases produced are carbon dioxide, water vapor, and small amounts of carbon monoxide and hydrogen
- Control the rate at which the electrode melts
- Act as a flux to protect the weld against the formation of oxides, nitrides, and other inclusions and, with the resulting slag, to protect the molten-weld pool
- Add alloying elements to the weld zone to enhance the properties of the joint—among these elements are deoxidizers to prevent the weld from becoming brittle.

Table 30.3: Designations for Mild-steel Coated Electrodes.

The prefix "E" designates arc-welding electrode.	
The first two digits of four-digit numbers and the first three digits of five-digit numbers indicate minimum tensile strength:	
E60XX	413 MPa
E70XX	482 MPa
E110XX	758 MPa
The next-to-last digit indicates position:	
EXX1X	All positions
EXX2X	Flat position and horizontal fillets
The last two digits together indicate the type of covering and the current to be used.	
The suffix (Example: EXXXX-A1) indicates the approximate alloy in the weld deposit:	
–A1	0.5% Mo
–B1	0.5% Cr, 0.5% Mo
–B2	1.25% Cr, 0.5% Mo
–B3	2.25% Cr, 1% Mo
–B4	2% Cr, 0.5% Mo
–B5	0.5% Cr, 1% Mo
–C1	2.5% Ni
–C2	3.25% Ni
–C3	1% Ni, 0.35% Mo, 0.15% Cr
–D1 and D2	0.25–0.45% Mo, 1.75% Mn
–G	0.5% min. Ni, 0.3% min. Cr, 0.2% min. Mo, 0.1% min. V, 1% min. Mn (only one element required)

The deposited coating or slag must be removed following each pass in order to ensure a good weld. Bare electrodes and wires, typically made of stainless steels and aluminum alloys, also are available, and are used as filler metals in various welding operations.

30.6 Electron-beam Welding

In *electron-beam welding* (EBW), developed in the 1960s, heat is generated by high-velocity, narrow-beam electrons. The kinetic energy of the electrons is converted into heat as they strike the workpiece to be welded. This process requires special equipment in order to focus the beam on the workpiece, typically in a vacuum. The higher the vacuum, the greater the depth the beam penetrates, and the greater becomes the depth-to-width ratio of the weld; thus, the methods are called EBW-HV (for high vacuum) and EBW-MV (for medium vacuum); some materials may also be welded by EBW-NV (for no vacuum).

Almost any metal can be welded and workpiece thicknesses range from foil to plate. Capacities of electron guns range up to 100 kW; the intense energy also is capable of producing holes in the workpiece. Generally, no shielding gas, flux, or filler metal is required.

The EBW process makes high-quality welds that are deep and narrow, and with small heat-affected zones (Section 30.9). Depth-to-width ratios are in the range of 10–30. The size of welds made are much smaller than those made by conventional processes. Using automation and servo controls, the processing parameters can be controlled accurately, at welding speeds as high as 12 m/min.

Almost any metal can be welded with this process, in butt or lap configurations, and at thicknesses up to 150 mm. Distortion and shrinkage are minimal, and weld quality is good. Typical applications include aircraft, missile, nuclear, and electronic components, and gears and shafts for the automotive industry. EBW equipment generates X-rays; thus, proper monitoring and periodic maintenance of the equipment are essential.

30.7 Laser-beam Welding

Laser-beam welding (LBW) utilizes a high-power laser beam as the source of heat to produce a fusion weld. Because it can be focused onto a very small area, the beam has high energy density and deep penetrating capability. The laser beam can be directed, shaped, and focused precisely, with laser spot diameters as low as 0.2 mm. LBW is suitable particularly for welding deep and narrow joints (Fig. 30.15) with depth-to-width ratios typically ranging from 4 to 10.

Laser-beam welding has become very widespread and is now used by most industries. The laser beam may be **pulsed** (in milliseconds), with power levels up to 100 kW, for applications such as spot welding of thin materials. **Continuous** multi-kW laser systems are used for deep welds on thick sections.

Laser beam welds have good quality, with minimal shrinkage or distortion. The welds have good strength and are generally ductile and free of porosity. The process can be automated and used on a variety of materials, with thicknesses up to 25 mm. As described in Section 16.2.2, *tailor-welded sheet-metal blanks* are joined principally by laser-beam welding, using robotics for precise control of the beam path.

Typical metals and alloys welded include aluminum, titanium, ferrous metals, copper, superalloys, and the refractory metals. Welding speeds range from 2.5 m/min to as high as 80 m/min for thin metals. Because of the nature of the process, welding can be done in otherwise inaccessible locations. As in other and similar automated welding systems, the operator skill required is minimal. Safety is particularly important in laser-beam welding due to the extreme hazards to the eye and the skin; solid-state (YAG) lasers also are dangerous.

While a filler wire can be used, laser-beam welding generally does not use a filler metal; instead, the laser melts the material which then solidifies to weld the components. In some arrangements (similar to laser-engineered net shaping, Section 20.8), powder can be blown in front of the laser, leading to deposition of metal.

The major advantages of LBW over EBW are:

- A vacuum is not required, and the beam can be transmitted through air.
- Laser beams can be shaped, manipulated, and focused by means of fiber optics; hence the process can easily be automated.
- The beams do not generate X-rays.
- The quality of the weld is better than in EBW; there is less part distortion and the weld has less tendency for incomplete fusion, spatter, and porosity.

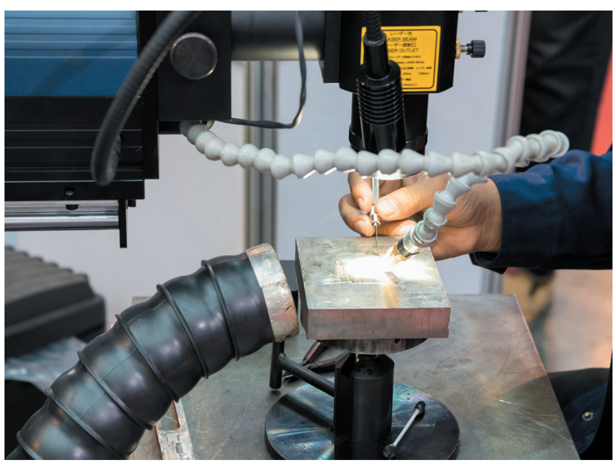


Figure 30.15: Laser beam welding in progress. *Source:* Alamy/Warut Sintapanon.

Figure 30.16: Detail of razor cartridge, showing laser spot welds. *Source:* Shutterstock/All About Space.

Example 30.2 Laser Welding of Razor Blades

The Gillette Sensor® razor cartridge has two narrow, high-strength blades with 13 pinpoint welds, 11 of which can be seen as darker spots, about 0.5 mm in diameter, on each blade. The welds are made with an Nd:YAG laser, equipped with fiber-optic delivery. This equipment provides very flexible beam manipulation and can target exact locations along the length of the blade. With a set of these machines, production is at a rate of 3 million welds per hour, with consistent weld quality.

Source: Courtesy of Lumonics Corporation, Industrial Products Division.

Laser GMAW. *Laser GMAW* is an emerging hybrid welding technology that combines the narrow heat-affected zone of laser welding with the high deposition rates of gas metal-arc welding. In this process, shown in Fig. 30.17, the laser is focused on the workpiece ahead of the GMAW arc, resulting in deep

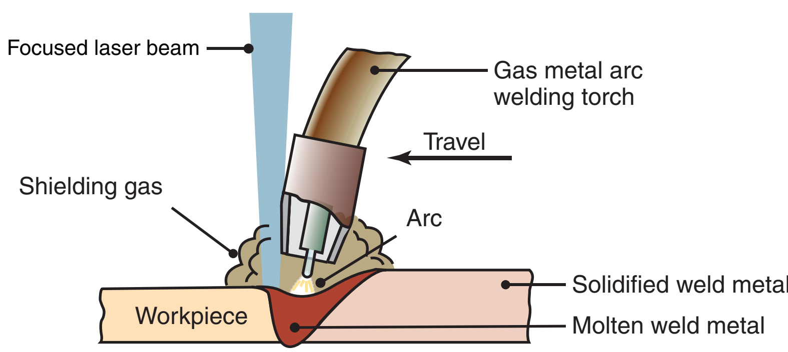


Figure 30.17: Schematic illustration of the Laser GMAW hybrid welding process. *Source:* Courtesy of Lincoln Electric.

penetration and allowing high travel speeds. In addition, the process can bridge gaps larger than in traditional laser welding, and the metallurgical quality of the weld is improved because of the presence of the shielding gas.

30.8 Cutting

In addition to being cut by mechanical means, as described in Part IV, a material can be cut into various contours by using a heat source that melts and removes a narrow zone in the workpiece. The sources of heat can be torches, electric arcs, or lasers.

Oxyfuel–gas Cutting. *Oxyfuel–gas cutting* (OFC) is similar to oxyfuel–gas welding (Section 30.2), but the heat source is now used to *remove* a narrow zone from a metal plate or sheet (Fig. 30.18a). OFC is suitable particularly for steels, where the basic reactions are

$$\mathrm{Fe} + \mathrm{O} \rightarrow \mathrm{FeO} + \mathrm{Heat}, \tag{30.6}$$

$$3\mathrm{Fe} + 2\mathrm{O}_2 \rightarrow \mathrm{Fe}_3\mathrm{O}_4 + \mathrm{Heat}, \tag{30.7}$$

and

$$4\mathrm{Fe} + 3\mathrm{O}_2 \rightarrow 2\mathrm{Fe}_2\mathrm{O}_3 + \mathrm{Heat}. \tag{30.8}$$

The greater heat is generated by the second reaction, with temperatures rising to about 870°C. However, because this temperature is not sufficiently high, the workpiece is first *preheated* with fuel gas, and then oxygen is introduced, as can be seen from the nozzle cross-section in Fig. 30.18a. The higher the carbon content of the steel, the higher is the required preheating temperature. Cutting takes place mainly by oxidation of the steel; some melting also takes place. Cast irons and steel castings also can be cut by this method. Cutting generates a **kerf**, similar to that produced in sawing with a saw blade or by wire electrical-discharge machining (see Fig. 27.12). Kerf width ranges from about 1.5 to 10 mm, with good control of dimensional tolerances. However, distortion caused by uneven temperature distribution can be a problem in OFC.

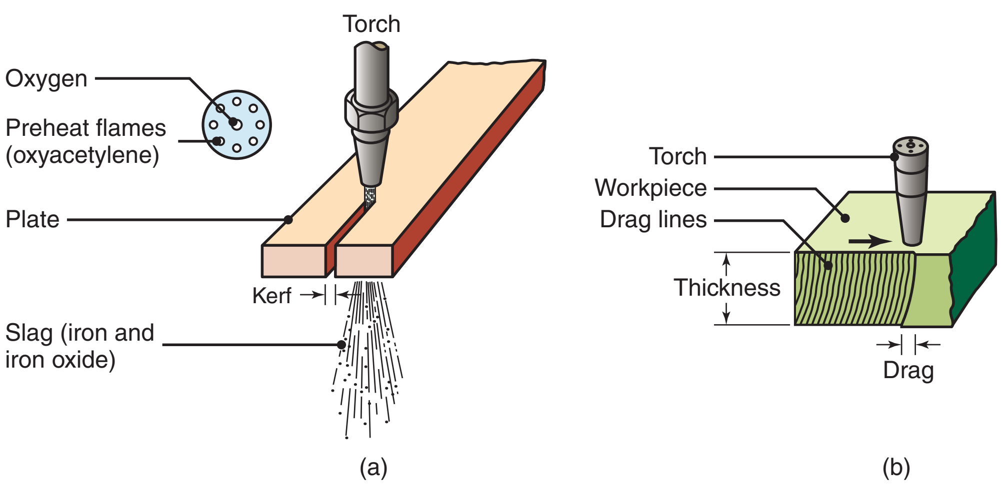


Figure 30.18: (a) Flame cutting of a steel plate with an oxyacetylene torch, and a cross-section of the torch nozzle. (b) Cross-section of a flame-cut plate, showing drag lines.

The maximum thickness that can be cut by OFC depends mainly on the gases used. With oxyacetylene gas, for example, the maximum thickness is about 300 mm, whereas, with oxyhydrogen, it is about 600 mm. The flame leaves **drag lines** on the cut surface (Fig. 30.18b), resulting in a rougher surface than that produced by such processes as sawing and blanking that use mechanical cutting tools. *Underwater cutting* is done with specially designed torches that produce a blanket of compressed air between the flame and the surrounding water. Torches may be guided along specified paths either manually, mechanically, or automatically by machines, using programmable controllers and robots.

Arc Cutting. *Arc-cutting* processes are based on the same principles as arc welding. A variety of materials can be cut at high speeds by arc cutting, although, as in welding, these processes also leave a heat-affected zone that has to be taken into account, particularly in critical applications.

In **air carbon-arc cutting** (CAC-A), a carbon electrode is used and the molten metal is blown away by a high-velocity air jet. The process is used especially for gouging and scarfing (removal of metal from a surface). However, it is noisy, and the molten metal can be blown substantial distances and can cause safety hazards.

Plasma-arc cutting (PAC) produces the highest temperatures, and is used for rapid cutting of nonferrous and stainless-steel plates. The productivity of this process is higher than that of oxyfuel–gas methods. PAC produces a good surface finish and with narrow kerfs. **Electron beams** and **lasers** also are used for very accurately cutting a wide variety of metals, as described in Sections 27.6 and 27.7. The surface finish is better than that of other thermal cutting processes, and the kerf is narrower.

30.9 The Weld Joint, Quality and Testing

Three distinct zones can be identified in a typical weld joint, as shown in Fig. 30.19:

1. **Base metal**
2. **Heat-affected zone**
3. **Weld metal**

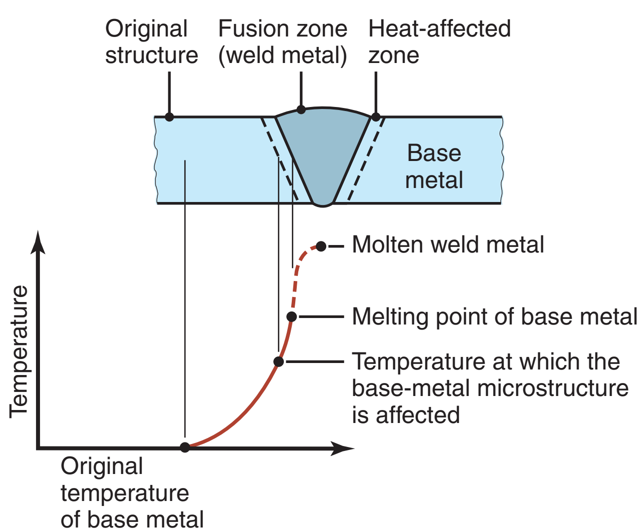


Figure 30.19: Characteristics of a typical fusion-weld zone in oxyfuel–gas and arc welding.

The metallurgy and properties of the second and third zones depend strongly on the type of metals joined, the particular joining process, the filler metals used (if any), and welding process variables. Recall that a joint produced without using a filler metal is called *autogenous*; its weld zone is composed of the *resolidified base metal*. A joint made with a filler metal has a central zone, called the *weld metal*, and is composed of a mixture of the base and the filler metals.

Solidification of the Weld Metal. After the application of heat and introducing the filler metal, if any, into the weld zone, the weld joint is allowed to cool to ambient temperature. The solidification process is similar to that in casting (Section 10.2); it begins with the formation of *columnar* (*dendritic*) grains, as shown in Fig. 10.3. These grains are relatively long and they form parallel to the heat flow. Because metals are much better thermal conductors than the surrounding air, the grains lie parallel to the plane of the two components being welded (Fig. 30.20a); in contrast, the grains in a shallow weld are as shown in Fig. 30.20b and c.

Grain structure and grain size depend on the specific metal alloy, the welding process employed, and the type of filler metal. Because it begins with a molten state, the weld metal basically has a *cast structure*, and since it has cooled slowly, the grains are coarse. Consequently, this structure generally has low strength, toughness, and ductility; however, with proper selection of filler-metal composition or of heat treatments following welding, the mechanical properties of the joint can be improved.

The resulting structure depends on the particular alloy, its composition, and the thermal cycling to which the joint is subjected. For example, cooling rates may be controlled and reduced by *preheating* the general weld area prior to welding it. Preheating is important, particularly for metals having high thermal conductivity, such as aluminum and copper (Table 3.2). Without preheating, the heat produced during welding dissipates rapidly through the rest of the parts being joined.

Heat-affected Zone. The *heat-affected zone* (HAZ) is within the base metal itself. It has a microstructure different from that of the base metal prior to welding, because it has been temporarily subjected to elevated

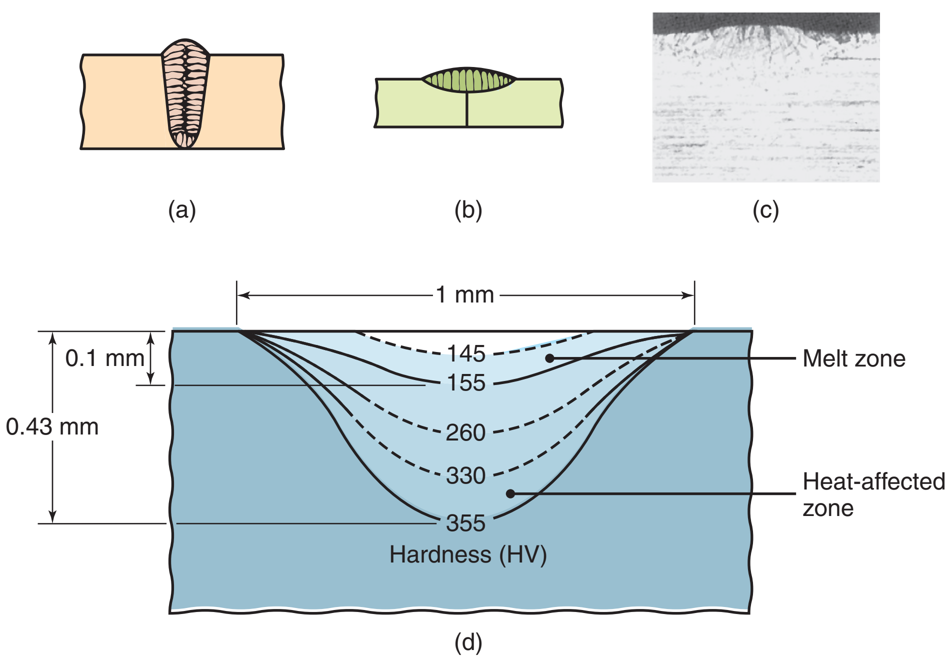


Figure 30.20: Grain structure in (a) a deep weld and (b) a shallow weld. Note that the grains in the solidified weld metal are perpendicular to their interface with the base metal. (c) Weld bead on a cold-rolled nickel strip produced by a laser beam. (d) Microhardness (HV) profile across a weld bead.

temperatures during welding. The portions of the base metal far away from the heat source do not undergo microstructural changes during welding, because of the much lower temperature to which they have been subjected.

The properties and microstructure of HAZ depend on (a) the rate of heat input and cooling and (b) the temperature to which this zone was raised. In addition to metallurgical factors, such as original grain size, grain orientation, and degree of prior cold work, physical properties, such as the specific heat and thermal conductivity of the metals, also influence the size and characteristics of HAZ.

The strength and hardness of HAZ (Fig. 30.20d) depend partly on how the original strength and hardness of the base metal was developed originally. As described in Chapters 2 and 4, they may have been developed by (a) cold working, (b) solid-solution strengthening, (c) precipitation hardening, or (d) heat treatments. The effects of these strengthening methods are complex.

The heat applied during welding *recrystallizes* the elongated grains of the cold-worked base metal. Grains that are away from the weld metal will recrystallize into fine, equiaxed grains; grains close to the weld metal have been subjected to elevated temperatures for a longer time, and thus they will grow in size (*grain growth*, Section 1.7). This region will be softer and have lower strength; such a joint will be weakest at its HAZ.

Joints made from dissimilar metals and for alloys strengthened by other methods, the effects of heat on HAZ are complex, and beyond the scope of this book. Details can be found in more advanced texts; see also the Bibliography at the end of this chapter.

30.9.1 Weld Quality

As a result of a history of thermal cycling and its attendant microstructural changes, a welded joint may develop various **discontinuities**. Welding discontinuities also can be caused by an inadequate or careless application of welding techniques or poor operator training. The major discontinuities that affect weld quality are described below.

Porosity. *Porosity* in welds may be caused by

- Gases released during melting of the weld area but trapped during solidification
- Chemical reactions during welding
- Contaminants.

Most welded joints have some porosity, generally in the shape of spheres or of elongated pockets (see also Section 10.6.1). The distribution of porosity in the weld zone may be random or the porosity may be concentrated in a certain region in the zone. Porosity in welds can be reduced by the following practices:

- Proper selection of electrodes and filler metals
- Improved welding techniques, such as preheating the weld area, or increasing the rate of heat input
- Proper cleaning and prevention of contaminants from entering the weld zone
- Reduced welding speeds, to allow time for gas to escape.

Slag Inclusions. *Slag inclusions* are compounds, such as oxides, fluxes, and electrode-coating materials, that are trapped in the weld zone. If shielding gases are not effective during welding, contamination from the environment also may contribute to such inclusions. Welding conditions also are important: with control of processing parameters, the molten slag will float to the surface of the molten weld metal, and thus it will not become entrapped.

Slag inclusions can be prevented by implementing the following practice:

- Cleaning the weld-bead surface with a wire brush (hand or power) or with a chipper before the next layer is deposited
- Providing sufficient shielding gas
- Redesigning the joint to permit sufficient space for proper manipulation of the puddle of molten weld metal.

Incomplete Fusion. *Incomplete fusion* produces poor weld beads, such as those shown in Fig. 30.21. A better weld can be obtained by implementing the following practices:

- Raising the temperature of the base metal
- Cleaning the weld area prior to welding
- Modifying the joint design
- Changing the type of electrode
- Providing sufficient shielding gas.

Incomplete penetration occurs when the depth of the welded joint is insufficient. Penetration can be improved by:

- Increasing the heat input
- Reducing the travel speed during welding
- Modifying the joint design
- Ensuring that the surfaces to be joined fit together properly.

Weld Profile. *Weld profile* is important not only because of its effects on the strength and appearance of the weld, but also because it can indicate incomplete fusion or the presence of slag inclusions in multiple-layer welds.

- **Underfilling** results when the joint is not filled with the proper amount of weld metal (Fig. 30.22a).
- **Undercutting** results from melting away of the base metal and the subsequent development of a groove in the shape of a sharp recess or notch (Fig. 30.22b). If it is deep or sharp, an undercut can act as a stress raiser, and thus reduce the fatigue strength of the joint and lead to premature failure.

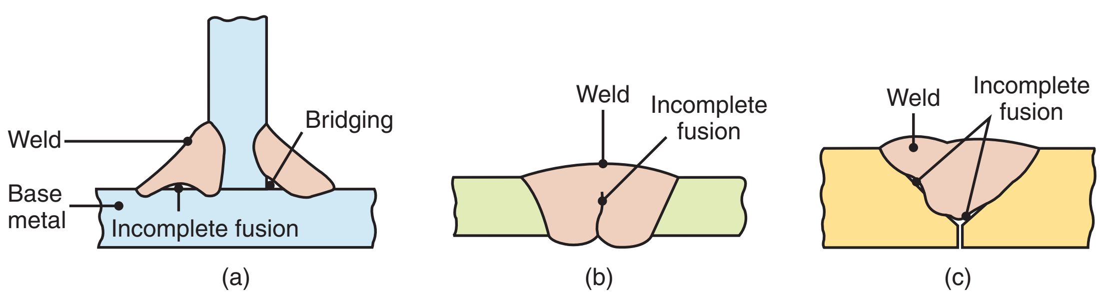


Figure 30.21: Examples of various discontinuities in fusion welds.

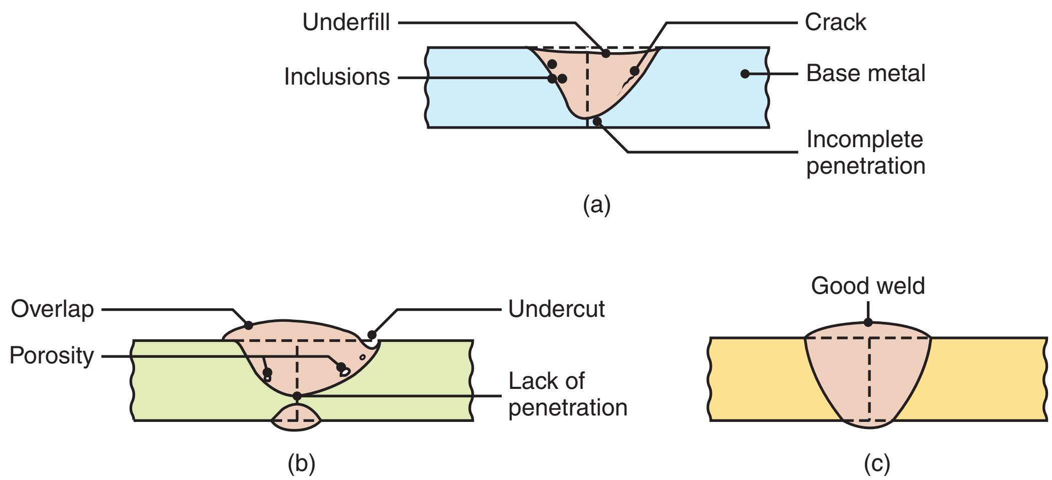


Figure 30.22: Examples of various defects in fusion welds.

- **Overlap** is a surface discontinuity (Fig. 30.22b), usually caused by poor welding practice or by selection of improper materials. Figure 30.22c shows a weld that would be considered to be good.

Cracks. *Cracks* may develop at various locations and directions in the weld area. Typical types of cracks are longitudinal, transverse, crater, underbead, and toe cracks (see Fig. 30.23). Cracks generally result from a combination of the following factors:

- Temperature gradients, causing thermal stresses in the weld zone
- Variations in the composition of the weld zone, causing different rates of contraction during cooling

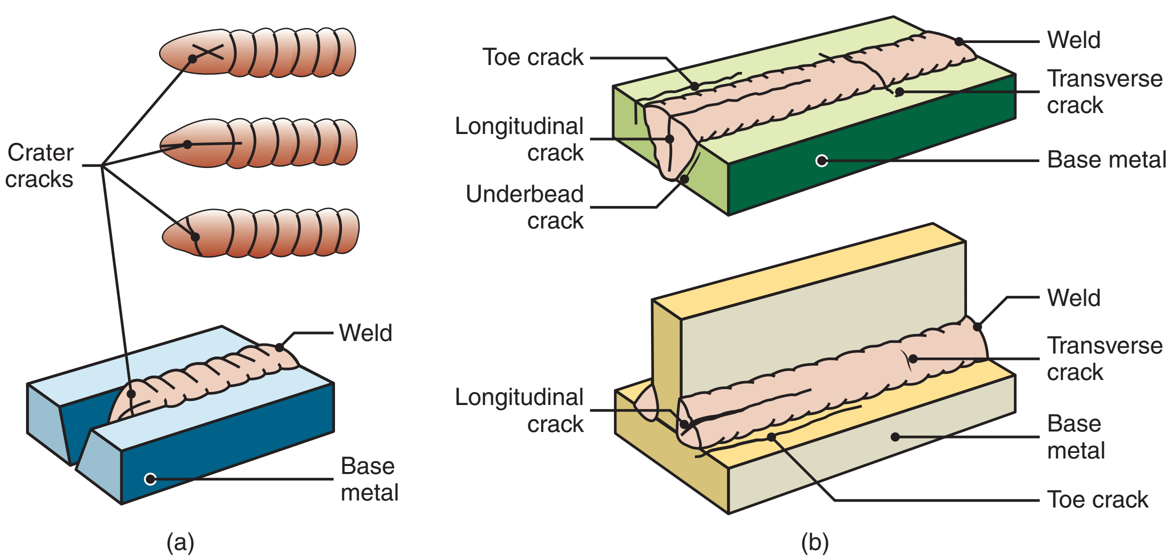


Figure 30.23: Types of cracks developed in welded joints. The cracks are caused by thermal stresses, similar to the development of hot tears in castings, as shown in Fig. 10.14.

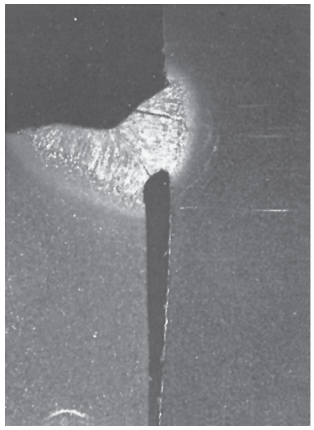


Figure 30.24: Crack in a weld bead. The two welded components were not allowed to contract freely after the weld was completed.

- Embrittlement of grain boundaries (Section 1.5.2), caused by the segregation of such elements as sulfur to the grain boundaries and occurring when the solid–liquid boundary moves as the weld metal begins to solidify
- Hydrogen embrittlement (Section 2.10.2)
- Inability of the weld metal to contract during cooling (Fig. 30.24), a situation similar to *hot tears* that develop in castings (Fig. 10.14) and is related to excessive restraint of the workpiece during the welding operation.

Cracks also are classified as **hot cracks** (developed while the joint is still at elevated temperatures) and **cold cracks** (after the weld metal has cooled). The basic crack-prevention measures in welding are:

- Modify the joint design to minimize stresses developed from shrinkage during cooling
- Change the parameters, procedures, and welding sequence
- Preheat the components to be welded
- Avoid rapid cooling of the welded joint.

Lamellar Tears. In describing the anisotropy of plastically deformed metals in Section 1.5, it was stated that the workpiece is weaker when tested in its thickness direction because of the alignment of non-metallic impurities and inclusions (*stringers*). This condition is observed particularly in rolled plates and structural shapes. In welding such components, *lamellar tears* may develop, because of shrinkage of the restrained components of the structure during cooling. Tears can be avoided by providing for shrinkage of the members or by modifying the joint design to make the weld bead penetrate the weaker component more deeply.

Surface Damage. Some of the hot metal may spatter during welding and be deposited, as small droplets, on adjacent surfaces. In arc-welding processes, the electrode may inadvertently touch the parts being welded at places other than the weld zone, called **arc strikes**. The associated surface discontinuities may be objectionable for reasons of appearance or in subsequent use or assembly of the welded structure. If severe, these discontinuities may adversely affect the properties of the welded structure, particularly notch sensitive metals. Using proper welding techniques and procedures is important in avoiding surface damage.

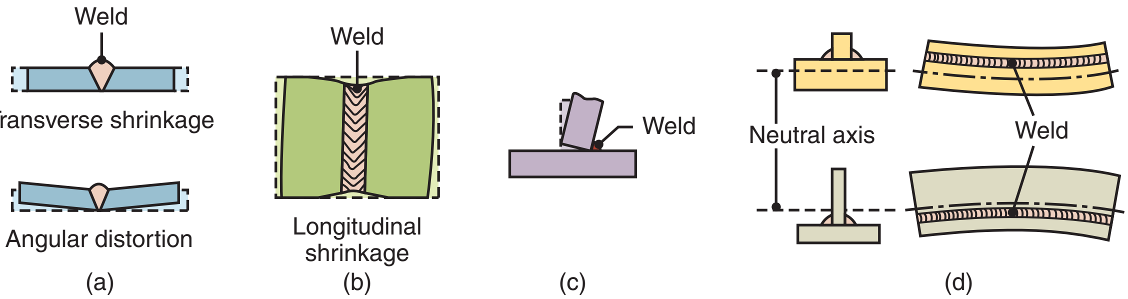


Figure 30.25: Distortion of parts after welding. Distortion is caused by differential thermal expansion and contraction of different regions of the welded assembly.

Residual Stresses. Because of localized heating and cooling during welding, the expansion and contraction of the weld area causes *residual stresses* (see also Section 2.11). Residual stresses can lead to the following defects:

- Distortion, warping, and buckling of the welded parts (Fig. 30.25)
- Stress-corrosion cracking (Section 2.10.2)
- Additional distortion if a portion of the welded structure is subsequently removed, such as by machining, drilling, or sawing
- Reduced fatigue life of the welded structure.

The type and distribution of residual stresses developed in welds is best described by referring to Fig. 30.26a. When two plates are being welded, a long narrow zone is subjected to elevated temperatures, while the plates, as a whole, are essentially at ambient temperature. After the weld is completed and as time elapses, heat from the weld zone dissipates laterally into the plates, while the weld area begins to cool. The plates then begin to expand longitudinally, while the welded length begins to contract (Fig. 30.25).

If the plate is not constrained, it will warp, as shown in Fig. 30.25a. If, however, the plate is not allowed to warp, it will develop residual stresses, which typically are distributed throughout the material (see stresses shown in Fig. 30.26b). Note that the magnitude of the compressive residual stresses in the

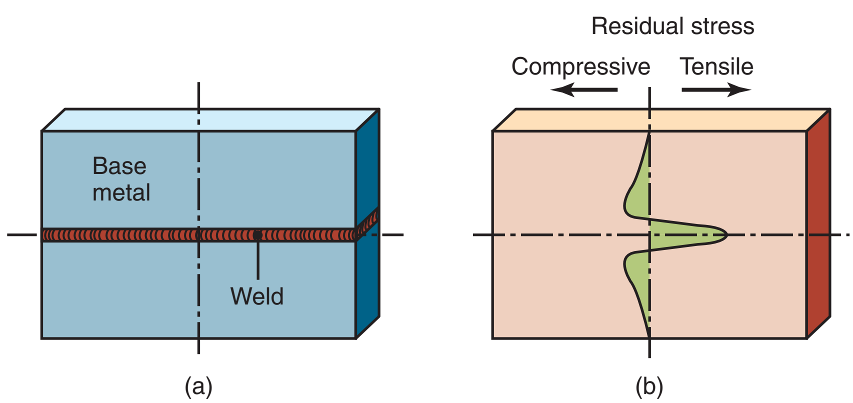


Figure 30.26: Residual stresses developed in (a) a straight-butt joint; note that the residual stresses shown in (b) must be balanced internally (see also Fig. 2.30.)

plates diminishes to zero at the top and bottom surfaces of the welded plate. Note also that because no external forces are acting on the welded plates, the tensile and compressive forces represented by these residual stresses must balance each other.

The sequence of events leading to the distortion of a simple tubular welded structure are shown in Fig. 30.27. Prior to welding, the structure is stress free, as shown in Fig. 30.27a, and it may be sufficiently rigid; some fixturing may be present to support the structure as part of a larger assembly if necessary. During welding, the molten metal fills the gap between the surfaces to be joined and forms a weld bead. As the weld begins to solidify, both the weld bead and the surrounding material begin to cool down to room temperature. As they cool, they would contract but are constrained by the rest of the weldment; as a result, the part distorts (Fig. 30.27c) and residual stresses develop.

The residual stresses produce the deformation shown in Fig. 30.27c and put the weld and the heat-affected zone into a state of residual tension, which is not desirable for fatigue performance. In general, HAZ is less fatigue resistant than the base metal. Because the residual stresses developed can be harmful, it is not unusual to stress relieve welds in highly stressed or fatigue-susceptible applications (see below). Recall that the weld itself may have porosity (see Fig. 30.22b), which also can act as a stress raiser and lead to fatigue crack growth.

In complex welded structures, residual-stress distributions are three dimensional, and difficult to analyze. Note that the two plates shown in Fig. 30.26 were not restrained from movement; in other words, the plates were not an integral part of a larger structure. If, however, they were restrained, reaction stresses would develop, because the plates are not free to expand or contract, a situation that arises particularly in structures with high stiffness.

Stress Relieving of Welds. Effects that residual stresses can cause, such as distortion, buckling, and cracking, can be reduced by **preheating** the base metal or the parts to be welded. Preheating reduces distortion by reducing the cooling rate following welding and the level of thermal stresses developed, by lowering the elastic modulus; this technique also reduces shrinkage and possible cracking of the joint.

For optimum results, preheating temperatures and cooling rates must be controlled in order to maintain acceptable strength and toughness of welded structures. Workpieces may be heated in several ways, including (a) in a furnace, (b) electrically, either resistively or inductively, or (c) by radiant lamps or hot-air blast, especially for thin sections. The temperature and time required for stress relieving depend on the type of material and on the magnitude of the stresses developed.

Other methods of stress relieving include *peening, hammering,* or *surface rolling* (Section 34.2) of the weld-bead area. These techniques induce compressive residual stresses, which, in turn, lower or eliminate tensile residual stresses in the weld. For multilayer welds, the first and last layers should not be peened, to protect them against possible peening damage on the surface.

Residual stresses also can be relieved or reduced by *plastically* deforming the structure itself by a small amount. This technique can be used in welded pressure vessels, by pressurizing the vessels internally, called

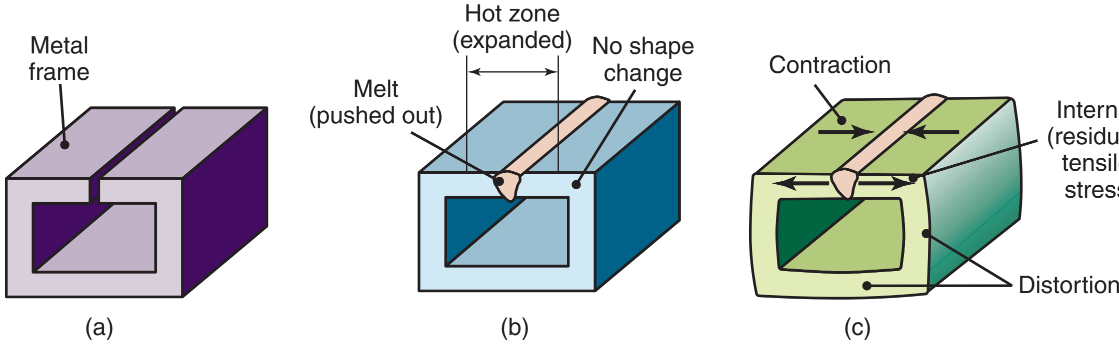


Figure 30.27: Distortion of a welded structure. *Source:* After J.A. Schey.

proof stressing. In order to reduce the possibility of sudden fracture under high internal pressure, the weld must be made properly and must be free of notches and discontinuities, which can act as stress raisers.

In addition to being preheated for stress relieving, welds may be heat treated by various techniques in order to modify other properties. These techniques include annealing, normalizing, quenching, and tempering of steels and solution treatment and aging of various alloys, as described in Chapter 4.

30.9.2 Weldability

The *weldability* of a metal is generally defined as its capacity to be welded into a specific structure that has certain properties and characteristics and will satisfactorily meet service requirements. Weldability involves a large number of variables, and thus generalizations are difficult. Recall that material characteristics, such as alloying elements, impurities, inclusions, grain structure, and processing history, of both the base metal and the filler metal, are all important. For example, weldability of steels decreases with increasing carbon content, because of martensite formation (see Section 4.7) and thus reduces the strength of the weld. Coated steel sheets (Chapter 34) also present various challenges in welding, depending on the type and thickness of the coating.

Because of the effects of melting and solidification and of the associated microstructural changes, a thorough consideration of the phase diagram and the response of the metal or alloy to sustained elevated temperatures is essential. Also influencing weldability are mechanical and physical properties: strength, toughness, ductility, notch sensitivity, elastic modulus, specific heat, melting point, thermal expansion, surface-tension characteristics of the molten metal, and corrosion resistance.

Preparation of *surfaces* for welding is important, as are the nature and properties of surface-oxide films and of adsorbed gases (see also Section 33.2). The specific welding process employed significantly affects the temperatures developed and their distribution in the weld zone. Other factors that affect weldability are shielding gases, fluxes, moisture content of the coatings on electrodes, welding speed, welding position, cooling rate, and level of preheating, as well as such post-welding techniques as stress relieving and heat treating.

Weldability of Ferrous Materials:

- *Plain-carbon steels:* Generally excellent for low-carbon steels, fair to good for medium-carbon steels, and poor for high-carbon steels.
- *Low-alloy steels:* Similar to medium-carbon steels.
- *High-alloy steels:* Generally good under well-controlled conditions.
- *Stainless steels:* Generally weldable by various processes.
- *Cast irons:* Generally weldable, although their weldability varies greatly.

Weldability of Nonferrous Materials:

- *Aluminum alloys:* Weldable at a high rate of heat input; an inert shielding gas and lack of moisture are important. Aluminum alloys containing zinc or copper generally are considered unweldable.
- *Copper alloys:* Depending on composition, generally weldable at a high rate of heat input; an inert shielding gas and lack of moisture are important.
- *Magnesium alloys:* Weldable using a protective shielding gas and fluxes.
- *Nickel alloys:* Similar to stainless steels; lack of sulfur is undesirable.
- *Titanium alloys:* Weldable with proper use of shielding gases.

- *Tantalum:* Similar to titanium.
- *Tungsten:* Weldable under well-controlled conditions.
- *Molybdenum:* Similar to tungsten.
- *Niobium (columbium):* Good weldability.

30.9.3 Testing of Welds

Several standardized tests and test procedures have been established and are available from organizations such as the American Society for Testing and Materials (ASTM), the American Welding Society (AWS), the American Society of Mechanical Engineers (ASME), the American Society of Civil Engineers (ASCE), and various federal agencies.

Welded joints may be tested either *destructively* or *nondestructively* (see also Sections 36.10 and 36.11). Each technique has certain capabilities and limitations, as well as process parameter sensitivity, reliability, and requirements for special equipment and operator skill.

Destructive Testing Techniques:

- **Tension test.** *Longitudinal and transverse tension tests* are performed on specimens removed from actual welded joints and from the weld-metal area. Stress–strain curves are then developed, using the procedures described in Section 2.2. These curves indicate the yield strength, ultimate tensile strength, and ductility of the welded joint (elongation and reduction of area) in different locations and directions.
- **Tension-shear test.** The specimens in the *tension-shear test* (Fig. 30.28a and b) are prepared to simulate conditions to which actual welded joints are subjected. The specimens are subjected to tension, so that the shear strength of the weld metal and the location of fracture can be determined.
- **Bend test.** Several bend tests have been developed to determine the ductility and strength of welded joints. In one common test, the welded specimen is bent around a fixture (*wraparound bend test*, Fig. 30.28c). In another, the specimens are tested in *three-point transverse bending* (Fig. 30.28d; see also Fig. 2.11a). These tests help to determine the relative ductility and strength of welded joints.
- **Fracture toughness test.** This test commonly utilizes impact testing techniques (Section 2.9). *Charpy V-notch* specimens are first prepared and tested for toughness. In the *drop-weight test*, the energy is supplied by a falling weight.
- **Creep and corrosion tests.** *Creep tests* (Section 2.8) are essential in determining the behavior of welded joints and structures subjected to elevated temperatures. Welded joints also may be tested for their resistance to *corrosion* (Section 3.8); because of the difference in composition and microstructure, *preferential corrosion* may take place in the weld zone.

Nondestructive Testing Techniques. Welded structures often have to be tested *nondestructively* (Section 36.10), particularly for critical applications in which weld failure can be catastrophic, such as in pressure vessels, load-bearing structural members, and power plants. Nondestructive testing techniques for welded joints generally consist of the following methods:

- Visual
- Radiographic (X-rays)
- Magnetic-particle
- Liquid-penetrant
- Ultrasonic.

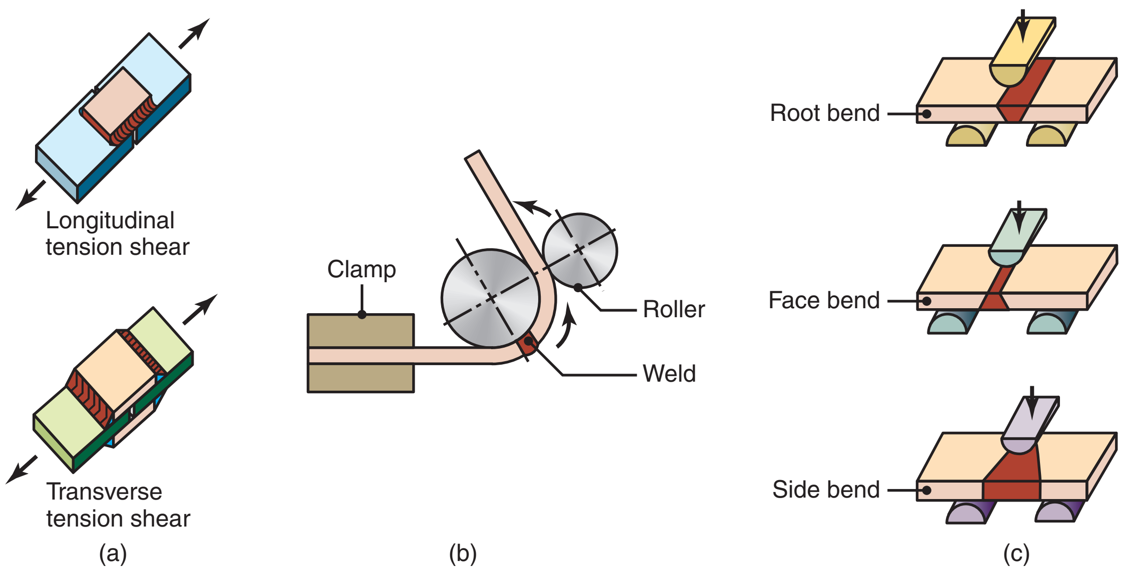


Figure 30.28: (a) Specimens for longitudinal tension-shear testing and for transfer tension-shear testing. (b) Wraparound bend-test method. (c) Three-point transverse bending of welded specimens.

As an example of another nondestructive method, testing for hardness distribution (see Section 2.6 and Figs. 16.3 and 30.20) in the weld zone also would be a useful indicator of weld strength and microstructural changes.

30.10 Joint Design and Process Selection

In describing individual welding processes, several examples were given regarding the types of welds and joints produced and their applications in various consumer and industrial products. Typical types of joints produced by welding, together with their terminology, are given in Fig. 30.29. Standardized symbols commonly used in engineering to describe the types of welds are shown in Fig. 30.30. These symbols identify the type of weld, groove design, weld size and length, welding process, sequence of operations, and various other essential information.

General design guidelines for welding are given in Fig. 30.31. Various other types of joint design are given in Chapters 31 and 32. Important design guidelines are summarized below.

- Product design should minimize the number of joints because, unless automated, welding can be time consuming and costly.
- Weld locations should be selected so as to avoid excessive local stresses or stress concentrations as well as for better appearance.
- Weld location should be selected so as not to interfere with any subsequent processing of the joined components or with their intended uses.
- The need for edge preparation should be minimized or avoided.
- Weld-bead size should be as small as possible while maintaining joint strength, in order to conserve weld metal and for better appearance.

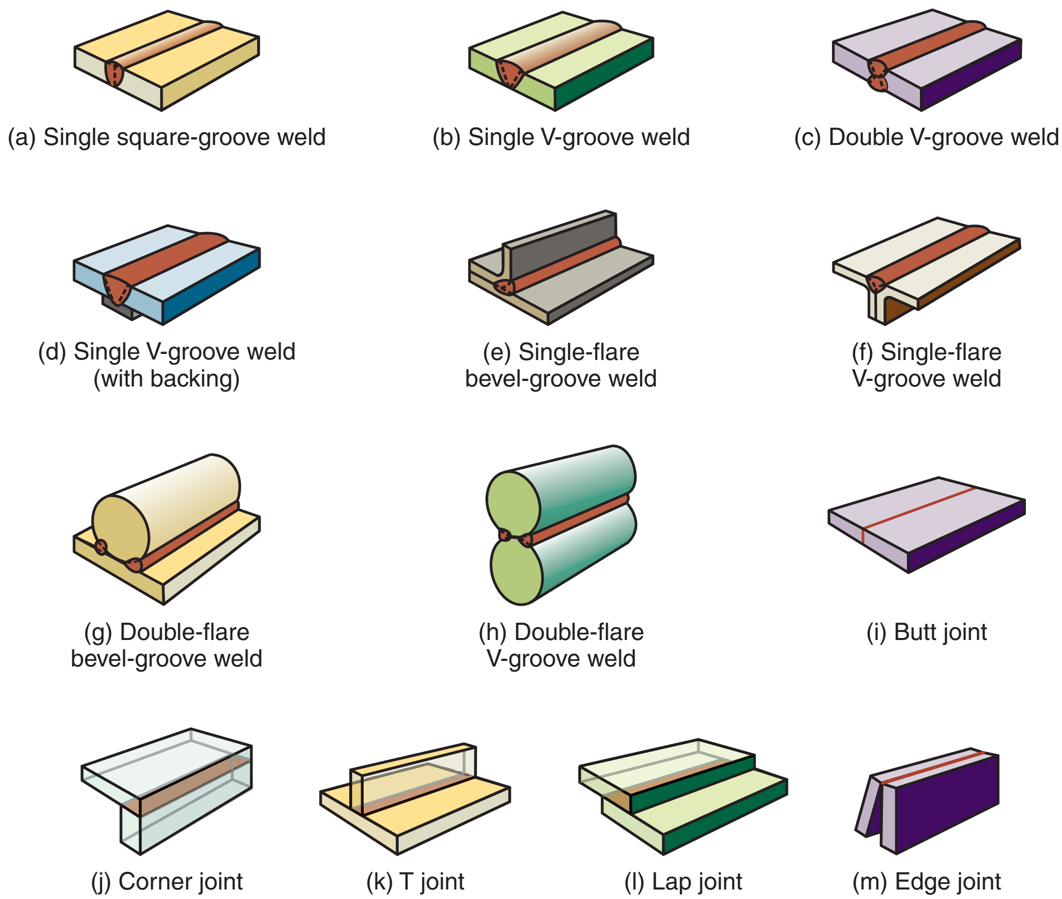


Figure 30.29: Examples of welded joints and their terminology.

Welding Process Selection. In addition to considering welding process characteristics, capabilities, and material considerations described thus far, selection of a weld joint and an appropriate process involve the following considerations (see also Chapters 31 and 32):

- Configuration of the parts to be joined, joint design, thickness and size of components, and number of joints required
- Methods used in making the components to be joined
- Types of materials involved
- Location, accessibility, and ease of joining
- Weld application and service requirements, including type of loading, stresses generated, and environment
- Effects of distortion, warping, appearance, discoloration, and service
- Costs involved in edge preparation, joining, and post-processing, including machining, grinding, and finishing operations
- Costs of equipment, materials, labor, and skills required of the whole operation.

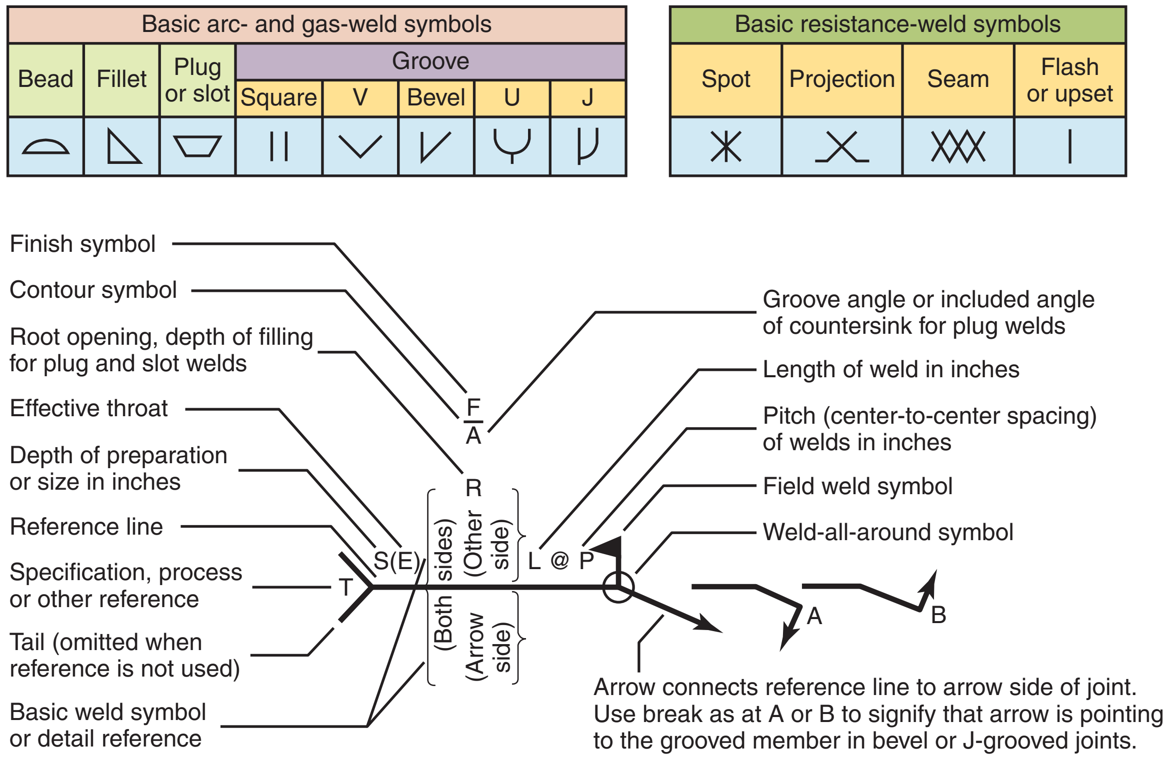


Figure 30.30: Standard identification and symbols for welds.

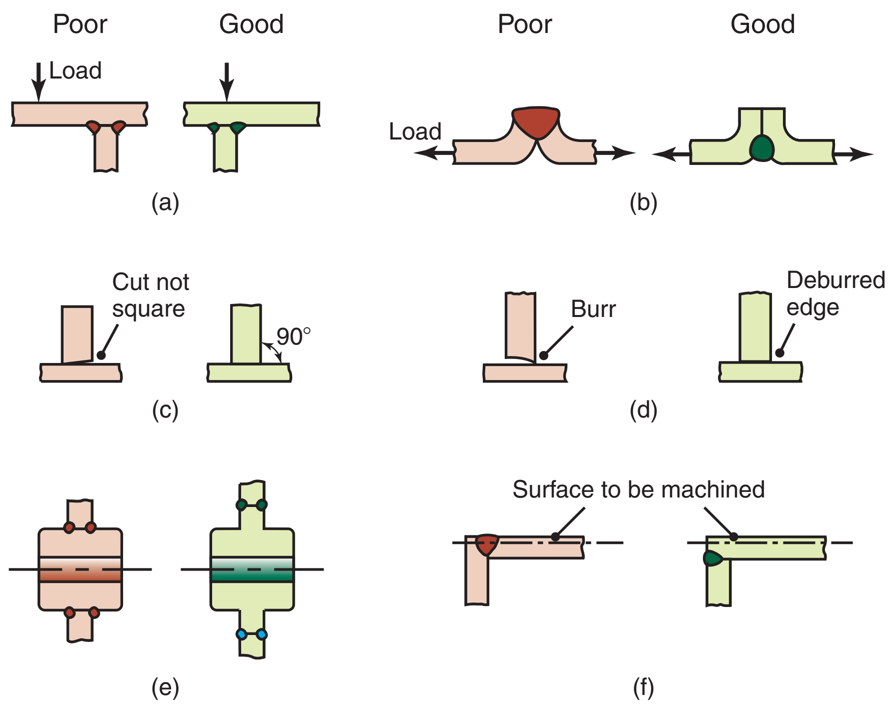


Figure 30.31: Some design guidelines for welds. *Source:* After J.G. Bralla.

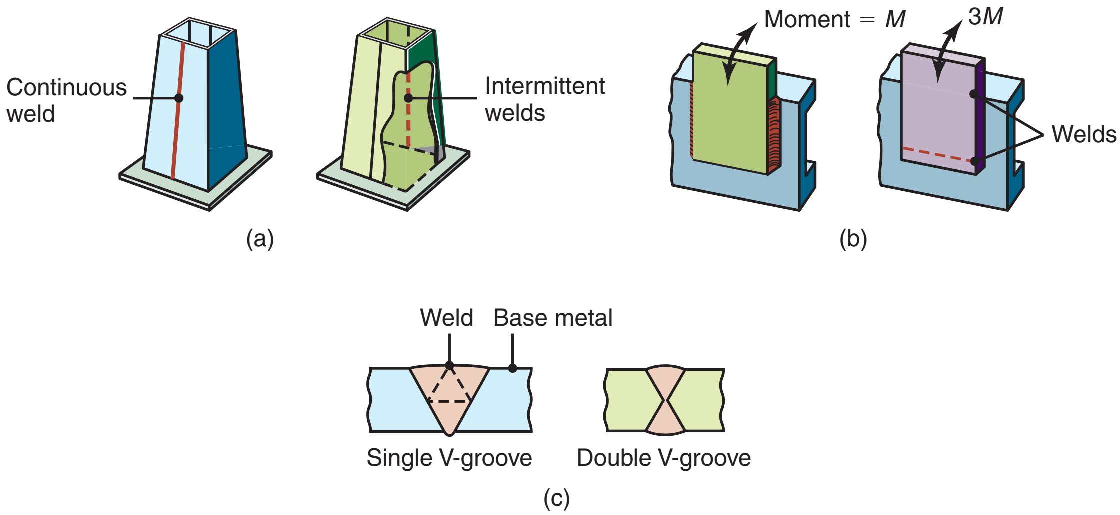


Figure 30.32: Examples of weld designs used in Example 30.3.

Example 30.3 Weld Design Selection

Three different types of weld designs are shown in Fig. 30.32. The two vertical joints in Fig. 30.32a can be welded either externally or internally. Note that full-length external welding will take considerable time and will require more weld material than the alternative design, which consists of intermittent internal welds. Moreover, in the alternative method the appearance of the structure is improved and distortion is reduced.

In Fig. 30.32b, it can be shown that the design on the right can carry three times the moment M of the one on the left. Note also that both designs require the same amount of weld metal and welding time. In Fig. 30.32c, the weld on the left requires about twice the amount of weld material than does the design on the right. Moreover, because more material must now be machined, the design on the left will require more time for edge preparation, and more base metal will be wasted.

Summary

- Oxyfuel–gas, arc, and high-energy-beam welding are among the most commonly used joining operations. Gas welding uses chemical energy to supply the necessary heat; arc and high-energy-beam welding use electrical energy.

- In all the processes described, heat is used to bring the joint being welded to a liquid state. Shielding gases are used to protect the molten-weld pool and the weld area against oxidation. Filler metals may or may not be used in oxyfuel–gas and arc welding.

- Selection of a welding process for a particular operation depends on the workpiece material, its thickness and size, its shape complexity, the type of joint required, the strength required, and the change in product appearance caused by welding.

- A variety of welding equipment is available, now mostly computer and robot controlled, with programmable features.

- Cutting of metals also can be done by the processes based on oxyfuel–gas and arc welding. The highest temperatures for cutting are obtained by plasma-arc cutting.

- The welded joint consists of solidified metal and a heat-affected zone; each has a wide variation in their microstructure and properties, depending on the metals joined and on the filler metals. The metallurgy of the welded joint is important in all welding processes, because it determines the strength, toughness, and quality of the joint.

- Discontinuities, such as porosity, inclusions, incomplete welds, tears, surface damage, and cracks, can develop in the weld zone. Residual stresses and relieving them are important considerations.

- Weldability of metals and alloys depends greatly on their composition, mechanical and physical properties, type of welding operation and process parameters employed, and the control of welding parameters.

- General guidelines are available for the selection of suitable and economical methods for a particular welding application.

Key Terms

Arc cutting
Arc welding
Atomic-hydrogen welding
Base metal
Carburizing flame
Coated electrode
Consumable electrode
Discontinuities
Drag lines
Electrode
Electrogas welding
Electron-beam welding
Electroslag welding
Filler metal
Flux
Flux-cored arc welding
Fusion welding
Gas metal-arc welding
Gas tungsten-arc welding
Heat-affected zone
Inclusions
Kerf
Keyhole technique
Laser-beam welding
Laser GMAW welding
Neutral flame
Nonconsumable electrode
Oxidizing flame
Oxyfuel–gas cutting
Oxyfuel–gas welding
Plasma-arc welding
Polarity
Porosity
Reducing flame
Residual stresses
Shielded metal-arc welding
Slag
Stick welding
Submerged-arc welding
Tears
Thermit welding
Weld profile
Weld metal
Weldability
Welding gun

Bibliography

ASM Handbook, Vol. 6: **Welding, Brazing, and Soldering**, ASM International, 1993.
ASM Handbook, Vol. 6A: **Welding Fundamentals and Processes**, ASM International, 2011.
Bohnart, E,R., **Welding: Principles and Practices**, 5th ed., McGraw-Hill, 2017.
Campbell, F.C., **Joining**, ASM International, 2011.
Katayama, S., (ed.), **Handbook of Laser Welding Technologies**, Woodhead Publishing, 2013.
Kou, S., **Welding Metallurgy**, 2nd ed., Wiley-Interscience, 2002.
Lippold, J.C., **Welding Metallurgy and Weldability**, Wiley, 2014.
Modern Welding Technology, 11th ed., Goodheart-Willcox, 2012.
Phillips, D.H., **Welding Engineering: An Introduction**, Wiley, 2016.
Welding Handbook, various volumes and editions, American Welding Society.
Welding Inspection Handbook, 4th ed., American Welding Society, 2015.
Weman, K., **Welding Processes Handbook**, 2nd ed., Woodhead Publishing, 2011.
Yadav, K.S., **Advanced Welding Technology**, Rajsons Publishing, 2018.

Review Questions

30.1. Describe fusion as it relates to welding operations.

30.2. Explain the features of neutralizing, reducing, and oxidizing flames. Why is a reducing flame so called?

30.3. What is stick welding?

30.4. Explain the basic principles of arc-welding processes.

30.5. Why is shielded metal-arc welding a commonly used process? Why is it also called stick welding?

30.6. What keeps the weld bead on a steel surface from oxidizing (rusting) during welding?

30.7. Describe the functions and characteristics of electrodes. What functions do coatings have? How are electrodes classified?

30.8. What are the similarities and differences between consumable and nonconsumable electrodes?

30.9. What properties are useful for a shielding gas?

30.10. What are the advantages to thermite welding?

30.11. Explain where the energy is obtained in thermite welding.

30.12. Explain how cutting takes place when an oxyfuel–gas torch is used. How is underwater cutting done?

30.13. What is the purpose of flux? Why is it not needed in gas tungsten-arc welding?

30.14. What is meant by weld quality? Discuss the factors that influence it.

30.15. How is weldability defined?

30.16. Why are welding electrodes generally coated?

30.17. Describe the common types of discontinuities in welded joints.

30.18. What types of destructive tests are performed on welded joints?

30.19. Explain why hydrogen welding can be used to weld tungsten without melting the tungsten electrode.

30.20. What materials can be welded by Laser SMAW hybrid welding?

Qualitative Problems

30.21. Explain the reasons that so many different welding processes have been developed over the years.

30.22. It has been noted that heat transfer in gas-metal arc welding is higher than in shielded-metal arc welding. Explain why this would be the case. Which process would lead to more heat-affected zone cracking in hardened steels?

30.23. Explain why some joints may have to be preheated prior to welding.

30.24. Describe the role of filler metals in welding.

30.25. List the processes that can be performed with two electrodes. What are the advantages in using two electrodes?

30.26. What is the effect of the thermal conductivity of the workpiece on kerf width in oxyfuel–gas cutting? Explain.

30.27. Describe the differences between oxyfuel–gas cutting of ferrous and of nonferrous alloys. Which properties are significant?

30.28. Could you use oxyfuel–gas cutting for a stack of sheet metals? (*Note:* For stack cutting, see Fig. 24.28e.) Explain.

30.29. What are the advantages of electron-beam and laser-beam welding compared with arc welding?

30.30. Describe the methods by which discontinuities in welding can be avoided.

30.31. Explain the significance of the stiffness of the components being welded on both weld quality and part shape.

30.32. Comment on the factors that influence the size of the two weld beads shown in Fig. 30.15.

30.33. Which of the processes described in this chapter are not portable? Can they be made so? Explain.

30.34. Thermit welding is commonly used for welding railroad rails. List the reasons that make thermit welding attractive for this application. Review your list and create a list of products that would be suitable for thermit welding, and then identify any difficulties you would expect in applying thermit welding to that application.

30.35. Describe your observations concerning the contents of Table 30.1.

30.36. What determines whether a certain welding process can be used for workpieces in horizontal, vertical, or upside-down positions or, for that matter, in any position (see Table 30.1)? Explain and give examples of appropriate applications.

30.37. Comment on the factors involved in electrode selection in arc-welding processes.

30.38. In Table 30.1, the column on the distortion of welded components is ordered from lowest distortion to highest. Explain why the degree of distortion varies among different welding processes.

30.39. Explain the significance of residual stresses in welded structures.

30.40. Rank the processes described in this chapter in terms of (a) cost and (b) weld quality.

30.41. Must the filler metal be made of the same composition as the base metal that is to be welded? Explain.

30.42. What is weld spatter? What are its sources? How can spatter be controlled? Explain.

30.43. Describe your observations concerning Fig. 30.20.

30.44. If the materials to be welded are preheated, is the likelihood for porosity increased or decreased? Explain.

30.45. Discuss the need for and role of fixtures in holding workpieces in the welding operations described in this chapter.

30.46. Why is the quality of welds produced by submerged arc welding very good?

30.47. Explain why the electroslag welding process is suitable for thick plates and heavy structural sections.

30.48. Explain why the grains in Fig. 30.20c grow in the particular directions shown.

Quantitative Problems

30.49. Plot the hardness in Fig. 30.20d as a function of the distance from the top surface, and discuss your observations.

30.50. A welding operation will take place on carbon steel. The desired welding speed is around 20 mm/s. If an arc-welding power supply is used with a voltage of 12 V, what current is needed if the weld width is to be 5 mm?

30.51. In Fig. 30.26b, assume that most of the top portion of the top piece is cut horizontally with a sharp saw. The residual stresses will now be disturbed and the part will change its shape, as was described in Section 2.11. For this case, how do you think the part will distort: curved downward or upward? Explain. (See also Fig. 2.30d.)

30.52. A welding operation takes place on an aluminum-alloy plate. A pipe 50 mm in diameter, with a 5 mm wall thickness and a 50 mm length, is butt welded onto an extruded L-section 150 mm by 150 mm by 5 mm thick, with a length of 1 m. If the weld zone in a gas tungsten arc welding process is approximately 10 mm wide, what would be the temperature increase of the entire structure due to the heat input from welding only? What if the process were an electron-beam welding operation, with a bead width of 2 mm? Assume that the electrode and aluminum alloy require 2.9 Joules to melt one cubic millimeter.

30.53. An arc welding operation is taking place on carbon steel. The desired welding speed is around 24 mm/sec. If the power supply is 12 V, what current is needed if the weld width is to be 6 mm?

30.54. In oxyacetylene, arc, and laser-beam cutting, the processes basically involve melting of the workpiece. If a 50 mm diameter hole is to be cut from a 250 mm diameter, 12 mm thick plate, plot the mean temperature rise in the blank as a function of kerf. Assume that one-half of the energy goes into the blank.

30.55. A submerged arc welding operations takes place on 10 mm thick stainless steel, producing a butt weld as shown in Fig. 30.29c. The weld geometry can be approximated as a trapezoid with 15 mm and 10 mm as the top and bottom dimensions, respectively. If the voltage provided is 40 V at 400 A, estimate the welding speed if a stainless steel filler wire is used.

30.56. 6061 aluminum plates with a 2.5 mm thickness are to be butt-welded by GMAW using a 1 mm diameter electrode. The applied voltage is 22 V, the current is 125A, and the arc travel speed is 16 mm/s. Calculate the power, the deposition rate of electrode material, and the required electrode feed rate.

30.57. Assume that you are asked to give a quiz to students on the contents of this chapter. Prepare three quantitative problems and three qualitative questions, and supply the answers.

Synthesis, Design, and Projects

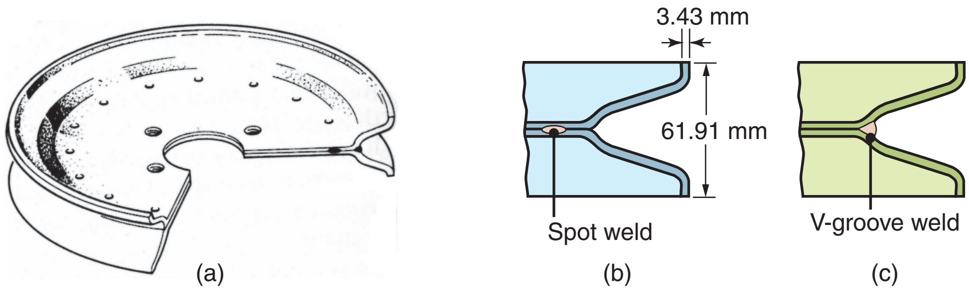
30.58. Comment on workpiece size and shape limitations for each of the processes described in this chapter.

30.59. Arc blow is a phenomenon where the magnetic field induced by the welding current passing through the electrode and workpiece in shielded metal arc welding interacts with the arc and causes severe weld splatter. Identify the variables that you feel are important in arc blow. When arc blow is a problem, would you recommend minimizing it by using AC or DC power?

30.60. Review the types of welded joints shown in Fig. 30.29 and give an application for each.

30.61. Comment on the design guidelines given in various sections of this chapter.

30.62. The accompanying figure shows a metal sheave that consists of two matching pieces of hot-rolled, low-carbon-steel sheets. These two pieces can be joined either by spot welding or by V-groove welding. Discuss the advantages and limitations of each process for this application.



30.63. You are asked to inspect a welded structure for a critical engineering application. Describe the procedure that you would follow in order to determine the safety of the structure.

30.64. Discuss the need for, and the role of, work-holding devices in the welding operations described in this chapter.

30.65. Make a list of welding processes that are suitable for producing (a) butt joints, where the weld is in the form of a line or line segment, (b) spot welds, and (c) both butt joints and spot welds. Comment on your observations.

30.66. Explain the factors that contribute to the differences in properties across a welded joint.

30.67. Explain why preheating the components to be welded is effective in reducing the likelihood of developing cracks.

30.68. Review the poor and good joint designs shown in Fig. 30.31, and explain why they are labeled so.

30.69. In building large ships, there is a need to weld thick and large sections of steel together to form a hull. Consider each of the welding operations discussed in this chapter, and list the benefits and drawbacks of that particular joining operation for this application.

30.70. Inspect various parts and components in (a) an automobile, (b) a major appliance, and (c) kitchen utensils, and explain which, if any, of the processes described in this chapter has been used in joining them.

30.71. Comment on whether there are common factors that affect the weldability, castability, formability, and machinability of metals, as described in various chapters of this book. Explain with appropriate examples.

30.72. If you find a flaw in a welded joint during inspection, how would you go about determining whether or not the flaw is significant?

30.73. Lattice booms for cranes are constructed from extruded cross-sections (see Fig. 15.2) that are welded together. Any warpage that causes such a boom to deviate from straightness will severely reduce its lifting capacity. Conduct a literature search on the approaches used to minimize distortion due to welding and how to correct it, specifically in the construction of lattice booms.

30.74. A common practice in repairing expensive broken or worn parts (such as those that may occur when a fragment is broken from a forging) is to fill the area with layers of weld beads and then to machine the part back to its original dimensions. Make a list of the precautions that you would suggest to someone who uses this approach.

30.75. Consider a butt joint that is to be welded. Sketch the weld shape you would expect for (a) SMAW; (b) laser welding; and (c) Laser-SMAW hybrid welding. Indicate the size and shape of the heat-affected zone you would expect. Comment on your observations.

30.76. Prepare a table listing the processes described in this chapter and providing, for each process, the range of welding speeds as a function of workpiece material and thickness.

30.77. Make an outline of the general guidelines for safety in welding operations described in this chapter. For each of the operations, prepare a poster which effectively and concisely gives specific instructions for safe practices in welding (or cutting). Review the various publications of the National Safety Council and other similar organizations.

30.78. Describe the reasons that fatigue failure generally occurs in the heat-affected zone of welds instead of through the weld bead itself.

Chapter 31

Solid-state Welding Processes

- This chapter describes a family of joining processes in which the parts welded do not undergo a phase change; if heat is involved, it is generated internally.
- The chapter begins with a description of cold welding, followed by ultrasonic welding and the friction-welding processes.
- Resistance welding is then described, followed by explosion welding and diffusion bonding. These three processes have unique capabilities and applications, suitable for a wide variety of materials and can be automated for large-scale production.

- The chapter then examines the capabilities of diffusion bonding, and the joining processes that are combined with superplastic forming.
- The final topic described concerns economic considerations in welding.

31.1 Introduction

This chapter describes **solid-state welding** processes, in which joining takes place without fusion at the interface of the two parts being welded. Unlike the fusion-welding processes described in Chapter 30, in solid-state welding no liquid or molten phase is required for joining. The principle of solid-state welding is demonstrated best by the following example: If two clean metal surfaces are brought into close contact with each other under sufficient pressure, they form a bond and produce a joint. For a strong bond, it is essential that the interface be free of contaminants, such as oxide films, residues, metalworking fluids, and even adsorbed layers of gas.

Solid-state bonding involves one or more of the following parameters:

- **Heat:** Applying external heat increases *diffusion* (the transfer of atoms across an interface) and improves the strength of the weld between the two surfaces being joined, as occurs in *diffusion bonding*. Heat may be generated (a) internally, by friction, as utilized in *friction welding*; (b) through electrical-resistance heating, as in resistance-welding, such as *spot welding*; and (c) externally, by induction heating (as in *butt-welding* of tubes).
- **Pressure:** The higher the contact pressure, the stronger is the interface, as in *roll bonding* and *explosion welding*, where plastic deformation occurs. Pressure and heat may be combined, as in *flash welding*, *stud welding*, and *resistance projection welding*.
- **Relative interfacial movements:** When sliding of the contacting surfaces, called *faying surfaces*, occurs (as in *ultrasonic welding*), even very small amplitudes will disturb the interface, breaking up any oxide films present, and generating new clean surfaces, thus improving weld strength.

Most joining processes are now automated, with *robotics, vision systems, sensors,* and *adaptive* and *computer controls* (described in Part VIII). The reasons are to reduce costs (Section 31.8) and increase consistency of operation, reliability of weld quality, and higher productivity.

31.2 Cold Welding and Roll Bonding

In *cold welding* (CW), pressure is applied to the workpieces through dies or rolls. Because of the *plastic deformation* involved, it is essential that at least one, but preferably both, of the mating parts be sufficiently ductile. Cold welding is usually performed on nonferrous metals or on soft iron with little, if any, carbon content. Prior to welding, the interface is first degreased, wire brushed, and wiped off to remove oxide.

During joining of two *dissimilar* metals that are *mutually soluble*, brittle *intermetallic compounds* may form (Section 4.2.2); these will produce a weak and brittle joint. An example of weak bonding of aluminum and steel. The best bond strength is obtained with two *similar* materials.

Roll Bonding. The pressure required for welding can be applied through a pair of rolls (Fig. 31.1), called *roll bonding* or *roll welding* (ROW). Developed in the 1960s, roll bonding is used for making U.S. coins (see Example 31.1). Surface preparation is important for good interfacial strength. The operation also can be carried out at elevated temperatures (*hot roll bonding*).

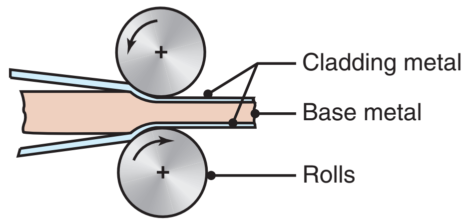


Figure 31.1: Schematic illustration of the roll-bonding, or cladding, process.

Typical examples of this process are the *cladding* of (a) pure aluminum over precipitation-hardened aluminum-alloy sheet (Alclad, a trade name), which has a corrosion-resistant surface with a strong inner core, typically used in the aircraft industry; (b) stainless steel over mild steel, for corrosion resistance; and (c) copper over steel, for coaxial cables, with steel for strength. A common application of roll bonding is in making *bimetallic strips*, for thermostats and similar control units, using two thin layers of materials with different thermal-expansion coefficients (see Table 3.1). Bonding of only selected areas of the interface can be achieved by depositing a *parting agent*, such as graphite or ceramic, called *stop-off* (Section 31.7).

Example 31.1 Roll Bonding of the U.S. Quarter

The technique used for making composite U.S. quarters is the roll bonding of (a) two outer layers of 75% Cu–25% Ni (cupronickel), where each layer is 1.2 mm thick and (b) with an inner layer of pure copper 5.1 mm thick. For good bond strength, the faying surfaces are first cleaned chemically and wire brushed. Then the strips are rolled, first to a thickness of 2.29 mm, then down to a thickness of 1.36 mm. The strips thus undergo a total reduction in thickness of 82%.

Because of the reduction in thickness, there is a major increase in the surface area between the layers. This extension in surface area under the high pressure applied by the rolls, combined with the *solid solubility* of nickel in copper (Section 4.2.1), produces a strong bond between the layers.

31.3 Ultrasonic Welding

In *ultrasonic welding* (USW), the faying surfaces of the two components are subjected to a normal force and oscillating shearing (tangential) stresses. The shearing stresses are applied by the tip of a **transducer** (Fig. 31.2a), which is similar to that used for ultrasonic machining (see Fig. 26.26a). The frequency of oscillation is generally in the range of 10–75 kHz, although a lower or higher frequency also can be employed. Proper coupling between the transducer and the tip (called **sonotrode**, from the words *sonic* and *electrode*, also called the *horn*), is important for efficient operation.

The shearing stresses cause plastic deformation at the interface of the two components, breaking up oxide films and contaminants, to allow for good contact and producing a strong solid-state bond. The temperature generated in the weld zone is usually one-third to one-half of the melting point (on the absolute scale) of the metals joined. Consequently, neither melting nor fusion takes place. In some situations, however, the temperature developed can be sufficiently high to cause metallurgical changes in the weld zone, thus affecting the strength of the bond.

The ultrasonic-welding process is versatile and reliable, and it can be used with a wide variety of metallic and nonmetallic materials, including dissimilar metals (as in bimetallic strips). It is used extensively for joining plastics (Section 32.6), packaging with metal foils, and lap welding of sheet, foil, and thin wire and in automotive and consumer electronics industries. The welding tip can be replaced with *rotating disks* (Fig. 31.2b) for seam welding of structures in which one component is sheet, foil, or polymer-woven material (a process similar to *resistance seam welding*, Section 31.5.2).

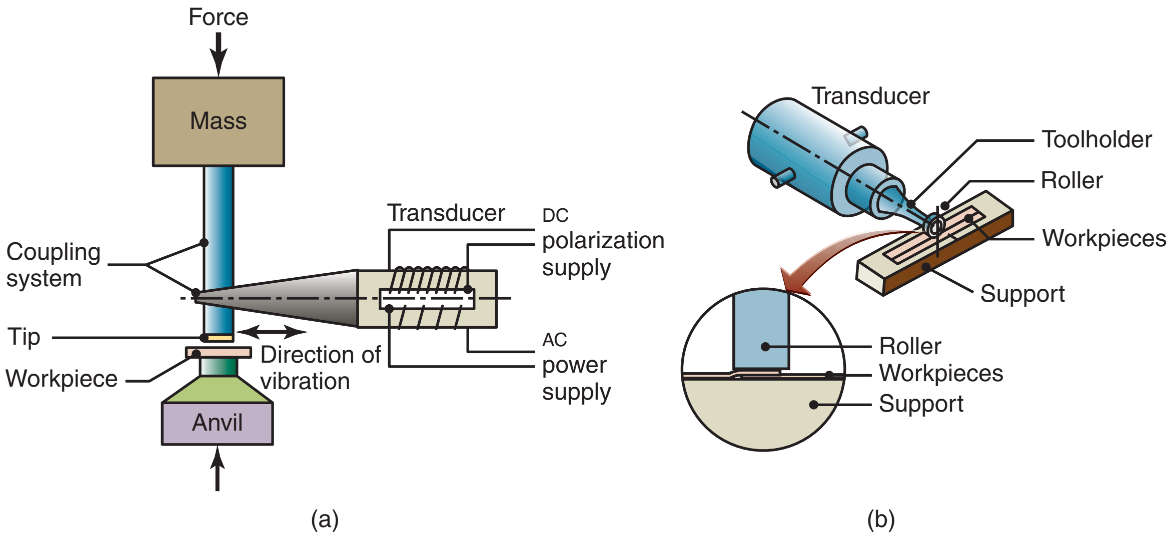


Figure 31.2: (a) Components of an ultrasonic-welding machine for making lap welds. The lateral vibrations of the tool tip cause plastic deformation and bonding at the interface of the workpieces. (b) Ultrasonic seam welding using a roller as the *sonotrode*.

31.4 Friction Welding

In the joining processes described thus far, the energy required for welding is supplied from external sources, typically chemical, electrical, or ultrasonic energy. In *friction welding* (FRW), the heat required is generated through *friction* at the interface of the two components to be joined.

Developed in the 1940s, one of the workpiece components in this process remains stationary while the other is placed in a chuck or collet, and rotated at a constant speed as high as 15 m/s. The two members are then brought into contact under an axial force (Fig. 31.3). After sufficient contact is established, the rotating member is brought to a quick stop (so that the weld is not destroyed by shearing) while the axial force is increased. Oxides and other contaminants at the interface are thus removed by the radially outward movement of the hot metal at the interface.

The pressure at the interface and the resulting friction produce sufficient heat to develop a strong joint. The weld zone is usually confined to a narrow region; its size and shape depend on the (a) level of heat generated, (b) thermal conductivity of the materials, (c) mechanical properties of the materials at elevated temperatures, (d) rotational speed, and (e) the axial pressure applied (Fig. 31.4).

Friction welding can be used to join a wide variety of materials, provided that one of the components has rotational symmetry. Solid or tubular parts can be joined, with good joint strength. Solid steel bars up to 100 mm in diameter and pipes up to 250 mm in outside diameter, have been friction welded successfully. Because of the combined heat and pressure, the interface in friction welding develops a *flash* by plastic deformation (upsetting) of the heated zone. If objectionable, it can easily be removed by machining or grinding. Friction-welding machines are fully automated, and the operator skill required is minimal, once individual cycle times for the complete operation are set properly.

Inertia Friction Welding. This process is a modification of friction welding, although the two terms have been used interchangeably. In *inertia friction welding*, the energy required for frictional heating is supplied by a *flywheel*. It is first accelerated to the proper speed, the two members are brought into contact, and an axial force is then applied. As friction at the interface begins to slow the flywheel, the axial force is increased.

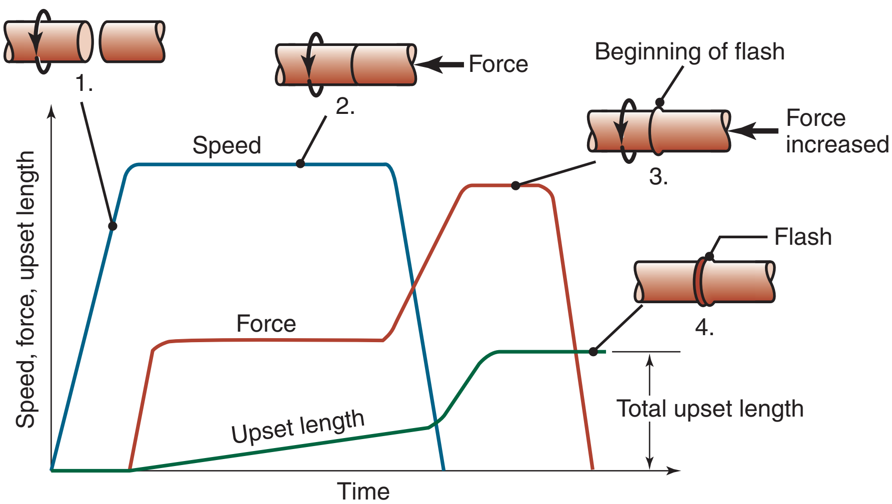


Figure 31.3: Sequence of operations in the friction-welding process: (1) The part on the left is rotated at high speed. (2) The part on the right is brought into contact with the part on the left under an axial force. (3) The axial force is increased, and the part on the left stops rotating; flash begins to form. (4) After a specified upset length or distance is achieved, the weld is completed. The *upset length* is the distance the two pieces move inward during welding after their initial contact; thus, the total length after welding is less than the sum of the lengths of the two pieces. The flash subsequently can be removed by machining or grinding.

The weld is completed when the flywheel has come to a stop; the timing of this sequence is important for good weld quality.

The rotating mass in inertia-friction-welding machines can be adjusted for applications requiring different levels of energy, depending on workpiece size and its properties. In one application, 10 mm diameter shafts are welded to automotive turbocharger impellers at a rate of one joint every 15 s.

Linear Friction Welding. In a further development of friction welding, the interface of the two components to be joined is subjected to a *linear reciprocating motion*, as opposed to a rotary motion. Thus, in this process, the components do not have to be circular or tubular in cross-section. One part is moved across the face of the other part using a balanced reciprocating mechanism. The process is capable of welding square, rectangular, or round components, and made of metals or plastics.

In one application, a rectangular titanium-alloy part was friction welded, at a linear frequency of 25 Hz with an amplitude of ± 2 mm under a pressure of 100 MPa acting on a 240 mm^2 interface. Rectangular cross-sections as large as 50 mm $\times$ 20 mm also have been welded successfully.

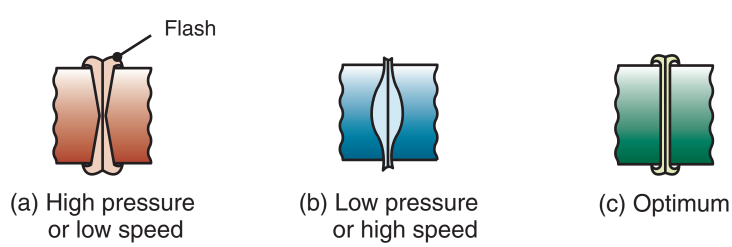


Figure 31.4: Shape of the fusion zones in friction welding as a function of the axial force applied and the rotational speed.

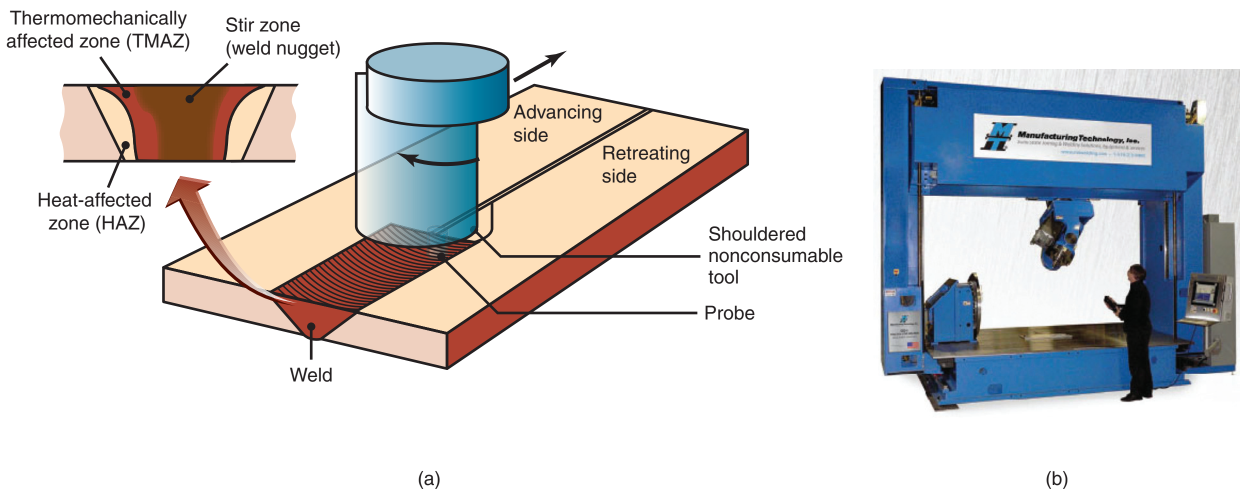


Figure 31.5: The friction-stir-welding process. (a) Schematic illustration of friction-stir-welding. Aluminum-alloy plates up to 75 mm thick have been welded by this process. (b) Multi-axis friction stir welding machine for large workpieces such as aircraft wing and fuselage structures. This machine can develop 67 kN axial forces and welding speeds up to 1.8 m/s. It is powered by a 15 kW spindle motor. *Source:* (b) Courtesy of Manufacturing Technology, Inc.

Friction Stir Welding. In *friction-stir-welding* (FSW), developed in 1991, a third body (called a *probe*) is plunged into the joint, and it rubs against the two surfaces to be joined. The nonconsumable rotating probe is typically made of cubic boron nitride (Section 8.2.3), 5 to 6 mm in diameter and 5 mm high (Fig. 31.5). The contact pressure causes frictional heating, raising the temperature to 230°C – 260°C. The tip of the rotating probe forces mixing or stirring of the material in the joint. No shielding gas or surface cleaning is required.

The thickness of the material can be as little as 1 mm and as much as 50 mm, welded in a single pass. Aluminum, magnesium, nickel, copper, steel, stainless steel, and titanium have been welded successfully; developments are taking place to extend FSW applications also to polymers and composite materials. The FSW process is being applied to aerospace, automotive, shipbuilding, and military vehicles, using sheets or plates. With developments in rotating-tool design, other possible applications include inducing microstructural changes, refining grain size in materials, and improving localized toughness in castings.

The equipment can be a conventional, vertical-spindle milling machine (see Fig. 24.18b), and the process is relatively easy to implement. For special applications, dedicated machinery for friction stir welding is available (Fig. 31.5b). Welds produced by FSW have high quality, with minimal pores and uniform structure. Because the welds are produced with low heat input, there is low distortion and little microstructural changes.

31.5 Resistance Welding

The category of *resistance welding* (RW) covers a number of processes in which the heat required for welding is produced by means of *electrical resistance* across the two components to be joined. These processes have major advantages, such as high-quality welds that do not require consumable electrodes, shielding gases, or flux, and can be produced at high rates. Resistance welding lends itself very well to automation, often using welding robots (see Section 37.6).

The heat generated in resistance welding is given by the general expression

$$H = I^2 Rt, \tag{31.1}$$

where H is the heat generated in joules (watt-seconds), I is the current (in amperes), R is the resistance (in ohms), and t is the time of current flow (in seconds). Equation (31.1) is often modified so that it represents the actual heat energy available in the weld, by including a factor K, which compensates for the energy losses through conduction and radiation. This equation then becomes

$$H = I^2 RtK, \tag{31.2}$$

where it can be noted that the value of K is less than unity.

The *total resistance* is the sum of the following (Fig. 31.6):

1. Resistances of the electrodes
2. Electrode–workpiece contact resistance
3. Resistances of the individual parts to be welded
4. Contact resistance between the faying surfaces of the two workpieces to be joined.

The actual temperature rise in the joint depends on the specific heat and thermal conductivity of the metals to be joined. For example, metals such as aluminum and copper have high thermal conductivity (see Table 3.1); hence they require high heat concentrations. Similar and dissimilar metals can be joined by this process. The current may be as high as 100,000 A, although the voltage is typically only 0.5 to 10 V. The strength of the bond developed depends on surface roughness and on the cleanliness of the mating

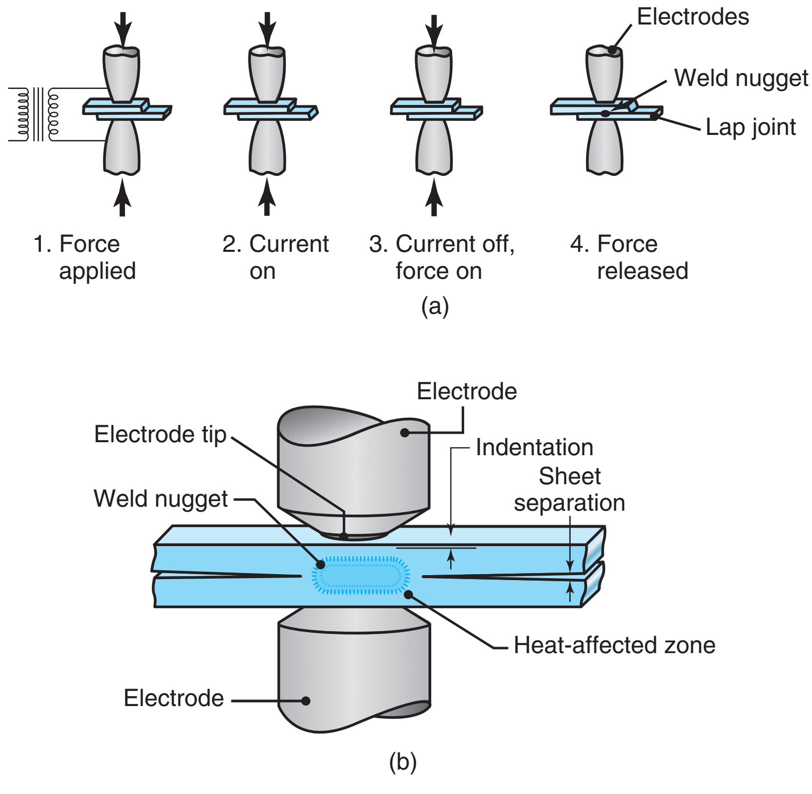


Figure 31.6: (a) Sequence of events in resistance spot welding of a lap joint. (b) Cross-section of a spot weld, showing the weld nugget and the indentation of the electrode on sheet surfaces. This is one of the most commonly used processes in sheet-metal fabrication and in automotive metal-body assembly.

surfaces. Oil films, paint, and thick oxide layers should therefore be first removed, although the presence of uniform, *thin* layers of oxide and other contaminants is not as critical for bond strength.

Developed in the early 1900s, resistance-welding processes require specialized machinery, now operated by programmable computer control. The machinery is generally not portable, and the process is suitable primarily for use in manufacturing plants and machine shops; operator skill required is minimal.

31.5.1 Resistance Spot Welding

In *resistance spot welding* (RSW) for a lap joint, the tips of two opposing solid, cylindrical electrodes touch the joint of two sheet metals, and resistance heating produces a spot weld (Fig. 31.6a). In order to obtain a strong bond in the **weld nugget**, pressure is applied until the current is turned off and the weld has solidified. Precise control and timing of the alternating current (AC) and of the pressure are essential for weld quality (see also *high-frequency resistance welding*, Section 31.5.3).

The surfaces of a spot weld has a slightly discolored *indentation*; the weld nugget (Fig. 31.6b) may be up to 10 mm in diameter. Currents range from 3000 to 40,000 A, depending on the materials being welded and their thicknesses. For example, the current is typically 10,000 A for steels and 13,000 A for aluminum. Electrodes are typically made of copper alloys and must have sufficient electrical conductivity and hot strength to maintain their shape after repeated uses.

The simplest and most commonly used resistance-welding process, spot welding may be performed by means of single or multiple pairs of electrodes (as many as a hundred or more); the required pressure is supplied through mechanical or pneumatic means. **Rocker-arm-type** spot-welding machines are typically used for smaller parts; **press-type** machines are used for larger workpieces. The shape and surface condition of the electrode tip and its accessibility are important factors. A variety of electrode shapes are used for areas that are difficult to reach (Fig. 31.7).

Spot welding is used widely for fabricating sheet-metal parts; examples range from attaching handles to stainless-steel cookware (Fig. 31.8a), to spot-welding mufflers (Fig. 31.8b), and to large sheet-metal structures. Modern spot-welding equipment is computer controlled for optimum timing of current and pressure, and the spot-welding guns are manipulated by programmable robots (Fig. 31.8c).

Testing Spot Welds. Spot-welded joints may be tested for weld-nugget strength by means of the following techniques (Fig. 31.9):

- Tension-shear
- Cross-tension
- Twist
- Peel

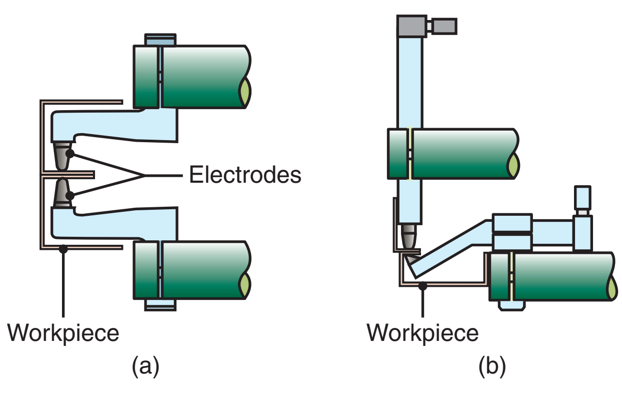


Figure 31.7: Two electrode designs for easy access to the components to be welded.

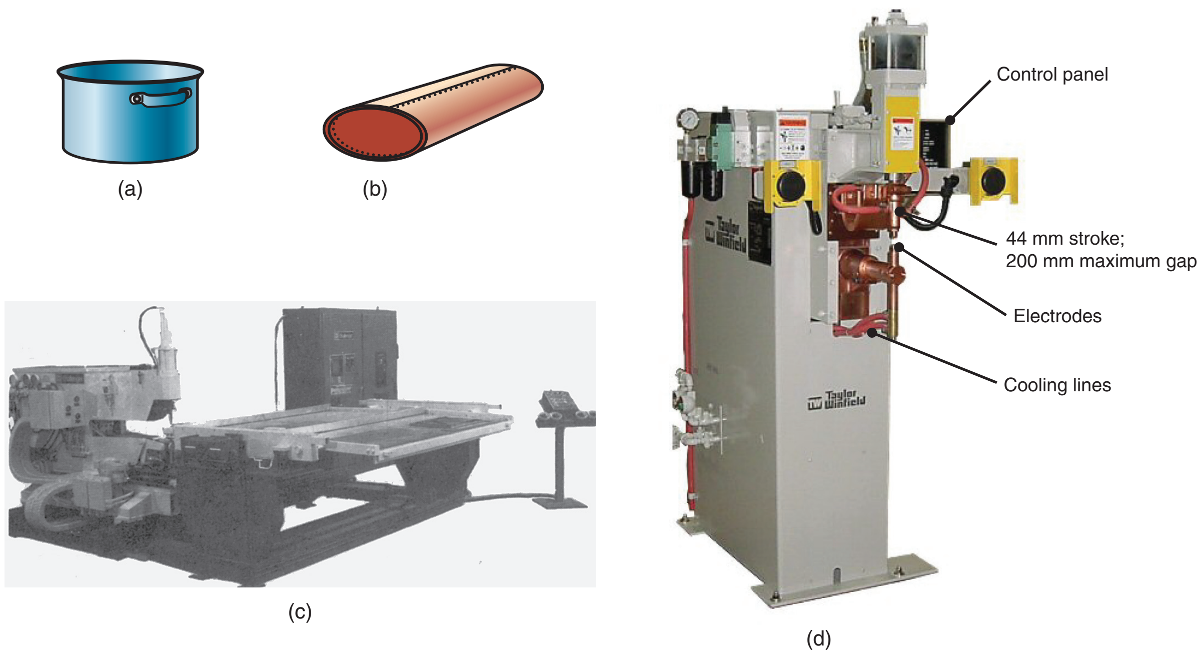


Figure 31.8: Spot-welded (a) cookware and (b) muffler. (c) An automated spot-welding machine. The welding tip can move in three principal directions. Sheets as large as 2.2×0.55 m can be accommodated in this machine with proper workpiece supports. (d) A typical spot welding machine. *Source:* (c) and (d) Courtesy of Taylor-Winfield Technologies, Inc.

The cross-tension and twist tests can indicate flaws, cracks, and porosity in the weld area. The peel test is commonly used for thin sheets. After the joint has been bent and peeled, the shape and size of the torn-out weld nugget are evaluated. Because they are easy to perform and inexpensive, these tests are commonly used in fabricating facilities.

Example 31.2 Heat Generated in Spot Welding

Given: Assume that two 1 mm thick steel sheets are being spot-welded at a current of 5000 A and over a current flow time of 0.1 s by means of electrodes 5 mm in diameter.

Find: Estimate the heat generated and its distribution in the weld zone if the effective resistance in the operation is 200 $\mu\Omega$.

Solution: From the information given, the weld-nugget volume can be estimated to be 30 mm^3. Assume that the density for steel (Table 3.1) is 8000 kg/m^3. Then the weld nugget has a mass of 0.24 g. The heat required to melt 1 g of steel is about 1400 J, so the heat required to melt the weld nugget is (1400)(1400)(0.24) = 336 J. The remaining heat (164 J) is dissipated into the metal surrounding the nugget.

31.5.2 Resistance Seam Welding

Resistance seam welding (RSEW) is a modification of spot welding wherein the electrodes are replaced by rotating wheels or rollers (Fig. 31.10a). Using a continuous AC power supply, the electrically conducting rollers produce a spot weld whenever the current reaches a sufficiently high level in the AC cycle. The typical welding speed is 1.5 m/min for thin sheets.

With sufficiently high frequency or slow traverse speed, spot welds actually overlap into a *continuous* seam and produce a joint that is liquid and gas tight (Fig. 31.10b). The RSEW process is used to make the longitudinal seam on steel cans for household products, mufflers, and gasoline tanks.

In **roll spot welding**, the current to the rolls is applied intermittently, producing a series of spot welds at specified intervals along the length of the seam (Fig. 31.10c). In **mash seam welding** (Fig. 31.10d), the overlapping welds are about one to two times the sheet thickness; the welded seam thickness is about 90% of the original sheet thickness. This process is also used in producing *tailor-welded sheet-metal blanks*, which can be made by laser welding as well (Section 16.2.2).

31.5.3 High-frequency Resistance Welding

High-frequency resistance welding (HFRW) is similar to seam welding, except that a high-frequency current of up to 450 kHz is employed. A typical application is the production of *butt-welded* tubing or pipe, where the current is conducted through two sliding contacts (Fig. 31.11a) to the edges of roll-formed tubes. The heated edges are then pressed together by passing the tube through a pair of squeeze rolls; flash that forms, if any, is then trimmed off.

Structural sections, such as I-beams, can be fabricated by HFRW, by welding the webs and flanges. Spiral pipe and tubing, finned tubes for heat exchangers, and wheel rims also can be made by this technique.

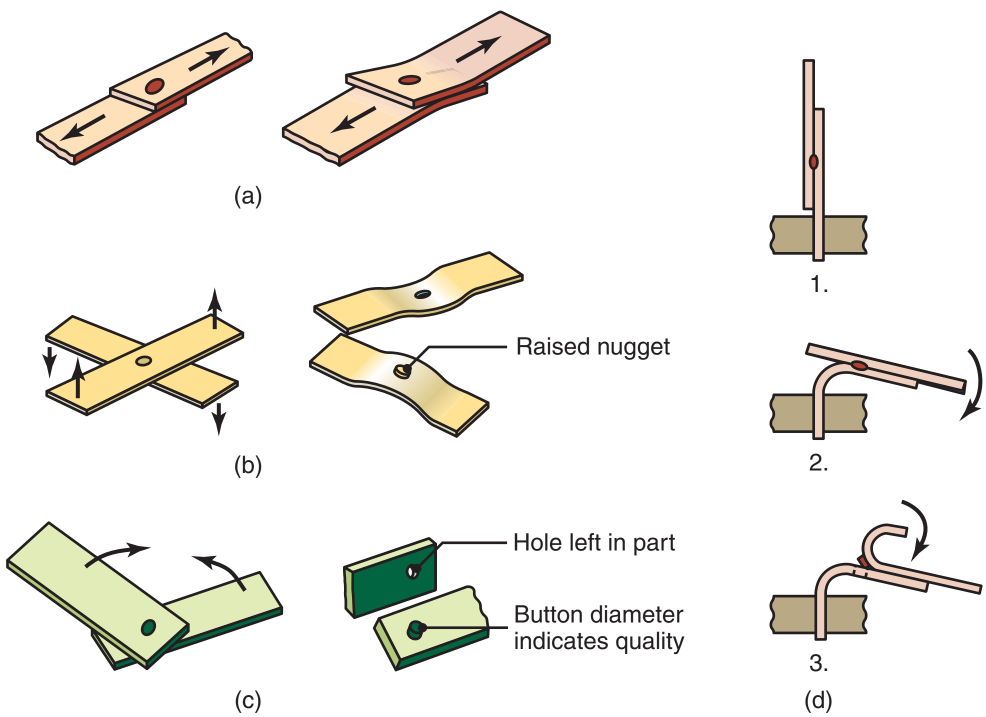


Figure 31.9: Test methods for spot welds: (a) tension-shear test, (b) cross-tension test, (c) twist test, (d) peel test (see also Fig. 32.9).

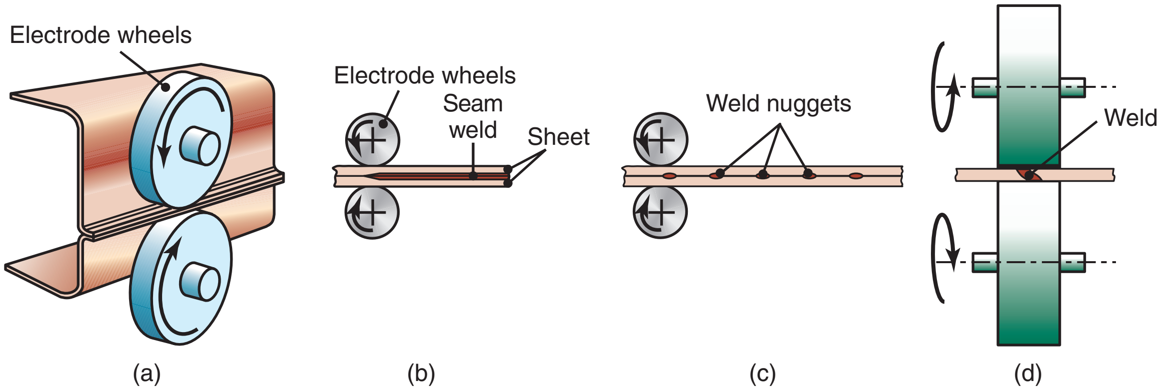


Figure 31.10: (a) Seam-welding process in which rotating rolls act as electrodes. (b) Overlapping spots in a seam weld. (c) Roll spot welds and (d) Mash seam welding.

In another method, called **high-frequency induction welding** (HFIW), roll-formed tubes (Section 16.6) are subjected to high-frequency induction heating, as shown in Fig. 31.11b.

31.5.4 Resistance Projection Welding

In *resistance projection welding* (RPW), high electrical resistance at the joint is developed by embossing one or more projections (*dimples*; see Fig. 16.39) on one of the surfaces to be welded (Fig. 31.12). The projections may be round or oval for design or strength purposes. High localized temperatures are generated at the projections, which are in contact with the flat mating part. Typically made of copper-based alloys, the electrodes are large and flat, and are water cooled to keep their temperature low. The weld nuggets are similar to those in spot welding; they are formed as the electrodes exert pressure to soften and compress and flatten the projections.

Spot-welding equipment can be used for resistance projection welding by modifying the electrodes. Although embossing of the workpieces adds to production cost, the operation produces several welds in one pass and extends electrode life; moreover, it is capable of welding metals of different thicknesses, such as a sheet welded over a plate. Nuts and bolts also can be welded to sheets and plates by this process (Fig. 31.12c and d), with projections that may be produced either by machining or forging. Joining a network of rods and wires [such as in making metal baskets, grills (Fig. 31.12e), oven racks, and shopping carts] is considered resistance projection welding, because of the small contact area between crossing wires (grids).

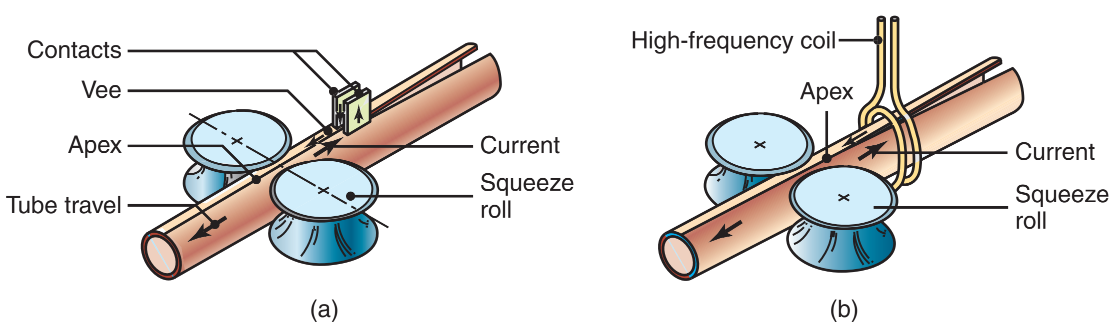


Figure 31.11: Two methods of high-frequency continuous butt welding of tubes.

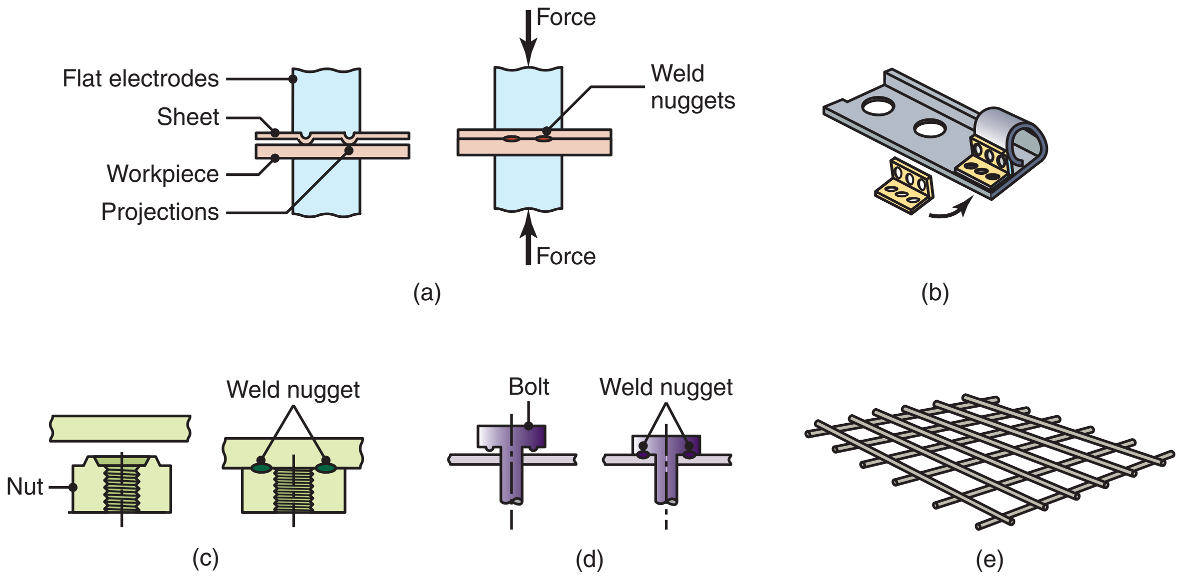


Figure 31.12: (a) Schematic illustration of resistance projection welding. (b) A welded bracket. (c) and (d) Projection welding of nuts or threaded bosses and studs. (e) Resistance-projection-welded grills.

31.5.5 Flash Welding

In *flash welding* (FW), also called **flash butt welding**, heat is generated rapidly from the arc as the ends of the two members begin to make contact, developing electrical resistance at the joint (Fig. 31.13a). After the proper temperature is reached and the interface begins to soften, an axial force is applied at a controlled rate, producing a weld by plastic deformation of the joint; joint quality is good. The mechanism involved is called *hot upsetting* (see Fig. 14.3); the term *upset welding* (UW) also is used for this process. Some molten metal is expelled from the joint as a shower of sparks during the process, thus the name *flash welding*. Because of the presence of an arc, the process can also be classified as arc welding.

Impurities and contaminants are squeezed out during this operation, and a significant amount of material may be burned off during welding. The machines for flash welding usually are automated, with a variety of power supplies, ranging from 10 to 1500 kVA.

The FW process is suitable for end-to-end or edge-to-edge joining of strips and sheets of similar or dissimilar metals, 0.2 to 25 mm thick and for end-joining bars 1 to 75 mm in diameter. Thin sections have a tendency to buckle under the axial force applied during welding. Rings made by forming, such as by the techniques shown in Fig. 16.22, can be flash butt welded. The process is also used to repair broken band-saw blades (Section 24.5), using fixtures mounted on the band-saw frame.

The process can be automated for reproducible welding operations. Typical applications are the joining of pipe and of tubular shapes for metal furniture, doors, and windows. FW is also used for welding the ends of sheets or wire in continuously operating rolling mills (Chapter 13) and in the feeding of wire-drawing equipment (Section 15.11). Some design guidelines for mating surfaces in flash welding are shown in Fig. 31.13d and e; note the importance of having uniform cross-sections at the joint.

31.5.6 Stud Welding

Stud welding (SW), also called *stud arc welding*, is similar to flash welding. The stud, which may be a threaded metal rod, hanger, or handle, serves as one of the electrodes while it is being joined to another component,

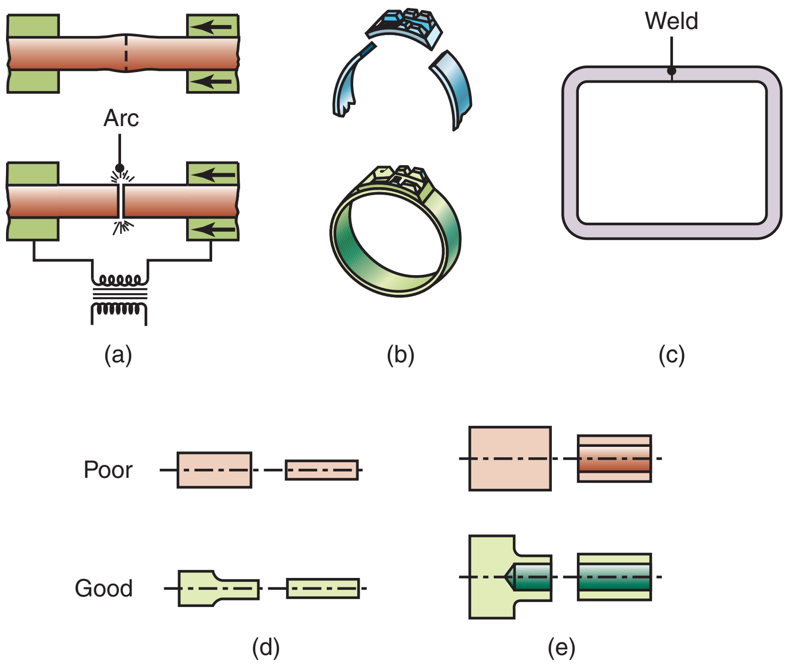


Figure 31.13: (a) Flash-welding process for end-to-end welding of solid rods or tubular parts. (b) and (c) Typical parts made by flash welding. (d) and (e) Some design guidelines for flash welding.

usually a flat plate (Fig. 31.14). Polarity for aluminum is typically direct-current electrode positive (DCEP); for steels, it is direct-current electrode negative (DCEN).

In order to concentrate the heat generated, and to prevent oxidation and retain the molten metal in the weld zone, a disposable ceramic ring (**ferrule**) is placed around the joint. The equipment for stud welding can be automated, with various controls for arcing and for applying pressure; portable stud-welding equipment is also available. Typical applications of stud welding include automobile bodies, electrical panels, shipbuilding, and in building construction.

In **capacitor-discharge stud welding**, a DC arc is produced from a capacitor bank; no ferrule or flux is required, because the welding time is on the order of only 1 to 6 milliseconds. The choice between this

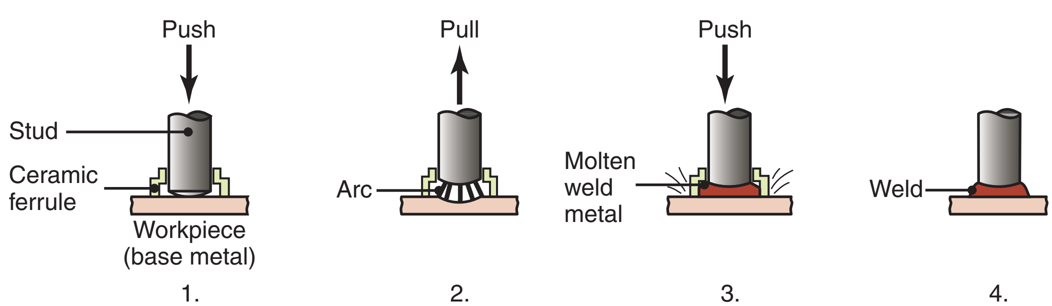


Figure 31.14: The sequence of operations in stud welding commonly used for welding bars, threaded rods, and various fasteners onto metal plates.

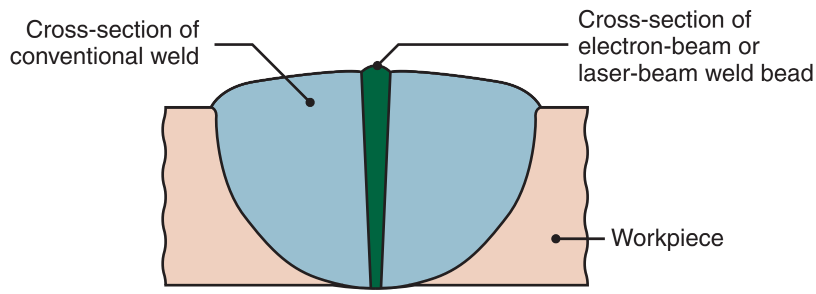


Figure 31.15: The relative sizes of the weld beads obtained by tungsten-arc and by electron-beam or laser-beam welding.

process and the stud arc welding process depends on such factors as the metals to be joined, workpiece thickness and cross-section, stud diameter, and the shape of the joint.

31.5.7 Percussion Welding

The resistance-welding processes described usually include an electrical transformer to meet the power requirements; alternatively, electrical energy for welding may be stored in a capacitor. *Percussion welding* (PEW) utilizes this technique, in which the power is discharged within 1 to 10 milliseconds, developing localized high heat at the joint. Percussion welding is useful where heating of the components adjacent to the joint is to be avoided, as, for example, in electronic assemblies and electrical wires.

Example 31.3 Resistance Welding vs. Laser-beam Welding in the Can-making Industry

The cylindrical bodies of cans for food and various household products have been resistance seam welded (with a lap joint up the side of the can) for many years. Beginning in the late 1980s, laser-beam welding technology was introduced into the can-making industry. The joints are welded by lasers, with the same productivity as in resistance welding but with the following advantages:

- As opposed to the lap joints suitable for resistance welding, laser welding utilizes butt joints; some material is thus saved. Multiplied by the billions of cans made each year, this amount becomes a very significant saving.
- Because laser welds have a very narrow zone (Fig. 31.15; see also Fig. 30.15), the unprinted area on the can surface (called *printing margin*) is greatly reduced. As a result, appearance and customer acceptance are improved.
- The resistance lap-welded joint can be subject to corrosion by the contents of the can, which can be acidic, such as orange or tomato juice, thus changing their taste; a butt joint, made by laser-beam welding, eliminates the problem.

Source: Courtesy of G.F. Benedict.

31.6 Explosion Welding

In *explosion welding* (EXW), pressure is applied by detonating a layer of explosive, placed over one of the components being joined, called the *flyer plate* (Figs. 31.16a and b). The contact pressures developed are extremely high, and the kinetic energy of the plate striking the mating component causes a wavy interface.

The impact mechanically interlocks the two surfaces (Figs. 31.16c and d), so that pressure welding by *plastic deformation* also takes place. The flyer plate is placed at an angle, and any oxide films present at the interface are broken up and propelled out from the interface. As a result, the bond strength from explosion welding is very high.

The explosive may consist of a flexible plastic sheet, cord, or in granulated or liquid form, which is cast or pressed onto the flyer plate. The detonation speed is in the range from 2400 to 3600 m/s, depending on the type of explosive, thickness of the explosive layer, and packing density of the layer. There is a minimum denotation speed necessary for welding. Detonation is carried out with a standard commercial blasting cap.

Explosive welding is suitable particularly for cladding a plate or a slab with a dissimilar metal. Plates as large as 6 m $\times$ 2 m have been clad explosively. They may then be rolled into thinner sections. Tubes and pipes can be joined to the holes in the header plates of boilers and heat exchangers by placing the explosive inside the tube; the explosion expands the tube. The process is inherently dangerous, thus it requires safe handling by well-trained and experienced personnel.

31.7 Diffusion Bonding

Diffusion bonding, or **diffusion welding** (DFW), is a process in which the strength of the joint results primarily from diffusion (movement of atoms across an interface) and secondarily from *plastic deformation* of

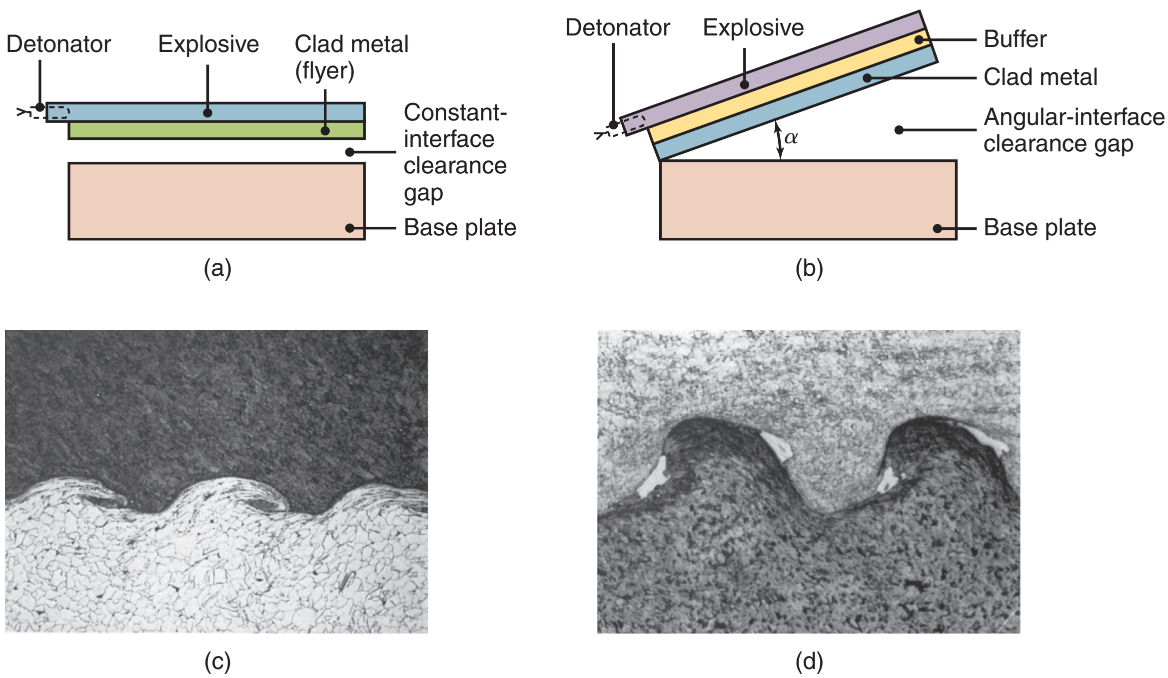


Figure 31.16: Schematic illustration of the explosion-welding process: (a) constant-interface clearance gap and (b) angular-interface clearance gap. (c) Cross-section of explosion-welded joint: titanium (top) and low-carbon steel (bottom). (d) Iron–nickel alloy (top) and low-carbon steel (bottom).

the faying surfaces. The process requires temperatures of about $0.5T_m$ (where T_m is the melting point of the metal, on the absolute scale) in order to have a sufficiently high diffusion rate between the parts being joined (Sections 1.7 and 1.8).

The interface in diffusion welding has essentially the same physical and mechanical properties as the base metal; its strength depends on (a) pressure, (b) temperature, (c) time of contact, and (d) cleanliness of the faying surfaces. These requirements can be relaxed by using a *filler metal* at the interface. Depending on the materials joined, *brittle intermetallic compounds* may form at the interface; they may be avoided by first electroplating (Section 34.9) the surfaces with suitable metal alloys. In diffusion bonding, pressure may be applied by dead weights, a press, differential gas pressure, or the thermal expansion of the parts. The parts usually are heated in a furnace or by electrical resistance; high-pressure autoclaves also are used for bonding complex parts.

Although DFW was developed in the 1970s as a modern welding technology, the principle of diffusion bonding dates back centuries when goldsmiths bonded gold over copper to develop a product called **filled gold**. First, a thin layer of gold foil is placed over copper, and pressure is applied by a weight on top of the foil. The assembly is then placed in a furnace and left there until a strong bond is developed, hence the process is also called *hot-pressure welding* (HPW).

Diffusion bonding generally is most suitable for joining dissimilar metals; it is also used for reactive metals (such as titanium, beryllium, zirconium, and refractory metal alloys) and for composite materials, such as metal-matrix composites (Section 9.5). Diffusion bonding is an important mechanism of sintering in powder metallurgy (Section 17.4). Because diffusion involves migration of the atoms across the joint, DFW is slower than other welding processes.

Although diffusion welding is used for fabricating complex parts in low quantities, for aerospace, nuclear, and electronics industries, it has been automated to also make it suitable and economical for moderate-volume production. Unless highly automated, significant operator training and skill are required.

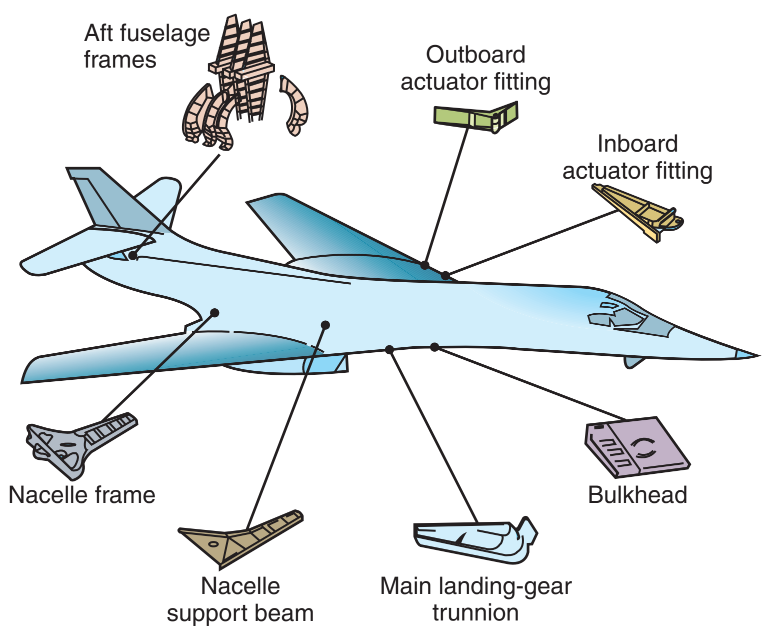


Figure 31.17: Aerospace diffusion bonding applications.

Example 31.4 Diffusion-bonding Applications

Diffusion bonding is especially suitable for such metals as titanium and the superalloys used in military aircraft. Design possibilities allow conservation of expensive strategic materials and reduction of manufacturing costs. The military aircraft illustrated in Fig. 31.17 has more than 100 diffusion-bonded parts, some of which are shown in the figure.

Diffusion Bonding–Superplastic Forming. Sheet-metal structures can be fabricated by combining *diffusion bonding* with *superplastic forming* (see also Section 16.10). Typical structures in which flat sheets are diffusion bonded and then shaped are shown in Fig. 31.18. After diffusion bonding of selected locations on the sheets, the unbonded (*stop-off*) regions are expanded in a mold either by air or fluid pressure. The structures made are thin, with high stiffness-to-weight ratios; thus they are particularly useful in aircraft and aerospace applications.

Diffusion bonding–superplastic forming improves productivity by eliminating the number of parts in a structure, mechanical fasteners, and reducing labor and manufacturing cost. It produces parts with good dimensional accuracy and low residual stresses. First developed in the 1970s, this technology is now well

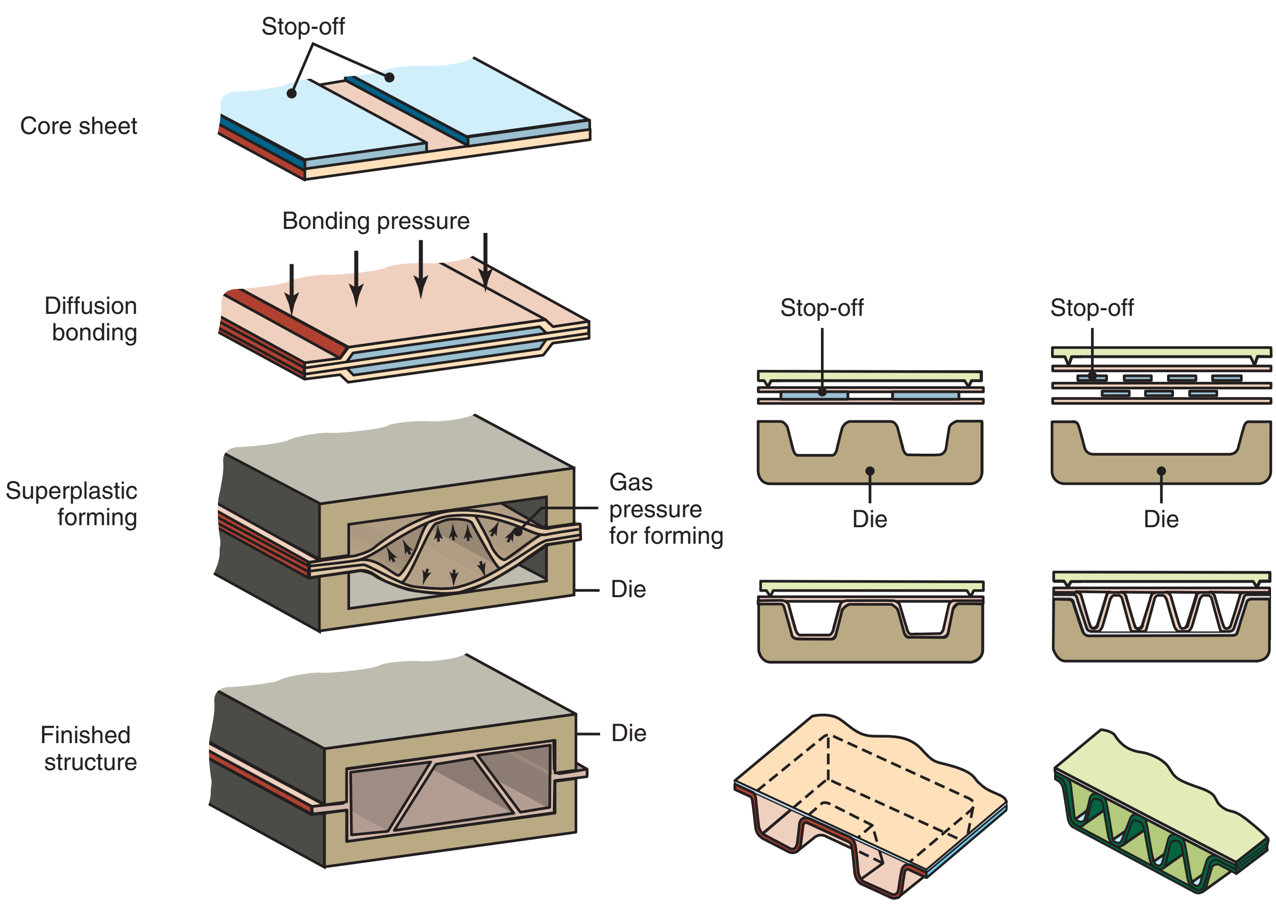


Figure 31.18: The sequence of operations in the fabrication of a structure by the diffusion bonding and superplastic forming of three originally flat sheets (see also Fig. 16.52).

advanced for titanium structures, typically using Ti-6Al-4V and 7475-T6, and various alloys for aerospace applications. Other welding processes also can be used with post-welding superplastic forming of plates, notably friction welding and friction stir welding.

31.8 Economics of Welding Operations

The characteristics, advantages, and limitations of the welding processes described thus far have included a brief introduction to welding costs. The relative costs of some selected processes are shown in Tables 30.1 and VI.1. As in all manufacturing operations, costs in welding and joining can vary widely, depending on such factors as equipment capacity, level of automation, labor skill required, weld quality, production rate, and preparation required, as well as on various other considerations specific to a specific joining operation.

Welding and joining costs for some common operations (all described throughout Chapters 30 through 32) may be summarized as:

- *High:* Brazing and fasteners (such as bolts and nuts), as they require hole-making operations and fastener costs.
- *Intermediate:* Arc welding, riveting, adhesive bonding.
- *Low:* Resistance welding, seaming, and crimping, as these operations are relatively simple to perform and to automate.

Equipment costs for welding may be summarized as:

- *High* ($100,000 to $200,000): Electron-beam and laser-beam welding.

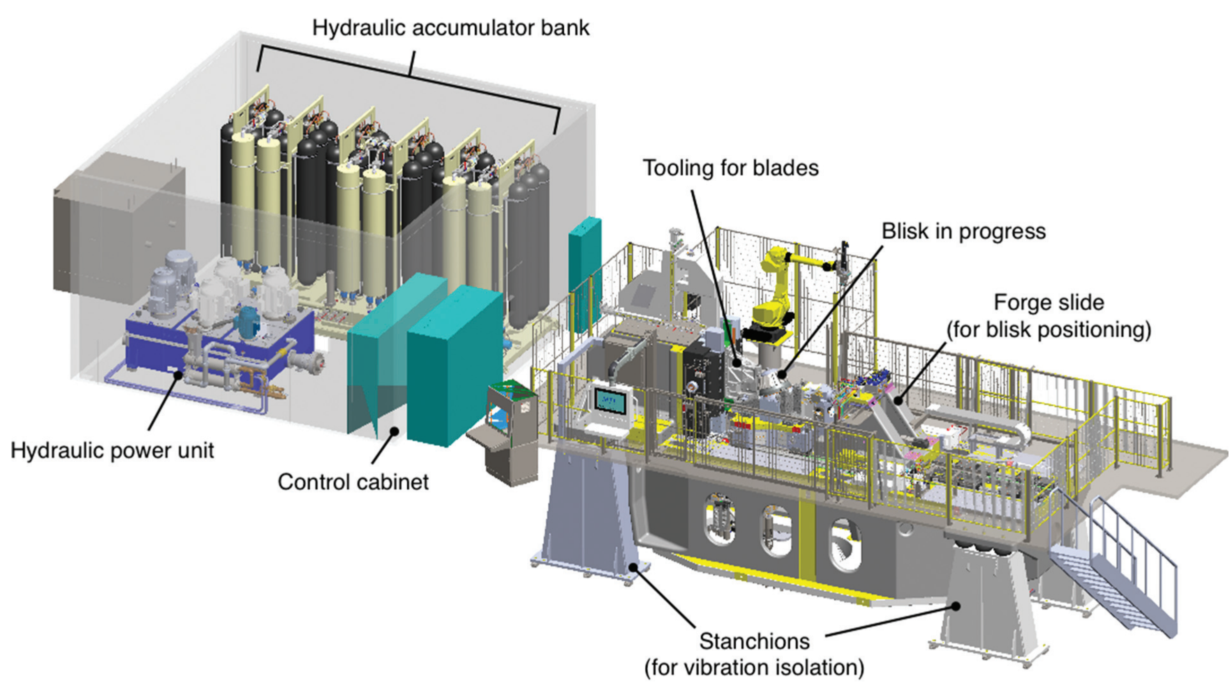


Figure 31.19: Illustration of a modern linear friction welding system for production of jet engine bladed disks (blisks). *Source:* After D. Adams, MTI Welding.

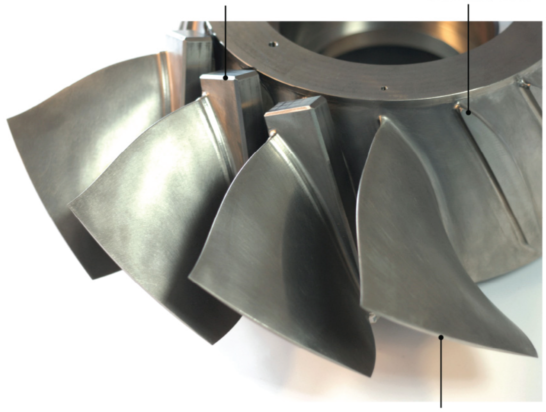


Figure 31.20: Detailed view of blades attached to a compressor disk by linear friction welding. During welding, the parts encounter plastic deformation; the block shown is later removed by machining. *Source:* Courtesy of ACB - An Aries Alliance Company.

- *Intermediate* ($5,000 to $50,000+): Spot, submerged arc, gas metal-arc, gas tungsten-arc, flux-cored arc, electrogas, electroslag, plasma arc, and ultrasonic welding.
- *Low* ($500+): Shielded metal-arc and oxyfuel–gas welding.

Labor costs in welding generally are higher than in other metalworking operations because of operator skill, welding time, and the preparations required. Much also depends on the level of automation of the equipment employed, including the wide use of robotics and computer controls, programmed to follow a prescribed path (*seam tracking*) during welding. It has been observed, for example, that in systems with robotic controls, the actual welding time reaches 80% of the total time, whereas in manual welding operations (Table 30.1), the actual time spent by the operator on welding is only about 30% of the total time.

Labor costs may be summarized as:

- *High to intermediate*: Oxyfuel–gas welding and shielded metal-arc welding.
- *High to low*: Electron-beam and laser-beam welding and flux-cored arc welding.
- *Intermediate to low*: Submerged-arc welding.

Case Study 31.1 Linear Friction Welding of Blanes and Blisks in a Jet Engine.

Titanium alloy Ti-6Al-4V bladed vanes (*blanes*) and bladed disks (*blisks*) are integral components of modern jet engines. Figure 31.19 shows a typical linear friction welding arrangement; Fig. 31.20 shows details of a typical blisk. Note that there are several blades mounted in close proximity to each other, and that very strict tolerances must be maintained for operating efficiency. Furthermore, the environment of a jet engine is very demanding; temperatures can easily exceed 1000°C, and loadings are unsteady, so that fatigue failure is an important issue.

Blanes and blisks traditionally required skilled machinists to attach them to a central hub using mechanical fasteners. This approach was time consuming and expensive, and product quality was difficult to control. Beginning with the 1990s, laser welding began to be used to fasten blades onto disks, with significant improvements in economics and performance; however, blade failures in the heat affected zone still occurred.

In 2001, linear friction welded (LFW) blanes and blisks began appearing in aerospace applications, and have seen steadily increasing use ever since. LFW involves reciprocating sliding motion under controlled pressure; the oscillation frequency is between 30 and 50 Hz, with an amplitude of 2.5 to 5.0 mm. As temperature increases, the load also increases, resulting in a pressure of around 100 MPa, sufficient to cause plastic deformation at the interface between the parts being joined. The deformation removes surface oxides and other defects from the joint. When the desired deformation is achieved, the relative motion between parts stops, resulting in a strong diffusion-based joint. Because the part cools fairly quickly, the joint is cold worked and has an advantageous microstructure for fatigue resistance.

Linear friction welding has several advantages for this application:

1. The properties of the welded joints are superior to traditional fusion-based welded joints, since friction welding does not melt the parent material. Melting causes a major change in a material's properties in the weld zone. The heat-affected zone (HAZ) of a friction-welded joint is narrow and fine grained, and with a smooth transition to the unaffected base material.
2. Complex geometries have a forged quality across the entire butt-welded area.
3. The welding process is very fast, 2 to 100 times faster than competing processes. Furthermore, it is possible to weld more than one blade at a time, thus reducing cycle times.
4. By welding the blades onto a disk, significant material savings can be achieved, as compared to designs involving machining from a single billet or block.
5. The process is energy efficient, because power requirements for LFW are as much as 20% lower than those for conventional welding. Also, the process is environmentally friendly since it requires no flux, filler metal, or shielding gases, and it does not emit smoke, fumes, or gases.
6. Being a steady-state process, LFW is extremely repeatable, and there is essentially no porosity, segregation, or slag inclusions in the weld area.

As shown in Fig. 31.20, the blades are produced with a relatively large block, and the disks have a shoe or pad for the blades. Following welding, the block and the flash have to be removed by machining them, resulting in high-quality blisks required in modern aircraft engines. The blades are attached, with a higher fatigue-resistant weld and without a heat-affected zone; as a result, the blanes and blisks are more reliable.

Source: D. Adams, Manufacturing Technology, Inc.

Summary

- In addition to the traditional joining processes of oxyfuel–gas and arc welding, several other joining processes that are based on producing a strong joint under pressure and/or heat also are available.
- Surface preparation and cleanliness are important in some of these processes. Pressure is applied mechanically or by explosives. Heat may be supplied externally, by means of electrical resistance or furnaces, or it may be generated internally, as in friction welding.
- Combining diffusion-bonding and superplastic-forming processes improves productivity and the capability to make complex parts economically.
- As in all manufacturing operations, certain hazards are inherent in welding operations. Some of these relate to the machinery and equipment used, others to the nature of the process itself, as in explosion welding. Proper safety precautions must always be taken in work areas.

Key Terms

Cold welding
Diffusion bonding (welding)
Explosion welding
Faying surfaces
Ferrule
Filled gold
Flash welding
Flyer plate
Friction stir welding
Friction welding
High-frequency resistance welding
Horn
Inertia friction welding
Linear friction welding
Percussion welding
Resistance projection welding
Resistance seam welding
Resistance spot welding
Resistance welding
Roll bonding
Roll spot welding
Roll welding
Seam welding
Solid-state welding
Sonotrode
Stud welding
Superplastic forming
Transducer
Ultrasonic welding
Weld nugget

Bibliography

ASM Handbook, Vol. 6A: **Welding Fundamentals and Processes**, ASM International, 2011.
Besharati., M.-K., and Asadi, P., **Advances in Friction-Stir Welding and Processing**, Woodhead Publishing, 2014.
Bohnart, E.R., **Welding: Principles and Practices**, 5th ed., McGraw-Hill, 2017.
Campbell, F.C., **Joining**, ASM International, 2011.
Kou, S., **Welding Metallurgy**, 2nd ed., Wiley-Interscience, 2002.
Lippold, J.C,, **Welding Metallurgy and Weldability**, Wiley, 2014.
Lohwasser, D., and Chen, Z.W., bf Friction Stir Welding, Woodhead Publishing, 2010.
Mishra, R.S., and De, P.S., **Friction Stir Welding and Processing**, Springer, 2014.
Modern Welding Technology, 11th ed., Goodheart-Willcox, 2012.
Phillips, D.H., **Welding Engineering: An Introduction**, Wiley, 2016.
Welding Handbook, various volumes and editions, American Welding Society.
Welding Inspection Handbook, 4th ed., American Welding Society, 2015.
Weman, K., **Welding Processes Handbook**, 2nd ed., Woodhead Publishing, 2011.
Yadav, K.S., **Advanced Welding Technology**, Rajsons Publishing, 2018.
Zhand, H., and Senkara, J., **Resistance Welding Fundamentals and Application**, 2nd ed., CRC Press,2011.

Review Questions

31.1. Explain what is meant by solid-state welding.

31.2. What is cold welding? Why is it so called?

31.3. What is (a) a ferrule, (b) filled gold, and (c) a flyer plate?

31.4. What are faying surfaces in solid-state welding processes?

31.5. What is the basic principle of (a) ultrasonic welding and (b) diffusion bonding?

31.6. Explain how the heat is generated in the ultrasonic welding of (a) metals and (b) thermoplastics.

31.7. Describe the advantages and limitations of explosion welding.

31.8. Describe the principle of resistance-welding processes.

31.9. What materials would you recommend for resistance welding electrodes?

31.10. What type of products are suitable for stud welding? Why?

31.11. What is the advantage of linear friction welding over inertia friction welding?

31.12. What are the main forms of friction welding?

31.13. Which processes in this chapter are applicable to polymers?

31.14. Describe how high-frequency butt welding operates.

31.15. What is diffusion bonding? What materials is it typically used with?

Qualitative Problems

31.16. Make a list of processes in this chapter, ranking them according to (a) the pressure achieved, (b) the maximum temperature, and (c) suitability for bonding dissimilar materials.

31.17. Make a list of standard abbreviations for welding processes. For example, cold welding is CW and roll welding is ROW.

31.18. Explain the reasons that the processes described in this chapter were developed.

31.19. Explain the similarities and differences between the joining processes described in this chapter and those described in Chapter 30.

31.20. Describe your observations concerning Fig. 31.16c and d.

31.21. Would you be concerned about the size of weld beads, such as those shown in Fig. 31.15? Explain.

31.22. What advantages does friction welding have over other methods described in this and in the preceding chapter?

31.23. What advantages do resistance welding processes have over others described in this chapter?

31.24. List the process parameters that you think will affect the weld strength of a friction weld, and explain why you think those parameters are important.

31.25. Describe the significance of faying surfaces.

31.26. Discuss the factors that influence the strength of (a) a diffusion-bonded and (b) a cold-welded component.

31.27. What are the sources of heat for the processes described in this chapter?

31.28. Can the roll-bonding process be applied to a variety of part configurations? Explain.

31.29. Why is diffusion bonding, when combined with the superplastic forming of sheet metals, an attractive fabrication process? Does it have any limitations?

31.30. List and explain the factors involved in the strength of weld beads.

31.31. Give some of the reasons that spot welding is used commonly in automotive bodies and in large appliances.

31.32. Explain the significance of the magnitude of the pressure applied through the electrodes during a spot-welding operation.

31.33. Give some applications for (a) flash welding, (b) stud welding, and (c) percussion welding.

31.34. Discuss the need for, and role of, work-holding devices in the welding operations described in this chapter.

31.35. Inspect Fig. 31.4, and explain why those particular fusion-zone shapes are developed as a function of pressure and speed. Comment on the influence of the material's properties.

31.36. Could the process shown in Fig. 31.11 also be applicable to part shapes other than round? Explain, and give specific examples.

31.37. In spot-weld tests, what would be the reason for weld failure to occur at the locations shown in Fig. 31.9?

31.38. Can friction stir welding be used for powder metal parts? Explain.

31.39. Do any of the processes described in this chapter use a filler metal? Explain.

31.40. Which processes in this chapter are not affected by an oxide film? Explain.

31.41. Consider the situation where two round components are welded together. You suspect that the components were friction welded, with the flash removed by machining. How could you confirm or disprove your suspicion?

31.42. Is there any advantage in preheating the workpieces in friction welding? Explain.

31.43. Inspect the edges of a U.S. quarter, and comment on your observations. Is the cross-section, i.e., the thickness of individual layers, symmetrical? Explain.

31.44. What does the strength of a weld nugget in resistance spot welding depend on?

31.45. Which applications could be suitable for the roll spot welding process shown in Fig. 31.10? Give specific examples.

Quantitative Problems

31.46. The energy required in ultrasonic welding is found to be related to the product of workpiece thickness and hardness. Explain why this relationship exists.

31.47. Two flat copper sheets (each 1.0 mm thick) are being spot welded by the use of a current of 7000 A and a current flow time of 0.3 s. The electrodes are 4 mm in diameter. Estimate the heat generated in the weld zone. Assume that the resistance is 200 $\mu\Omega$.

31.48. Calculate the temperature rise in Problem 31.47, assuming that the heat generated is confined to the volume of material directly between the two round electrodes and the temperature is distributed uniformly.

31.49. Calculate the range of allowable currents in Problem 31.47 if the temperature should be between 0.7 and 0.8 times the melting temperature of copper. Repeat this problem for carbon steel.

31.50. A resistance projection welding machine is used to join two 1.5-mm-thick sheets with eight 6-mm-diameter spot welds produced simultaneously. If 12 seconds are needed for the welding operation, determine (a) the welding current (b) the required kVA if the applied voltage is 12 V, and (c) the electrical energy consumption for each weld.

31.51. The energy applied in friction welding is given by the formula $E = IS^2/C$, where I is the moment of inertia of the flywheel, S is the spindle speed in rpm, and C is a constant of proportionality. ($C = 5873$ when the moment of inertia is given in kg-m^2.) For a spindle speed of 600 rpm and an operation in which a steel tube with a 88 mm outside diameter and a 6 mm wall thickness is welded to a flat frame, what is the required moment of inertia of the flywheel if all of the energy is used to heat the weld zone, approximated as the material 6 mm deep and directly below the tube? Assume that 1.9 J is needed to melt the electrode.

Synthesis, Design, and Projects

31.52. Comment on workpiece size and shape limitations (if any) for each of the processes described in this chapter.

31.53. Explain how you would fabricate the structures shown in Fig. 31.18 by methods other than diffusion bonding and superplastic forming.

31.54. Which materials can be friction stir welded, and which cannot? Explain your answer.

31.55. Describe part shapes that cannot be joined by the processes described in this chapter. Gives specific examples.

31.56. Comment on the feasibility of applying explosion welding in a factory environment.

31.57. Assume that you are asked to inspect a friction weld for a critical application. Describe the procedure you would follow. If you find a flaw during your inspection, how would you go about determining whether or not this flaw is important for the particular application?

31.58. Discuss your observations concerning the welding design guidelines illustrated in Fig. 31.13d and e.

31.59. Referring to Fig. 14.12b, could you use any of the processes described in Chapters 30 and 31 to make a large bolt by welding the head to the shank? Explain the advantages and limitations of this approach.

31.60. Explain how the projection-welded parts shown in Fig. 31.12 could be made by any of the processes described in this book.

31.61. Using a magnifier, inspect the cross-sections of coins such as the U.S. dime and nickel, and comment on your observations.

31.62. Describe the methods you would use for removing the flash from welds, such as those shown in Fig. 31.4. How would you automate these methods for a high-production facility?

31.63. In the roll-bonding process shown in Fig. 31.1, how would you go about ensuring that the interfaces are clean and free of contaminants so that a good bond is developed? Explain.

31.64. Inspect several metal containers for household products and for food and beverages. Identify those which have utilized any of the processes described in this chapter. Describe your observations.

31.65. Inspect the sheet-metal body of an automobile, and comment on the size and frequency of the spot welds applied. How would you go about estimating the number of welds in an automobile?

31.66. Alclad stock is made from 5182 aluminum alloy and has both sides coated with a thin layer of pure aluminum. The 5182 provides high strength, while the outside layers of pure aluminum provide good corrosion resistance because of their stable oxide film. Hence, Alclad is commonly used in aerospace structural applications. Investigate other common roll-bonded metals and their uses, and write up a summary table.

31.67. Design a test method for evaluating the bond strength in roll welding.

31.68. Review Figure 31.4 and sketch the flash pattern you would expect if (a) two tubular parts were inertia friction welded, (b) two elliptical parts were inertia friction welded, and (c) a butt weld was created with linear friction welding.

31.69. Sketch the microstructure you would expect if a butt joint were created by (a) linear friction welding, (b) friction stir welding, (c) mash seam welding, and (d) flash welding.

31.70. Design a machine that can perform friction welding of two cylindrical pieces, as well as remove the flash from the welded joint. (See Fig. 31.4.)

31.71. Assume that you are asked to give a quiz to students on the contents of this chapter. Prepare three quantitative problems and three qualitative questions, and supply the answers.

Chapter 32

Brazing, Soldering, Adhesive-bonding, and Mechanical Fastening Processes

- In brazing and soldering, no diffusion takes place at the interface; bond strength depends on adhesive forces, often increased through the use of a filler metal that produces a strong joint.
- Brazing and soldering are differentiated by the melting temperature of filler metals: brazing takes place above 450°C and produces stronger joints than soldering, whereas soldering involves lower temperatures. Soldering is widely applied in the electronics industry.
- Adhesive bonding is versatile, and a wide variety of adhesives is available.
- Mechanical joining approaches utilize fasteners such as bolts, nuts, and rivets in assembly operations.
- The chapter ends describing economic considerations in joining operations.

32.1 Introduction

In most joining processes described in Chapters 30 and 31, the faying (mating) surfaces of the components are heated by various external or internal means, to cause fusion and bonding at the joint. However, what if the parts to be joined are fragile or intricate, or they are made of two or more materials with very different characteristics and dimensions, or the components to be joined cannot withstand high temperatures, such as electronic components?

This chapter first describes two joining processes, *brazing* and *soldering*, that require lower temperatures than those used for fusion welding. Filler metals are placed in or supplied to the joint, and are melted by an external source of heat; upon solidification, a strong joint is developed (Fig. 32.1). The two processes are distinguished arbitrarily by temperature, and are lower for soldering, and higher for brazing. Moreover, the strength of a soldered joint is much lower than in brazing.

The chapter also describes the principles and types of *adhesive-bonding* processes. The ancient method of joining parts with animal-derived glues (typically employed in bookbinding, labeling, and packaging) has been developed into an important joining technology, for both metallic and nonmetallic materials. Modern adhesives consist of advanced polymers or composites; they are rarely animal based. The joining process has wide application in numerous consumer and industrial products, as well as in the aircraft and aerospace industries. Bonding materials, such as thermoplastics, thermosets, ceramics, and glasses, either to each other or to other materials, presents major challenges.

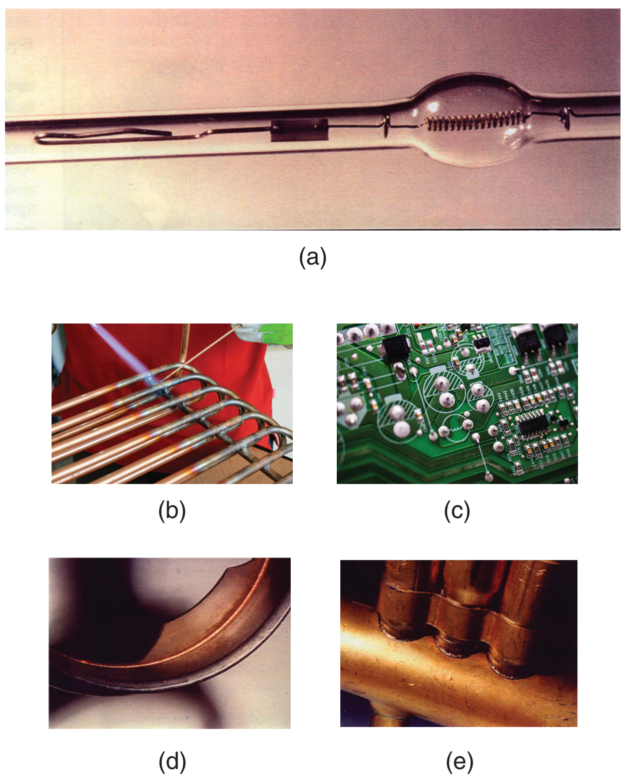


Figure 32.1: Examples of brazed and soldered joints. (a) Torch brazing of heat exchanger tubes; (b) a circuit board showing soldered components. *Source:* Courtesy of (a) Shutterstock/Bildagentur Zoonar GmbH; (b) Shutterstock/Chaikom.

Although the joints described thus far are all of a permanent nature, in many applications joined components have to be taken apart for purposes such as replacement of worn or broken components, general maintenance, or repair. Although many joints are designed not to be permanent, such as those using fasteners, they still must be strong and reliable. *Mechanical fastening* involves using bolts, screws, nuts, and a variety of special fasteners.

32.2 Brazing

First used as far back as 3000 to 2000 B.C., **brazing** is a joining process in which a **filler metal** is placed along the periphery of or between interfaces of the faying surfaces to be joined. The temperature is then raised sufficiently to melt the filler metal, but not the components (the base metal), as would be the case in fusion welding (Chapter 30). Brazing is derived from the word *brass*, an archaic word meaning to harden. It will be noted that brazing is a *liquid–solid-state bonding* process. Upon cooling and solidification of the filler metal, a strong joint is developed. Filler metals for brazing typically melt above 450°C, which is below the melting point (*solidus temperature*) of the metals to be joined (see Fig. 4.4).

A typical brazing operation is shown in Fig. 32.2a, in which a *braze metal* in the form of wire is first placed along the periphery of the components to be joined. Heat is then applied by various external means, melting the braze metal. The braze metal fills the closely fitting space (called *joint clearance*) at the interfaces through capillary action (Fig. 32.2b). In **braze welding**, filler metal (typically brass) is deposited at the joint by a technique similar to oxyfuel–gas welding (see Fig. 30.1d; see also Section 32.2.1).

Examples of joints made by brazing and soldering are shown in Fig. 32.3. Intricate, lightweight shapes can be joined rapidly with little distortion and good joint strength.

Filler Metals. Several filler metals are available, with a range of brazing temperatures (Table 32.1). Note that, unlike those for the welding operations described in the two previous chapters, filler metals for brazing generally have compositions that are significantly different from those of the metals to be joined; they are available as wire, rod, ring, shim stock, and filings. The selection of the type of filler metal and its composition are important to avoid *embrittlement* of the joint by (a) grain-boundary penetration of liquid metal (Section 1.5.2); (b) the formation of *brittle intermetallic compounds* (Section 4.2.2); and (c) *galvanic corrosion* in the joint (Section 3.8).

Because of *diffusion* between the filler metal and the base metal, the mechanical and metallurgical properties of a brazed joint can change with time as a result of subsequent processing or during service. For example, when titanium is brazed with pure tin as the filler metal, it is possible for the tin to diffuse completely into the titanium base metal when it is subjected to subsequent aging or to heat treatment.

Fluxes. Using *flux* is essential in brazing, because it prevents oxidation and removes oxide films. Brazing fluxes generally are made of borax, boric acid, borates, fluorides, and chlorides. *Wetting agents* may be added to improve both the wetting characteristics of the molten filler metal and capillary action.

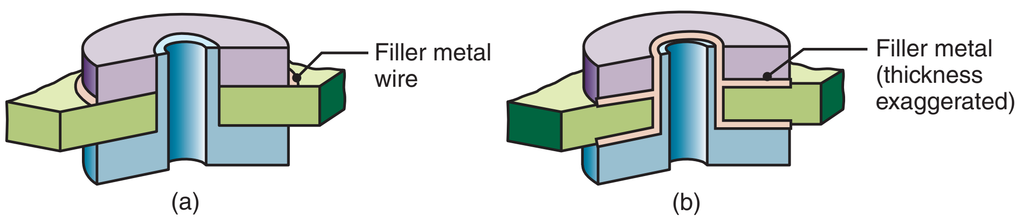


Figure 32.2: An example of furnace brazing (a) before and (b) after brazing. The filler metal is a shaped wire and the molten filler moves into the interfaces by capillary action, with the application of heat.

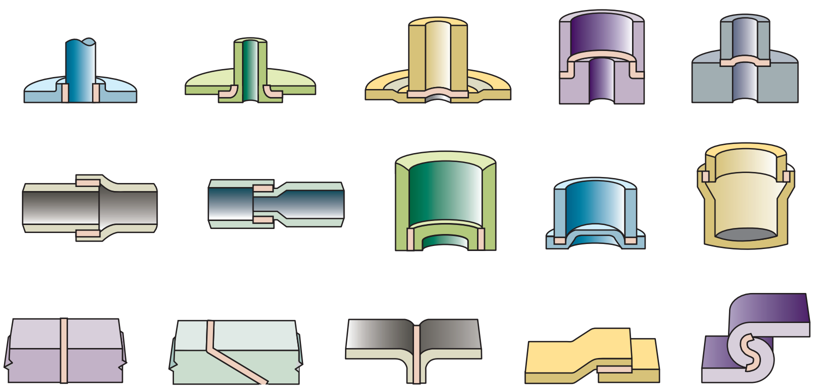


Figure 32.3: Joint designs commonly used in brazing operations. The clearance between the two parts being brazed is an important factor in joint strength. If the clearance is too small, the molten braze metal will not fully penetrate the interface; if it is too large, there will be insufficient capillary action for the metal to fill the interface.

It is essential that the surfaces to be brazed are clean and free from rust, oil, and other contaminants, in order (a) for effective wetting and distribution (spreading) of the molten filler metal in the joint interfaces and (b) to develop maximum bond strength. Grit blasting (Section 26.8) may be used to improve the surface finish of the faying surfaces. Because they are corrosive, fluxes must be removed after brazing, typically by washing with hot water.

Brazed Joint Strength. The strength of the brazed joint depends on (a) joint clearance, (b) joint area, and (c) the nature of the bond at the interfaces between the components and the filler metal. Clearances typically range from 0.025 to 0.2 mm; the smaller the gap, the higher is the *shear strength* of the joint (Fig. 32.4). Note that there is an optimum gap for achieving maximum *tensile strength* of the joint. The shear strength can reach 800 MPa by using brazing alloys containing silver (called *silver solder*). Because clearances in brazing are very small, the roughness of the faying surfaces becomes important (see also Section 33.3).

Table 32.1: Typical Filler Metals for Brazing Various Metals and Alloys.

Base metal	Filler metal	Brazing temperature (°C)
Aluminum and its alloys	Aluminum–silicon	570–620
Magnesium alloys	Magnesium–aluminum	580–625
Copper and its alloys	Copper–phosphorus and gold-copper-phosphorus	700–925
Ferrous and nonferrous (except aluminum and magnesium)	Silver and copper alloys, copper–phosphorus, copper-zinc	620–1150
Iron-, nickel-, and cobalt-based alloys	Gold-copper and gold-paladium	900–1100
Stainless steels, nickel- and cobalt-based alloys	Nickel–silver	925–1200

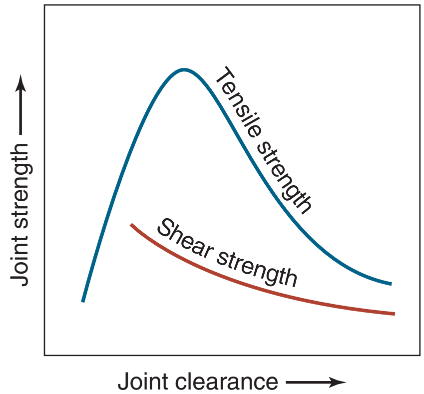


Figure 32.4: The effect of joint clearance on the tensile and shear strength of brazed joints. Note that, unlike tensile strength, the shear strength continually decreases as the clearance increases.

32.2.1 Brazing Methods

As described below, the heating methods used in brazing identify the various processes.

Torch Brazing. The heat source in *torch brazing* (TB) is oxyfuel gas with a carburizing flame (see Fig. 30.1c). Brazing is performed by first heating the joint with the torch and then depositing the brazing rod or wire at the interface. Part thicknesses are typically in the range from 0.25 to 6 mm. Torch brazing is difficult to control and requires skilled labor; it can be automated as a production process by using multiple torches.

Furnace Brazing. The parts in *furnace brazing* (FB) are first cleaned and preloaded with brazing metal in appropriate configurations; the assembly is then placed in a furnace where it is heated uniformly. Furnaces may be either batch type for complex shapes or continuous type for high production runs, especially for small parts with simple joint designs. Vacuum furnaces or neutral atmospheres are used for metals that react with the environment. Hydrogen can be used to reduce oxides in metals that are not affected by hydrogen embrittlement (Section 2.10.2). Skilled labor is not required, and complex shapes can be brazed because the whole assembly is heated uniformly in the furnace.

Induction Brazing. The source of heat in *induction brazing* (IB) is induction heating, by high-frequency AC current. Parts are preloaded with filler metal and are placed near the induction coils for rapid heating (see Fig. 4.24). Fluxes are generally required, unless a protective (neutral) atmosphere is utilized. Part thicknesses typically are less than 3 mm. Induction brazing is particularly suitable for continuous brazing of parts (Fig. 32.5).

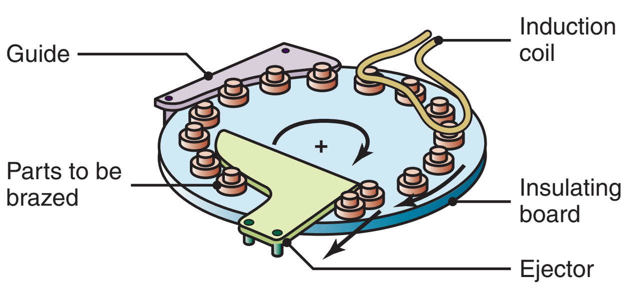


Figure 32.5: Schematic illustration of a continuous induction brazing setup for increased productivity.

Resistance Brazing. In *resistance brazing* (RB), the source of heat is the electrical resistance of the components to be brazed; electrodes are utilized in this method, as they are in resistance welding. Parts with thicknesses of 0.1 to 12 mm are either preloaded with filler metal or the metal is supplied externally during brazing. The operation is rapid, heating zones can be confined to very small areas, and the process can be automated to produce reliable and uniform joint quality.

Dip Brazing. In *dip brazing* (DB), an assembly of two or more parts are joined by dipping in a bath of filler metal, or by immersing in a bath of molten salt. In the latter case, a filler metal has to be part of the assembly. The molten salt acts as a flux, so that bonding occurs on oxide-free surfaces. The molten filler metal or the molten salt bath (Section 4.12) is at a temperature just above the melting point of the filler metal, so that all surfaces are coated with the filler. Dip brazing in metal baths is typically used for small parts, such as sheet, wire, and fittings, usually less than 5 mm in thickness or diameter.

Depending on the size of the parts and the bath size, as many as 1000 joints can be made at one time. Dip brazing usually requires self-jigging (self assembling) parts, but tack welding or pinning can be used; lap joints are preferred, although butt joints can also be made.

Infrared Brazing. The heat source in *infrared brazing* (IRB) is a high-intensity quartz lamp. The radiant energy is focused on the joint, and brazing can be carried out in a vacuum. *Microwave heating* also can be used. The process is particularly suitable for brazing very thin components, usually less than 1 mm thick, including metal honeycomb structures (Section 16.13).

Diffusion Brazing. *Diffusion brazing* (DFB) is carried out in a furnace where, with proper control of temperature and time, the filler metal diffuses into the faying surfaces of the components to be joined. The brazing time required may range from 30 min to as much as 24 hrs. This process is used for strong lap or butt joints and for difficult-to-join materials. More complex alloys may produce intermetallic compounds at the joint that can compromise joint strength. Because the rate of diffusion at the interface does not depend on the thickness of the components, part thicknesses may range from foil to as much as 50 mm.

High-energy Beams. For specialized and high-precision applications and with high-temperature metals and alloys, *electron-beam* or *laser-beam* heating may be used, as described in Sections 27.6 and 27.7.

Braze Welding. The joint in *braze welding* is prepared as in fusion welding (Chapter 30). While an oxyacetylene torch with an oxidizing flame is being used, filler metal is deposited at the joint rather than drawn in by capillary action. As a result, considerably more filler metal is used than in other forms of brazing. Temperatures in braze welding generally are lower than those in fusion welding, hence part distortion is minimal. Using a flux is essential in this process. The principal use of braze welding is for maintenance and repair work, such as on ferrous castings and steel components, although the process can be automated for mass production.

32.2.2 Design for Brazing

As in all joining processes, *joint design* is important in brazing; some guidelines are given in Fig. 32.6. Strong joints require a larger contact area for brazing than for welding. A variety of special fixtures and workholding devices and fixtures (see also Section 37.8) may be required to hold the parts together during brazing; some fixtures allow for thermal expansion and contraction during brazing.

32.3 Soldering

In *soldering*, the filler metal (called **solder**) melts at a relatively low temperature. As in brazing, the solder fills the joint by capillary action between closely fitting or closely placed components. Heat sources for soldering are typically soldering irons, torches, or ovens. The word "solder" is derived from the Latin *solidare*, meaning to make solid. Soldering with copper–gold and tin–lead alloys was first practiced as far back as 4000 to 3000 B.C.

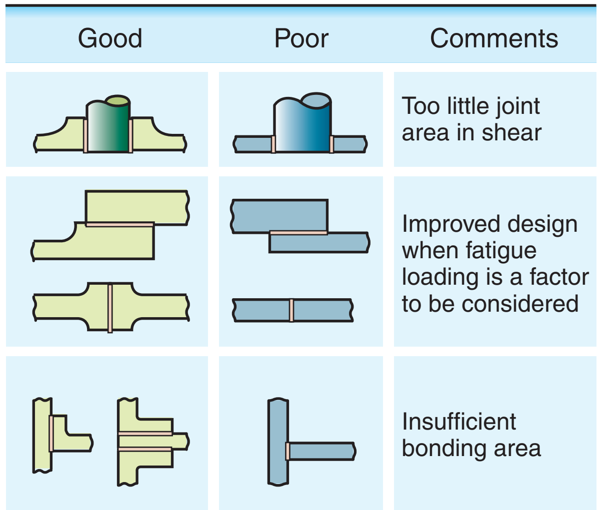


Figure 32.6: Examples of good and poor designs for brazing. *Source:* American Welding Society.

32.3.1 Types of Solders and Fluxes

Two important characteristics of solders are low surface tension and high wetting capability. Solders melt at the *eutectic point* of the solder alloy (see, for example, Fig. 4.8). Solders traditionally have been tin–lead alloys, in various proportions. A solder of 61.9% Sn–38.1% Pb composition, for example, melts at 188°C, whereas tin melts at 232°C and lead at 327°C. For special applications and higher joint strength, especially at elevated temperatures, other common solder compositions are tin–zinc, lead–silver, cadmium–silver, and zinc–aluminum alloys (Table 32.2).

Because of *toxicity* of lead, and its adverse effects on the environment, **lead-free solders** have been developed. Since the European Union prohibited intentional addition of lead to consumer electronics in 2006, tin-silver-copper solders have come into wide use, with a typical composition of 96.5% tin, 3.0% silver and 0.5% copper. A fourth element, such as zinc or manganese, is often added to provide desired mechanical or thermal characteristics. For non-electrical applications, several types of solders are available, also incorporating cadmium, gold, bismuth, and indium.

Fluxes for soldering have the same purposes as they do in welding and brazing (Section 32.2), and also serve to assist wetting of surfaces by solder. Fluxes for soldering are generally of two types:

1. *Inorganic acids* or *salts*, such as zinc–ammonium-chloride solutions, which clean the surface rapidly. To avoid corrosion, the flux residues should be removed after soldering, by washing the joint thoroughly with water.

2. Noncorrosive *resin-based fluxes*, used typically in electrical applications.

Table 32.2: A Selection of Common Solders and Their Typical Applications.

Tin–lead	General purpose
Tin–zinc	Aluminum
Lead–silver	Strength at higher than room temperature
Cadmium–silver	Strength at high temperatures
Zinc–aluminum	Aluminum, corrosion resistance
Tin–silver	Electronics
Tin–bismuth	Electronics

32.3.2 Soldering Techniques

The more common soldering techniques are the following:

1. **Torch soldering** (TS)
2. **Furnace soldering** (FS)
3. **Iron soldering** (INS)
4. **Induction soldering** (IS)
5. **Resistance soldering** (RS)
6. **Dip soldering** (DS)
7. **Infrared soldering** (IRS)

Other soldering techniques, for special applications, are the following:

8. **Ultrasonic soldering**. In this process, a transducer subjects the molten solder to ultrasonic cavitation, removing the oxide films from the surfaces to be joined. This also eliminates the need for a flux, hence this process is also known as *fluxless soldering*.
9. **Reflow (paste) soldering** (RS)
10. **Wave soldering** (WS)

The last two techniques are widely used for bonding and packaging in *surface-mount technology*, as described in Section 28.11. Because they are significantly different from other soldering methods, they are described next in some detail.

Reflow Soldering. *Solder pastes* are solder–metal particles held together by flux, binder, and wetting agents. The pastes are semisolid in consistency, have high viscosity, and thus are capable of maintaining their shape for relatively long periods. The paste is placed directly onto the joint or on flat objects. For finer detail, it can be applied via a *screening* or *stenciling* technique, as shown in Fig. 32.7a. Stenciling is common in attaching electrical components to printed circuit boards. An additional benefit of reflow soldering is that the high surface tension of the paste helps keep surface-mount packages aligned on their pads, a feature that improves the reliability of solder joints (see also Section 28.11).

After the paste has been placed and the joint is assembled, it is heated in a furnace where soldering takes place. In reflow soldering, the product is heated in a controlled manner, whereby the following sequence of events take place:

1. Solvents present in the paste are evaporated
2. The flux in the paste is activated, and fluxing action occurs
3. The components are preheated
4. The solder particles are melted and wet the joint
5. The assembly is cooled at a low rate to prevent thermal shock and possible fracture of the joint.

Although it appears to be straightforward, this process has several variables, thus good control over temperatures and durations must be maintained at each stage to ensure proper joint strength.

Wave Soldering. This is a common technique for attaching circuit components to their boards (Section 28.11). Although slowly being replaced by reflow soldering, this process is still widely used in industrial practice.

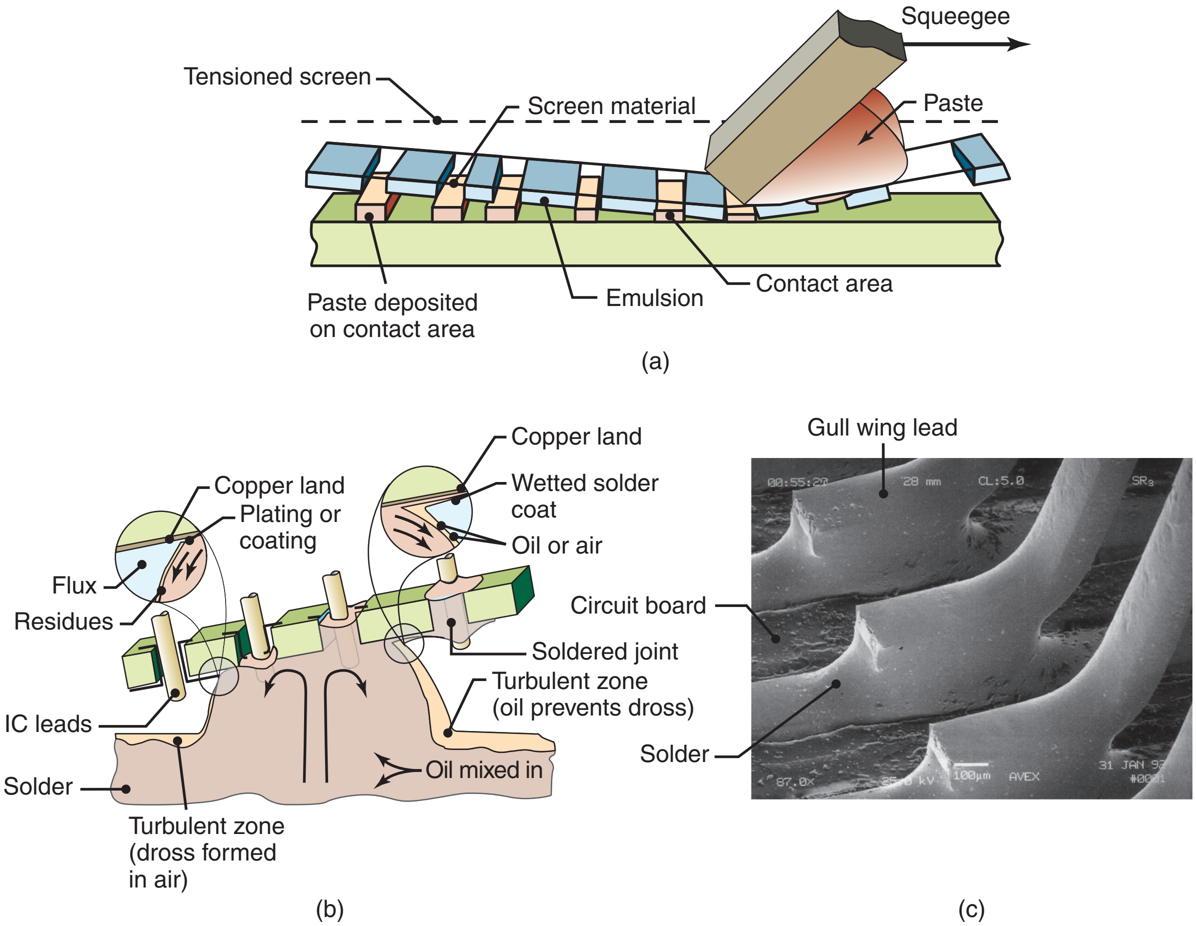


Figure 32.7: (a) Screening solder paste onto a printed circuit board in reflow soldering. (b) Schematic illustration of the wave-soldering process. (c) SEM image of a wave-soldered joint on surface-mount device. *Source:* (a) After V. Solberg.

It is important to note that because the molten solder does not wet all surfaces, it will not adhere to most polymer surfaces, and is easy to remove while in the molten state. The solder wets metal surfaces and forms a good bond, but only when the metal is preheated above a specific temperature. Wave soldering requires separate fluxing and preheating operations.

A typical wave-soldering operation is illustrated in Fig. 32.7b. A standing laminar wave of molten solder is first generated by a pump; preheated and prefluxed circuit boards are then conveyed over the wave. The solder wets the exposed metal surfaces, but (a) it does not remain attached to the polymer package for integrated circuits and (b) it does not adhere to the polymer-coated circuit boards. An *air knife* (a high-velocity jet of hot air; Section 34.11) blows excess solder away from the joint, to prevent bridging between adjacent leads.

When surface-mount packages are to be wave soldered, they must be bonded adhesively to the circuit board before soldering can begin. Bonding usually is accomplished by the following sequence: (1) screening or stenciling epoxy onto the boards, (2) placing the components in their proper locations, (3) curing the epoxy, (4) inverting the board, and (5) wave soldering. A scanning-electron-microscope (SEM) photograph of a typical surface-mount joint is shown in Fig. 32.7c.

Example 32.1 Soldering of Components onto a Printed Circuit Board

Computer and consumer electronics industries place extremely high demands on electronic components. Integrated circuits and other electronic devices are expected to function reliably for extended periods of time, during which they may be subjected to significant temperature variations and to vibration (see also Section 28.12). It is thus essential that solder joints be sufficiently strong and reliable, and also that the joints be applied extremely fast, using automated equipment.

A continuing trend in the computer and the consumer electronics industries is toward the reduction of chip sizes and increasing compactness of circuit boards. Space savings are achieved by mounting integrated circuits into surface-mount packages, allowing tighter packing on a circuit board. More importantly, the technique allows components to be mounted on *both* sides of the board.

A challenging problem arises when a printed circuit board has both surface-mount and in-line circuits on the same board, and it is essential to solder all joints via a highly reliable, automated process. It is important to recognize that, for efficiency of assembly, all of the in-line circuits be inserted from *one side* of the board.

The basic steps in soldering the connections on such a board are (see Fig. 32.7b and c):

1. Apply solder paste to one side of the board
2. Place the surface-mount packages onto the board, and insert in-line packages through the primary side of the board
3. Reflow the solder
4. Apply adhesive to the secondary side of the board
5. Using an adhesive, attach the surface-mount devices onto the secondary side
6. Cure the adhesive
7. Wave solder the other side, to produce an electrical attachment of the surface mounts and the in-line circuits to the board.

Applying solder paste is done with chemically-etched stencils or screens, so that the paste is placed only onto the designated areas of a circuit board. Stencils are used more widely for fine-pitched devices, as they produce a more uniform paste thickness. Surface-mount circuit components are then placed on the board; the board is then heated in a furnace to around 200°C, to reflow the solder and to form strong connections between the surface mount and the circuit board.

At this stage, the components with leads are inserted into the primary side of the board, their leads are crimped, and the board is flipped over. A dot of epoxy at the center of a surface mount component location is printed onto the board. The surface-mount packages are then placed onto the adhesive by high-speed automated, computer-controlled systems. The adhesive is cured, the board is flipped, and is wave soldered.

Wave soldering simultaneously joins the surface-mount components to the secondary side, and it solders the leads of the in-line components from the board's primary side. The board is then cleaned and inspected prior to electronic quality checks.

32.3.3 Solderability

Solderability may be defined in a manner similar to weldability (Section 30.9.2). Special fluxes have been developed to improve the solderability of metals and alloys. As a general guide:

- *Copper, silver,* and *gold* are easy to solder.
- *Iron* and *nickel* are more difficult to solder.

- *Aluminum* and *stainless steels* are difficult to solder, because of their thin, strong oxide films.
- *Steels*, *cast irons*, *titanium*, and *magnesium*, as well as *ceramics* and *graphite*, can be soldered, by first plating them with suitable metals to induce interfacial bonding. This method is similar to that used for joining carbides and ceramics (Section 32.6.3). An example is *tinplate*, which is steel sheet coated with tin, thus making it very easy to solder.

32.3.4 Soldering Applications and Design Guidelines

Soldering is used extensively in the electronics industry. However, because soldering temperatures are relatively low, a soldered joint has very limited use at elevated temperatures. Moreover, since they generally do not have much strength, solders cannot be used for structural (load bearing) members. Joint strength can be improved significantly by mechanical *interlocking* of the joint (Fig. 32.8).

Design guidelines for soldering are similar to those for brazing (Section 32.2.2). Some frequently used joint designs are shown in Fig. 32.8. Note the importance of large contact surfaces (because of the low strength of solders) for developing sufficient joint strength in soldered products. Since the faying surfaces generally would be small, solders are rarely used to make butt joints.

32.4 Adhesive Bonding

A very versatile joining process uses **adhesives** made of rubber or a polymer as a filler material. A common example of *adhesive bonding* is plywood, where several layers of wood are bonded with wood glue. Modern plywood was developed in 1905, but the practice of adhesive bonding of wood layers, using animal glue, dates back to 3500 B.C.

Adhesive bonding has gained increased acceptance in manufacturing ever since its first use on a large scale: the assembly of load-bearing components in aircraft during World War II (1939–1945). Adhesives are available in liquid, paste, solution, emulsion, powder, tape, and film. When applied, adhesives typically are about 0.1 mm thick.

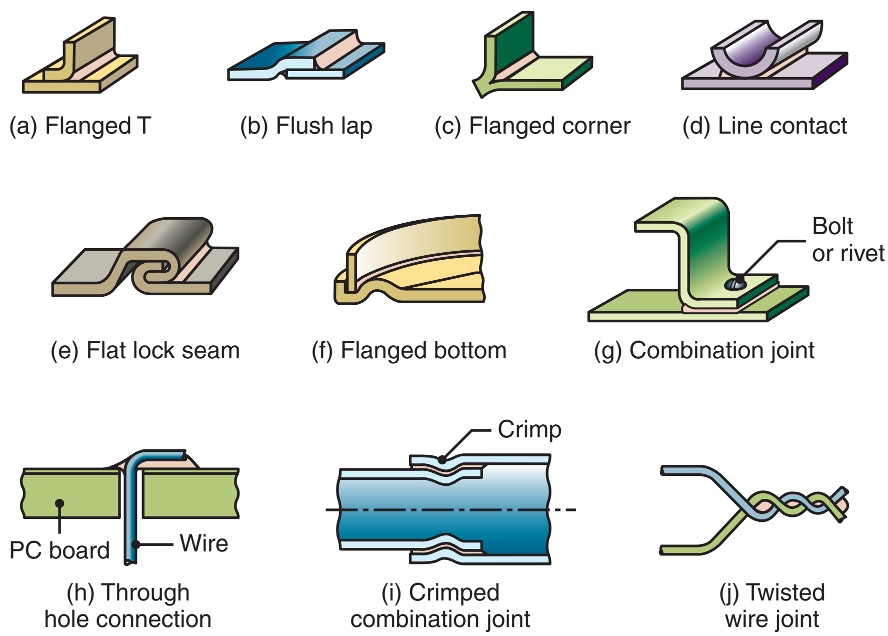


Figure 32.8: Joint designs commonly used for soldering.

Table 32.3: Typical Properties and Characteristics of Chemically Reactive Structural Adhesives.

	Epoxy	Polyurethane	Modified acrylic	Cyanoacrylate	Anaerobic
Impact resistance	Poor	Excellent	Good	Poor	Fair
Tension-shear strength, MPa	15–22	12–20	20–30	18.9	17.5
Peel strength*, N/m	< 523	14,000	5250	< 525	1750
Substrates bonded	Most	Most smooth, nonporous	Most smooth, nonporous	Most nonporous metals or plastics	Metals, glass, thermosets
Service temperature range, °C	–55 to 120	–40 to 90	–70 to 120	–55 to 80	–55 to 150
Heat cure or mixing required	Yes	Yes	No	No	No
Solvent resistance	Excellent	Good	Good	Good	Excellent
Moisture resistance	Good Excellent	Fair Fair	Good Good	Poor Poor	Good Good
Gap limitation, mm	None	None	0.5	0.25	0.60
Odor	Mild	Mild	Strong	Moderate	Mild
Toxicity	Moderate	Moderate	Moderate	Low	Low
Flammability	Low	Low	High	Low	Low

*Peel strength varies widely, depending on surface preparation and quality.

To meet the requirements of a particular application, an adhesive may require one or more of the following properties (Table 32.3):

- Strength: shear and peel
- Toughness
- Resistance to various fluids and chemicals
- Resistance to environmental degradation, including heat and moisture
- Capability to wet the surfaces to be bonded.

32.4.1 Types of Adhesives and Adhesive Systems

Several types of adhesives are available, and more continue to be developed that provide adequate joint strength, including fatigue strength (Table 32.4). Three basic types of adhesives are:

1. **Natural adhesives**, such as starch, soya flour, animal products, and dextrin (a gummy substance obtained from starch)
2. **Inorganic adhesives**, such as sodium silicate and magnesium oxychloride
3. **Synthetic organic adhesives**, which may be thermoplastics (used for nonstructural and some structural bonding) or thermosetting polymers (used primarily for structural bonding).

Because of their strength, synthetic organic adhesives are the most important adhesives in manufacturing operations, particularly for load-bearing applications. They are classified as:

- **Chemically reactive:** Polyurethanes, silicones, epoxies, cyanoacrylates, modified acrylics, phenolics, and polyimides; also included are anaerobics (which cure in the absence of oxygen), such as Loctite® for threaded fasteners (see also Case Study 32.1).

Table 32.4: General Characteristics of Adhesives.

Type	Comments	Applications
Acrylic	Thermoplastic; quick setting; tough bond at room temperature; two components; good solvent chemical and impact resistance; short work life; odorous; ventilation required	Fiberglass and steel sandwich bonds, tennis racquets, metal parts, and plastics
Anaerobic	Thermoset; easy to use; slow curing; bonds at room temperature; curing occurs in absence of air; will not cure where air contacts adherents; one component; not good on permeable surfaces	Close-fitting machine parts, such as shafts and pulleys, nuts and bolts, and bushings and pins
Epoxy	Thermoset; one or two components; tough bond; strongest of engineering adhesives; high tensile and low peel strengths; resists moisture and high temperature; difficult to use	Metal, ceramic, and rigid plastic parts
Cyanoacrylate	Thermoplastic; quick setting; tough bond at room temperature; easy to use; colorless	"Krazy Glue"; bonds most materials; especially useful for ceramics and plastics
Hot melt	Thermoplastic; quick setting; rigid or flexible bonds; easy to apply; brittle at low temperatures; based on ethylene vinyl acetate, polyolefins, polyamides, and polyesters	Bonds most materials; packaging, book binding, and metal can joints
Pressure sensitive	Thermoplastic variable strength bonds; primer anchors adhesive to roll tape backing material—a release agent on the back of web permits unwinding; made of polyacrylate esters and various natural and synthetic rubbers	Tapes, labels, and stickers
Phenolic	Thermoset; oven cured; strong bond; high tensile and low impact strength; brittle; easy to use; cures by solvent evaporation	Acoustical padding, brake lining and clutch pads, abrasive grain bonding, and honeycomb structures
Silicone	Thermoset; slow curing; flexible; bonds at room temperature; high impact and peel strength; rubberlike	Gaskets and sealants
Formaldehyde (Urea, Melamine, Phenol, Resorcinol)	Thermoset; strong with wood bonds; urea is inexpensive, is available as powder or liquid, and requires a catalyst; melamine is more expensive, cures with heat, and the bond is waterproof; resorcinol forms a waterproof bond at room temperature. Types can be combined	Wood joints, plywood, and bonding
Urethane	Thermoset; bonds at room temperature or oven cure; good gap-filling qualities	Fiberglass body parts, rubber, and fabric
Water-based (Animal, Vegetable, Rubbers)	Inexpensive, nontoxic, nonflammable	Wood, paper, fabric, leather, and dry seal envelopes

- **Pressure sensitive:** Natural rubber, styrene–butadiene rubber, butyl rubber, nitrile rubber, and polyacrylates.

- **Hot melt:** Thermoplastics (such as ethylene–vinyl acetate copolymers, polyolefins, polyamides, and polyester) and thermoplastic elastomers.

- **Reactive hot melt:** A thermoset portion (based on urethane's chemistry) with improved properties.

- **Evaporative or diffusion:** Vinyls, acrylics, phenolics, polyurethanes, synthetic rubbers, and natural rubbers.

- **Film and tape:** Nylon, epoxies, elastomer epoxies, nitrile phenolics, vinyl phenolics, and polyimides.

- **Delayed tack:** Styrene–butadiene copolymers, polyvinyl acetates, polystyrenes, and polyamides.

- **Electrically** and **thermally conductive**: Epoxies, polyurethanes, silicones, and polyimides. Electrical conductivity is obtained by the addition of fillers, such as silver (used most commonly), copper, aluminum, and gold. Fillers that improve the electrical conductivity of adhesives generally also improve their thermal conductivity.

Adhesive Systems. These systems may be classified on the basis of their specific chemistries:

- **Epoxy-based systems:** These systems have high strength and high-temperature properties up to 200°C; typical applications include automotive brake linings and bonding agents for sand molds for casting.

- **Acrylics:** These adhesives are suitable for general purpose applications, and are insensitive to substrates cleanliness.

- **Anaerobic systems:** Curing of these adhesives is done under oxygen deprivation, and the bond is usually hard and brittle; curing times can be reduced by external heat or by ultraviolet (UV) radiation.

- **Cyanoacrylate:** The bond lines are thin and the bond sets within 5 to 40 s.

- **Urethanes:** These adhesives have high toughness and flexibility at room temperature, and are used widely as sealants.

- **Silicones:** Highly resistant to moisture and solvents, these adhesives have high impact and peel strength; however, curing times are typically in the range from 1 to 5 days.

Many of these adhesives can be combined to optimize their properties, such as the combinations of *epoxy–silicon*, *nitrile–phenolic*, and *epoxy–phenolic*. The least expensive adhesives are epoxies and phenolics, followed by polyurethanes, acrylics, silicones, and cyanoacrylates. High-temperature adhesives such as polyimides and polybenzimidazoles are generally the most expensive and useful up to about 260°C. Most adhesives have an optimum temperature, ranging from about room temperature to about 200°C.

32.4.2 Electrically Conducting Adhesives

Although the majority of adhesive-bonding applications require mechanical strength, electrically conducting adhesives can replace lead-based solder alloys, particularly in the electronics industry. These adhesives require curing or setting temperatures that are lower than those required for soldering. Applications of electrically conducting adhesives include calculators, remote controls, control panels, electronic assemblies, liquid-crystal displays, and electronic games.

In these adhesives, the polymer is the matrix and contains metal fillers in such forms as flakes and particles (see also *electrically conducting polymers*, Section 28.14). There is a minimum proportion of fillers necessary to make the adhesive electrically conducting, typically in the range of 40% to 70% by volume.

The size, shape, and distribution of the metallic particles, the method of heat and pressure application, and the individual particle contact geometry can be controlled to impart isotropic or anisotropic electrical conductivity to the adhesive. The metals are typically silver, nickel, copper, and gold, as well as carbon. More recent developments include polystyrene coated with thin films of silver or gold. Graphite also can be used as a filler, usually to produce an electrically-conductive adhesive that is nonmagnetic, and provides electromagnetic interference (EMI) shielding for electronic components. Matrix materials are generally epoxies and thermoplastics, available as film or paste.

It should be noted that there are additional strategies for creating electrically conductive polymers, as described in Section 28.14. However, suspensions of silver nanoparticles or the use of conductive organic polymers such as PEDOT are not useful for adhesive applications.

32.4.3 Surface Preparation, Process Capabilities, and Applications

Surface preparation is very important in adhesive bonding, as joint strength depends on the absence of dirt, dust, oil, and various other contaminants. Observe, for example, when attempting to put an adhesive tape over a dusty or oily surface that the tape cannot develop any appreciable bond strength. Contaminants also affect the wetting ability of adhesives and prevent uniform spreading of the adhesive over an interface. Thick, weak, or loose oxide films on surfaces are detrimental to adhesive bonding. On the other hand, a porous or a thin and strong oxide film may be desirable, particularly one with some surface roughness (see Section 33.3) to improve adhesion or mechanical locking. The roughness must not be too high, because air may be trapped, reducing joint strength. Various compounds and primers are available that modify surfaces to improve bond strength.

Process Capabilities. Adhesives can be used for bonding a wide variety of similar and dissimilar metallic and nonmetallic materials, and components with different shapes, sizes, and thicknesses. Adhesive bonding can be combined with *mechanical joining* methods (Section 32.5) to further improve bond strength. Joint designs and bonding methods require care and skill; special equipment is usually required, such as fixtures, presses, tooling, and autoclaves and ovens for curing.

Nondestructive inspection of the quality and strength of adhesively bonded components can be difficult. Some of the techniques described in Section 36.10, such as acoustic impact (tapping), holography, infrared detection, and ultrasonic testing, are effective testing methods for adhesive bonds.

Testing of Adhesives. Recall that adhesives are most successful when they support shear stresses, and are less successful under other loading conditions. Further, many adhesives are weak when loaded by tensile stresses. Recognizing that loadings can be complex, a large number of test configurations have been developed to evaluate adhesives, depending on the particular application and the stresses encountered (Fig. 32.9). *Tapered cantilever* and *wedge tests* are particularly useful for high-strain-rate evaluations. Wedge tests can develop combined shear and normal stresses when the two members have different thicknesses. The most common test is the *peel test*, shown in Figs. 32.9b and 32.10, which also illustrates the strengths and limitations of adhesives. Note, for example, how easy it is to peel adhesive tape from a surface, yet it is very difficult to slide it along the surface. During peeling, the behavior of an adhesive may be brittle or ductile and tough, thus requiring high forces to peel the adhesive from a surface.

Applications. Major industries that use adhesive bonding extensively are aerospace, automotive, home appliance, and construction. Applications include automotive brake-lining assemblies, laminated windshield glass, component mounting, helicopter blades, honeycomb structures, aircraft bodies, and control surfaces.

An important consideration in using adhesives is curing time, which can range from a few seconds (at high temperatures) to several hours (at room temperature), particularly for thermosetting adhesives. Production rates can be low as compared with those of other joining processes. Moreover, adhesive bonds for structural applications rarely are suitable for service above 250°C.

Major advantages of adhesive bonding are the following:

- The interfacial bond has sufficient strength for structural applications, although it is also used for nonstructural purposes, such as sealing, insulation, prevention of electrochemical corrosion between dissimilar metals, and reduction of vibration and noise (by means of internal damping at the joints).
- Adhesive bonding effectively distributes the load at an interface, thereby eliminating localized stresses that usually result from joining the components with mechanical fasteners. Moreover, structural integrity of the sections is maintained, because no holes are required.
- The external appearance of the bonded components is unaffacted.
- Very thin and fragile components can be bonded without significant increase in their weight.
- Porous materials and those with very different properties and sizes can be joined.

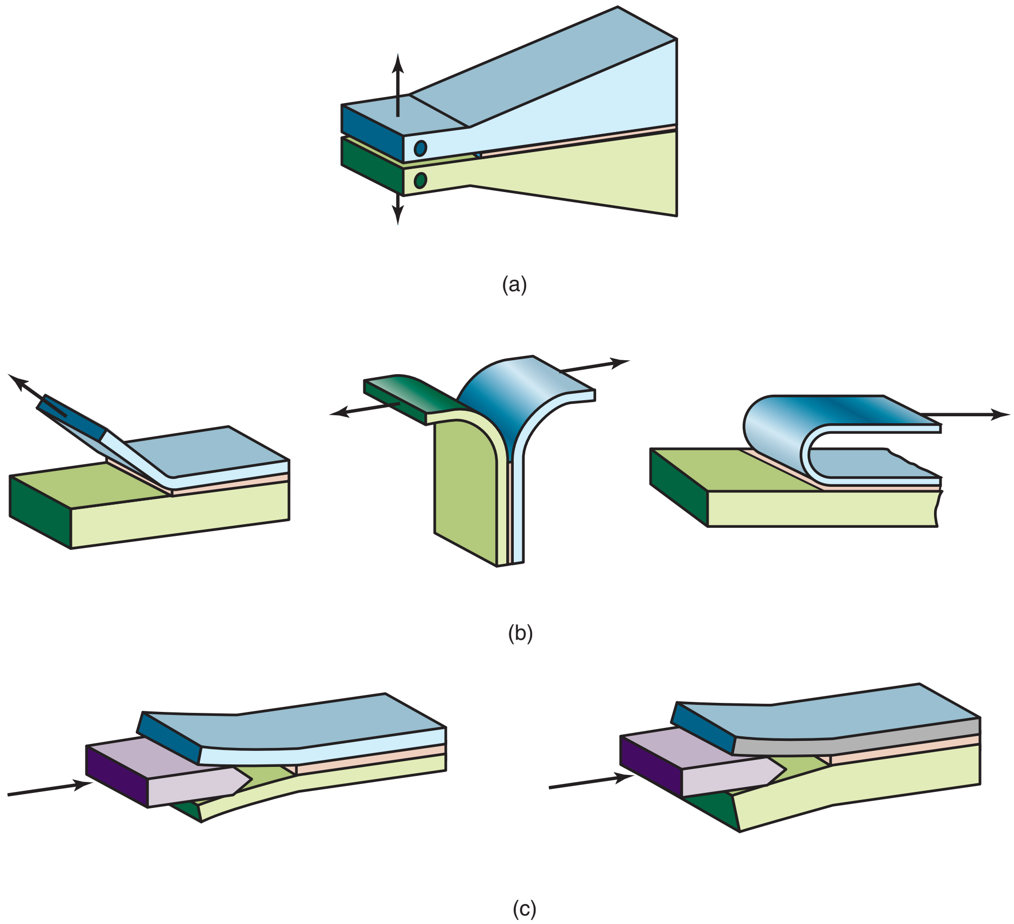


Figure 32.9: Common arrangements for evaluating adhesives: (a) tapered double cantilever beam, (b) peel tests and (c) wedge tests.

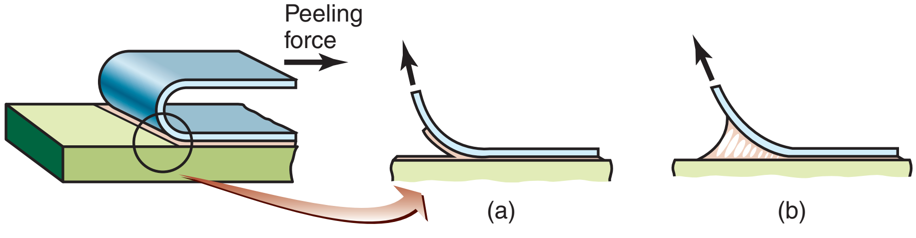


Figure 32.10: Characteristic behavior of (a) brittle and (b) tough adhesives in a peeling test. This test is similar to the peeling of adhesive tape from a solid surface.

- Because adhesive bonding is usually carried out at a temperature between room temperature and about 200°C, there is no significant distortion of the components or changes in their original properties.

The major limitations of adhesive bonding are:

- Limited range of service temperatures.
- Bonding time can be long.
- The need for great care in surface preparation.
- Bonded joints are difficult to test nondestructively, particularly for large structures.
- Limited reliability of adhesively-bonded structures during their service life and significant concerns regarding hostile environmental conditions, such as degradation by temperature, oxidation, stress corrosion, radiation, or dissolution.

The cost of adhesive bonding depends on the particular operation. In many cases, the overall economics of the process make adhesive bonding an attractive alternative, and sometimes it may be the only one that is feasible or practical. The cost of equipment varies greatly, depending on the size and type of application.

32.4.4 Design for Adhesive Bonding

- Several joint designs for adhesive bonding are shown in Figs. 32.11 to 32.13; they vary considerably in strength. The selection of appropriate design is important, and should include such considerations as the type of loading and the environment.
- Designs should ensure that joints are preferentially subjected only to compressive or shear forces, although limited tension can be supported.
- Peeling and cleavage should be avoided.

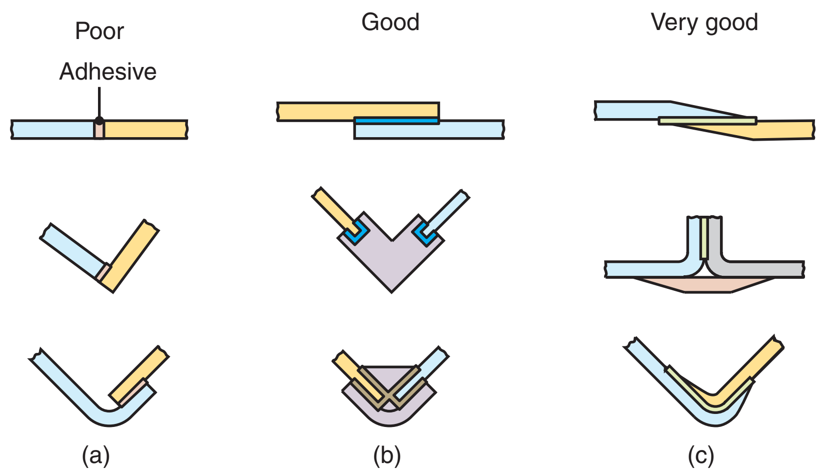


Figure 32.11: Various joint designs in adhesive bonding. Note especially that good designs require large contact areas between the members to be joined.

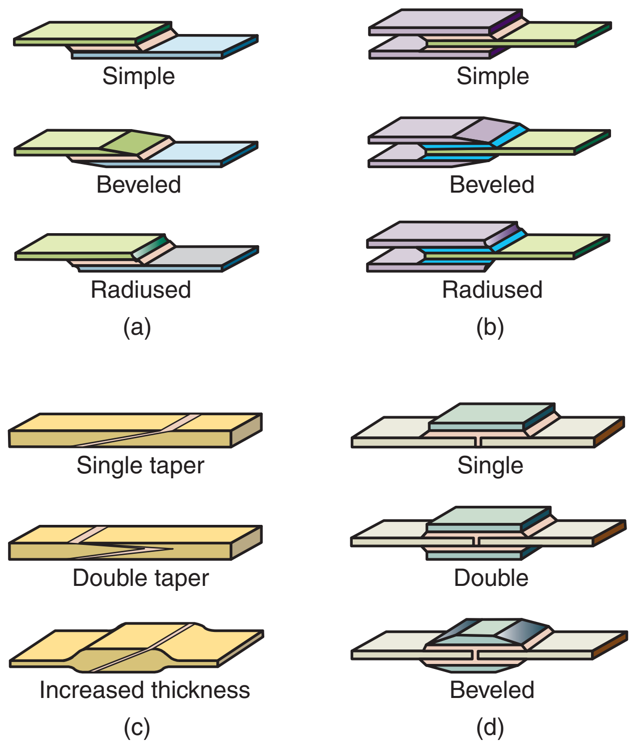


Figure 32.12: Desirable configurations for adhesively bonded joints: (a) single lap, (b) double lap, (c) scarf, and (d) strap.

- Butt joints require large bonding surfaces; tapered (scarf) joints should be used whenever feasible. Simple lap joints tend to distort under tension, because of the force couple at the joint (see Fig. 31.9.). If this is a concern, double lap joints or straps can be used (Fig. 32.12b and d).

- The coefficients of thermal expansion (Table 3.1) of the individual components to be bonded should preferably be close to each other, in order to avoid internal stresses during adhesive bonding. Thermal cycling can cause differential movements across the joint, and should be avoided.

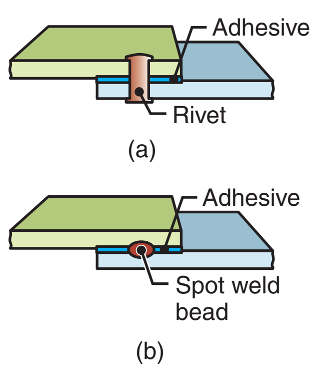


Figure 32.13: Two examples of combination joints, for purposes of improved strength, air or liquid tightness, and resistance to crevice corrosion.

32.5 Mechanical Fastening

Two or more components may have to be joined or fastened in such a way that they can be taken apart during the product's service life or its life cycle, such as shaft couplings, car wheels, appliances, engines, and bicycles. *Mechanical fastening* may be preferred over other joining methods for the following reasons:

- Ease of manufacturing
- Ease of assembly and transportation
- Ease of disassembly, maintenance, parts replacement, or repair
- Ease in creating designs that require movable joints, such as hinges, sliding mechanisms, adjustable components, and fixtures
- Lower overall cost of manufacturing the product.

The most common method of mechanical fastening is by using **fasteners**. These may be pins, rivets or keys; *threaded* fasteners, such as bolts, nuts, screws, and studs; or other types, such as various integrated fasteners. Also known as **mechanical assembly**, mechanical fastening typically requires that the components have *holes* through which fasteners are inserted. The joints may be subjected to both shear and tensile stresses, and should thus be designed to resist such forces.

Hole Preparation. An important aspect of mechanical fastening is *hole preparation*. As described in Chapters 16, 23, and 27, a hole can be produced by several means, such as punching, drilling, chemical and electrical means, and high-energy beams. Recall from Parts II and III that holes also may be *produced integrally* in products during processing, such as casting, forging, extrusion, powder metallurgy, or additive manufacturing. For improved accuracy and surface finish, many of these operations may be followed by finishing processes, such as shaving, deburring, reaming, and honing, as described in various sections of Part IV.

Because of the fundamental differences in their characteristics, each hole-making process produces a hole with different surface finish, surface properties, and dimensional accuracy. The most significant influence of a hole in a solid body is its tendency to reduce the component's fatigue life, because of stress concentrations (Section 2.7). Fatigue life can be best improved by inducing *compressive residual stresses* on hole surface in its hoop direction. These stresses usually are developed by pushing a round rod (*drift pin*) through the hole, expanding it by a very small amount. This operation plastically deforms the cylindrical surface of the hole, in a manner similar to shot peening or in roller burnishing (Section 34.2).

Threaded Fasteners. Bolts, screws, and studs are among the most commonly used *threaded fasteners*. Numerous standards and specifications include thread dimensions, dimensional tolerances, pitch, strength, and the quality of the materials used to make these fasteners.

Bolts are used with through holes and depend on a nut to develop a preload. Screws use a threaded hole or they may be *self-tapping*, whereby the screw either cuts or forms the thread into the part to be fastened. The self-tapping method is particularly effective and economical in plastic products. If the joint is to be subjected to vibration, such as in aircraft, machinery, engines, and appliances, several specially designed nuts and lock washers are available, or an anaerobic adhesive can be used.

Rivets. The most common method of permanent or semipermanent mechanical joining is by *riveting* (Fig. 32.14). Design guidelines for riveting are illustrated in Fig. 32.15. Rivets may be solid or tubular. Installing a solid rivet takes two steps: placing the rivet in the hole (usually punched or drilled) and then plastically deforming the end of its shank by upsetting it (*heading*; see Fig. 14.12). When a hole can be accessed only from one side, a *blind rivet* can be used, which uses a tubular rivet with an internal mandrel. After inserting it in a hole, the mandrel is pulled back, resulting in a flared end that locks the rivet in place (Fig. 32.14c). Specially designed rivets can drill their own holes and develop a strong joint, through

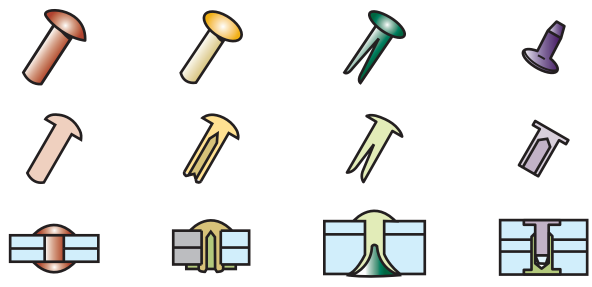


Figure 32.14: Examples of rivets: (a) solid, (b) tubular, (c) split or bifurcated, and (d) compression.

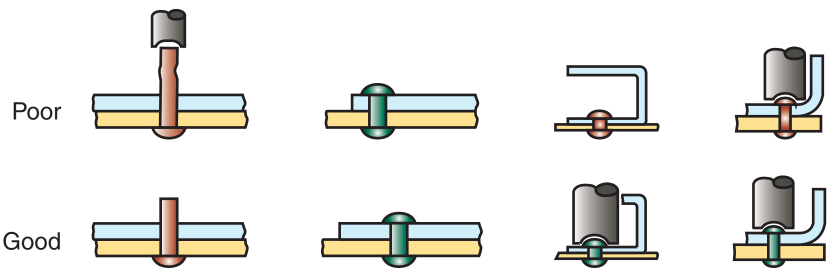


Figure 32.15: Design guidelines for riveting. (a) Exposed shank is too long; the result is buckling instead of upsetting. (b) Rivets should be placed sufficiently far from edges of the parts to avoid stress concentrations. (c) Joined sections should allow ample clearance for the riveting tools. (d) Section curvature should not interfere with the riveting process. *Source:* After J.G. Bralla.

the additional effect of friction stir welding (see Section 31.4). In another version, explosives can be placed within the rivet cavity and detonated, expanding the end of the rivet. Riveting operations can be performed manually or by mechanized means, including the use of programmable robots.

32.5.1 Various Fastening Methods

Numerous other techniques are used in joining and assembly applications.

Metal Stitching and Stapling. Illustrated in Fig. 32.16, this process is much like that of ordinary stapling of paper. The operation is fast, and it is particularly suitable for joining thin metallic and nonmetallic materials, including wood; a common example is the stapling of cardboard containers. In *clinching*, two or more materials are plastically deformed by a punch and die to produce an *interlocking* geometry. The fastener material must be sufficiently thin and ductile to withstand the large localized deformation.

Seaming. *Seaming* (Fig. 32.17) is based on the simple principle of folding two thin pieces of material together, much like joining two pieces of paper by folding them together at their top left corners. Common examples of seaming are found at the tops of beverage cans (Fig. 16.40), in containers for food and household products, and in sheet-metal ducts. The materials should be capable of undergoing bending and folding at very small radii without cracking (see Section 16.5). The performance and reliability of seams may be improved by the addition of adhesives or polymeric coatings and sealing materials or by soldering. Such approaches also make seams impermeable.

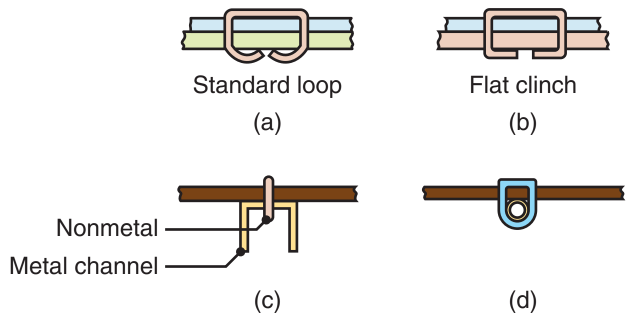


Figure 32.16: Typical examples of metal stitching.

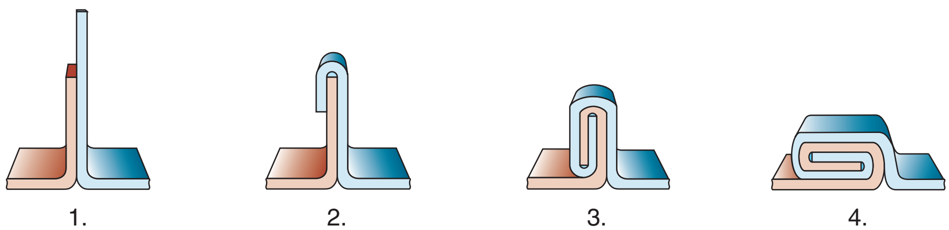


Figure 32.17: Stages in forming a double-lock seam.

Crimping. This process is a method of joining without using fasteners. It can be done with beads or dimples (Fig. 32.18), which can be produced by shrinking or swaging operations (Section 14.4). Crimping can be done on both tubular and flat components, provided that the materials are sufficiently thin and ductile, in order to undergo large localized deformations. Metal caps on glass bottles are attached by crimping; other examples include crimping connectors over electrical wiring. To provide a stronger joint, crimping can also be done using a sleeve around the parts to be joined.

Spring and Snap-in Fasteners. Several types of such fasteners are shown in Fig. 32.19. These fasteners are widely used in automotive bodies and household appliances; they are economical and permit easy and rapid assembly. *Integrated snap fasteners* are increasingly common because they ease assembly since they can be molded at the same time as the part they are to fasten.

Shrink and Press Fits. Components may be assembled by shrink or press fitting. In *shrink fitting*, a component is heated so that it expands and can be mounted over a shaft or another component; upon cooling,

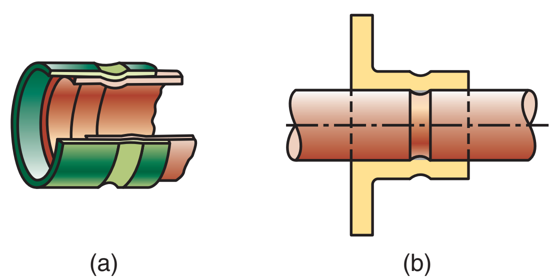


Figure 32.18: Two examples of mechanical joining by crimping.

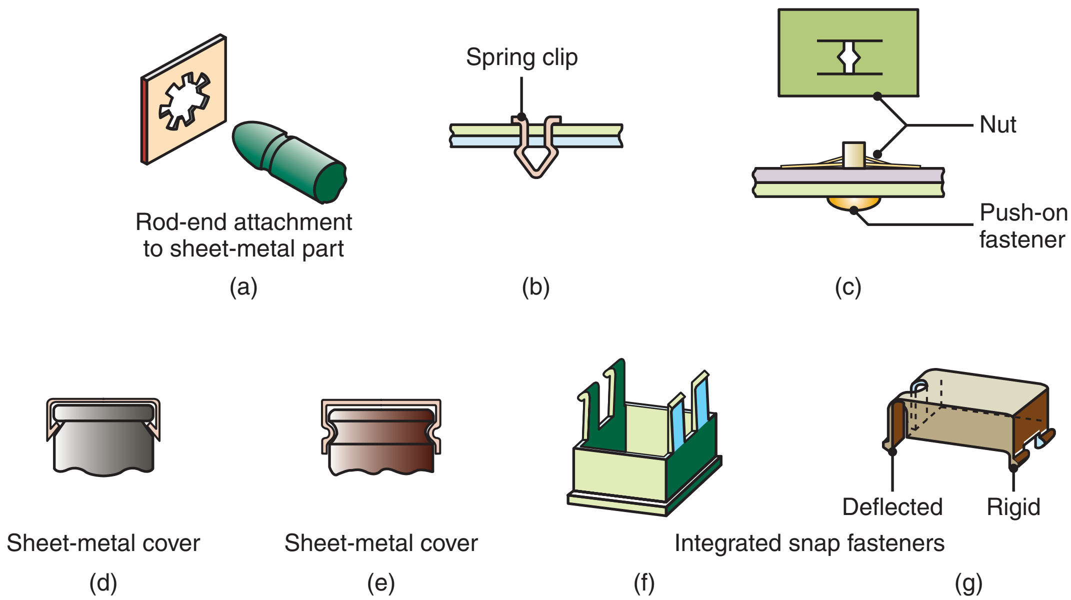


Figure 32.19: Examples of spring and snap-in fasteners, used to facilitate assembly.

it contracts and develops high contact stresses. Typical applications are assembling die components and mounting gears and cams onto shafts. In *press fitting*, one component is forced over another; when designed properly, this process results in high joint strength.

Shape-memory Alloys. The characteristics of these materials (Section 6.14) are their unique capability to recover their shape. They can be used for fasteners, with advanced applications as couplings in the assembly of titanium-alloy tubing for aircraft.

32.5.2 Design for Mechanical Fastening

The design of mechanical joints requires considerations of the type of loading to which the structure will be subjected and the size and spacing of holes. General design guidelines for mechanical joining include the following (see also Section 37.10):

- It is often beneficial to use fewer but larger fasteners than using a large number of small ones.
- Part assembly should be accomplished with a minimum number of fasteners.
- Fit between parts to be joined should be as loose as possible, to reduce costs and facilitate assembly.
- Standard size fasteners should be used whenever possible.
- Holes should not be too close to each other, the edges, or the corners, to avoid the possibility of tearing the material when subjected to external forces.

Compatibility of the fastener material with that of the components to be joined is important, as otherwise it may lead to *galvanic corrosion*, also known as *crevice corrosion* (see Section 3.8). For example, in a system in which a steel bolt or rivet is used to fasten copper sheets, the bolt is anodic and the copper plate is cathodic, a combination that causes rapid corrosion and loss of joint strength. Aluminum or zinc fasteners on copper products also react in a similar manner.

32.6 Joining Plastics, Ceramics, and Glasses

Plastics can be joined by many of the methods described for joining metals and nonmetallic materials, especially adhesive bonding and mechanical fastening.

32.6.1 Joining Thermoplastics

Thermoplastics can be joined by thermal means, adhesive bonding, solvent bonding, and mechanical fastening.

Thermal Methods. Thermoplastics (Section 7.3) soften and melt as the temperature is increased. Consequently, they can be joined when heat is generated at the interface, from either an external or internal source, allowing **fusion** to take place. The heat softens the thermoplastic at the interface to a viscous or molten state, ensuring a good bond with application of pressure.

Because of the low thermal conductivity of thermoplastics (Table 3.2), however, the heat source may burn the surfaces of the components if applied at too high a rate. Burning or charring can cause difficulties in developing sufficiently deep fusion for proper joint strength. *Oxidation* also can be a problem in joining some polymers, such as polyethylene, because it causes *degradation*. An inert shielding gas, such as nitrogen, can be used to prevent oxidation.

External heat sources may be chosen from among the following, depending on the compatibility of the polymers to be joined:

- **Hot air** or *inert gases*.
- **Hot-tool welding** or *hot-plate welding*, where heated tools and dies are pressed against the surfaces to be joined, heating them by the inter-diffusion of molecular chains. This process is commonly used in butt-welding of plastic pipes and tubing.
- **Infrared radiation** (from high-intensity quartz heat lamps) is focused into a narrow beam onto the surfaces to be joined.
- **Radio waves** are particularly useful for thin polymer films; frequencies are in the range of 100 to 500 Hz.
- **Dielectric heating**, at frequencies of up to 100 MHz, are effective for through heating of such polymers as nylon, polyvinyl chloride, polyurethane, and rubber.
- **Electrical resistance** elements (such as wires or braids, or carbon-based tapes, sheets, and ropes) are placed at the interface to create heat by the passing of electrical current, known as *resistive-implant welding*. In induction welding, these elements at the interface may be subjected to radio-frequency exposure. In both cases, the elements at the interface must be compatible with the use of the joined product, because they are left in the weld zone.
- **Lasers** emitting defocused beams at low power prevent degradation of the polymer.

Internal heat sources are developed by the following means:

- **Ultrasonic welding** (Section 31.3) is the most common for thermoplastics, particularly such amorphous polymers as acrylonitrile-butadiene-styrene (ABS) and high-impact polystyrene; frequencies are in the range of 20 to 40 kHz.
- **Friction welding** (also called *spin welding* for polymers) and *linear friction welding*, also called *vibration welding*, are particularly useful for joining those polymers with a high degree of crystallinity, such as acetal, polyethylene, nylons, and polypropylene.
- **Orbital welding** is similar to friction welding, with the exception that the rotary motion of one component is in an orbital path.

The fusion method is particularly effective with plastics that cannot be bonded easily using adhesives; such plastics as PVC, polyethylene, polypropylene, acrylics, and ABS can be joined in this manner. Specially designed portable fusion-sealing systems are used to allow in-field joining of plastic pipes, usually made of polyethylene and used for natural-gas delivery.

Coextruded multiple food wrappings consist of different types of films, bonded by heat during extrusion (Section 19.2.1). Each film has a different function; for example, one film may keep out moisture, another may keep out oxygen, and a third film may facilitate heat sealing in the packaging process. Some wrappings have as many as seven layers, all bonded together during production of the film.

Adhesive Bonding. This method is best illustrated in joining of sections of PVC pipe (used extensively in plumbing systems) and ABS pipe (used in drain, waste, and vent systems). A primer that improves adhesion is first used to apply the adhesive to the connecting sleeve and pipe surfaces (a step much like that of using primers in painting), then the pieces are pushed together.

Adhesive bonding of polyethylene, polypropylene, and polytetrafluoroethylene (*Teflon*) can be difficult, because adhesives do not bond well to them. The surfaces of parts made of these materials usually have to be treated chemically to improve bonding. Using adhesive primers or double-sided adhesive tapes also is effective.

Mechanical Fastening. This method is particularly effective for most thermoplastics (because of their inherent toughness and resilience) and for joining plastics to metals. Plastic or metal screws may be used in fastening, and the use of self-tapping metal screws is a common practice. *Integrated snap fasteners* greatly simplify assembly operations. Typical fastener geometries are shown in Fig. 32.19f and g. Because it can be molded directly at the same time as the plastic is molded, the fastener adds very little to the cost of assembly.

Solvent Bonding. This method consists of the following sequence of steps:

1. Roughening the surfaces with an abrasive
2. Wiping and cleaning the surfaces with a solvent appropriate for the particular polymer
3. Pressing the surfaces and holding them together until sufficient joint strength is developed.

Electromagnetic Bonding. Thermoplastics also may be joined by magnetic means, by embedding tiny metal particles on the order of 1 μm in diameter, in the polymer. A high-frequency electric field then causes induction heating of the polymer, melting it at the interfaces to be joined.

32.6.2 Joining Thermosets

Thermosetting plastics, such as epoxy and phenolics, can be joined by the following techniques:

- **Threaded** or molded-in **inserts**
- **Mechanical fasteners**, particularly self-tapping screws and integrated snap fasteners
- **Solvent bonding**
- **Co-curing**, in which the two components to be joined are placed together and cured simultaneously
- **Adhesive bonding**.

Case Study 32.1 Light Curing Acrylic Adhesives for Medical Products

Cobe Cardiovascular, Inc., manufactures blood collection and processing systems, as well as extra-corporeal systems for cardiovascular surgery. The company, like many other device manufacturers, traditionally used solvents for bonding device components and subassemblies. However, several federal agencies began to encourage industries to avoid using solvents. Cobe particularly wanted to eliminate using methylene chloride for environmental and occupational safety reasons. Towards this goal, the company began to redesign most of its assemblies and use light-curing (ultraviolet or visible) adhesives. Most of their devices were made of transparent plastics; consequently, its engineers needed clear adhesive bonds for aesthetic purposes and with no tendency for stress cracking or crazing.

As an example of a typical product, Cobe's blood salvage or collection reservoir is an oval polycarbonate device, approximately 300 mm tall, 200 mm in major diameter, and 100 mm deep. The reservoir is a one-time use, disposable device; its purpose is to collect and hold the blood during open-heart or chest surgery or for arthroscopic and emergency room procedures. Up to 3000 cc of blood may be stored in the reservoir while the blood awaits passage into a 250-cc centrifuge, which cleans the blood and returns it to the patient after the surgical procedure is completed. The collection reservoir consists of a clear, polycarbonate lid joined to a polycarbonate bucket. The joint has a tongue-and-groove configuration, with the goal of having a strong, elastic joint that could withstand repeated stresses with no possibility of leakage.

Light-cured acrylic adhesives offer a range of performance properties that make them well suited for this application because, first and foremost, they achieve high bond strength to the thermoplastics typically used to shape medical-device housings. For example, Loctite® 3211 (see *anaerobic adhesives*, Section 32.4.1) achieves shear strengths of 11 MPa on polycarbonate. As important as the initial shear strength may be, it is even more important that the adhesive be able to maintain high bond strength after its sterilization.

Another consideration that makes light-cured adhesives well suited for this application is their availability in formulations that allow them to withstand large strains prior to yielding; Loctite® 3211, for example, yields at elongations in excess of 200%.

Flexibility is critical, because the bonded joints are typically subjected to large bending and flexing stresses when the devices are pressurized during qualification testing and during use. If an adhesive is too rigid, it will fail in this type of testing, even if it offers higher shear strength than a comparable and more flexible adhesive. Light-cured acrylics are widely available in formulations that meet international quality standard certification (ISO; Section 36.6), meaning that when processed properly, they will not cause biocompatibility difficulties in the final assembly.

It is important to note that the joint be designed properly in order to maximize performance. If the enclosure is bonded with a joint consisting of two flat faces in intimate contact, the peeling stresses (see Fig. 32.10) will be acting on the bond whenever the vessel is pressurized. These stresses are the most difficult type for an adhesive joint to withstand, because the entire load will be concentrating on the leading edge of the joint.

The tongue-and-groove design that the company adopted addressed this concern, with the groove acting to hold and contain the adhesive during the dispensing operation. When the parts are mated and the adhesive is cured, this design allows much of the load on the joint (when the device is pressurized) to be translated into shear forces, which the adhesive is much better suited to withstand. The gap between the tongue and the groove can vary widely, because most light-cured adhesives can quickly be cured to depths in excess of 5 mm. This feature allows the manufacturer to have a robust joining process, meaning that wide dimensional tolerances can be accommodated.

With the new design and use of this adhesive, the environmental concerns and the issues associated with solvent bonding were eliminated, with the accompanying benefits of a safer, faster, and more consistent bond. The light-curing adhesive provided the aesthetic-bond line the company wanted, one that was clear and barely perceptible. The design also provided the structural strength required, thus maintaining a competitive edge for the company in the marketplace.

Source: Courtesy of P.J. Courtney, Loctite Corporation.

32.6.3 Joining Ceramics and Glasses

Ceramics and glasses often are assembled into components or subassemblies and are joined, either with the same type of material or with different metallic or nonmetallic materials. Generally, ceramics, glasses, and many similar materials can be joined by adhesive bonding. A typical example is assembling broken ceramic pieces, using a two-component epoxy, which is dispensed from two separate tubes and is mixed just prior to its application. Other joining methods include mechanical means, such as fasteners and spring or press fittings.

Ceramics. As described in Chapter 8, ceramics have properties that are very different from metallic and nonmetallic materials, especially regarding stiffness, hardness, brittleness, resistance to high temperatures, and chemical inertness. Joining them to each other or to other metallic or nonmetallic materials requires special considerations; several highly specialized joining processes are now available.

A common technique that is effective in joining difficult-to-bond combinations of materials consists of first applying a coating of a material that bonds itself well to one or both components, thus acting as a bonding agent. For example, the surface of *alumina ceramics* can be *metallized* (Section 34.5). In this technique, known as the *Mo–Mn* process, the ceramic part is first coated with a slurry of oxides of molybdenum and manganese. Next, the part is fired, forming a glassy layer on its surfaces. This layer is then plated with nickel; because the part now has a metallic surface, it can be brazed to another metal surface by using an appropriate filler metal.

Tungsten carbide and titanium carbide can easily be brazed to other metals, because they both have a metallic matrix: WC has a matrix of cobalt and TiC has nickel–molybdenum alloy as a matrix (Chapter 22). Common applications include brazing cubic boron nitride or diamond tips over carbide inserts (Fig. 22.10) and carbide tips over masonry drills (Figs. 23.22). Depending on their particular structure, ceramics and metals also can be joined by *diffusion bonding*, although it may be necessary to first place a metallic layer at the joint to make it stronger.

Ceramic components can also be joined or assembled together during their primary shaping process (Section 18.2); a common example is attaching handles to coffee mugs prior to firing them. Thus, shaping of the whole product is done *integrally* rather than as an additional operation after the part is already made.

Glasses. As evidenced by the availability of numerous glass objects, glasses can easily be bonded to each other. This is commonly done by first heating and softening the surface to be joined, then pressing the two pieces together, and cooling them. Glass can be bonded to metals, because of diffusion of metal ions into the amorphous surface structure of the glass. However, the differences in the coefficients of thermal expansion of the two materials must be taken into account.

32.7 Economics of Joining Operations

As in the economics of welding operations (Section 31.8), the joining processes discussed in this chapter depend greatly on several considerations. From Table VI.1, it can be seen that, in relative terms, the cost distribution for some of these processes are:

- *Highest:* Brazing, bolts, nuts, and other fasteners
- *Intermediate:* Riveting and adhesive bonding
- *Lowest:* Seaming and crimping

The variety of processes and the general costs involved are described below. For **brazing**,

- Manual brazing: basic equipment costs about $300, but can be over $50,000 for automated systems.
- Furnace brazing: costs vary widely, ranging from about $2000 for simple batch furnaces to $300,000 or higher for continuous vacuum furnaces.

- Induction brazing: for small units, the cost is about $10,000.
- Resistance brazing: equipment costs range from $1000 for simple units, to more than $10,000 for larger, more complex units.
- Dip brazing: equipment costs vary widely, from $2000 to more than $200,000; more expensive equipment include various computer-control features.
- Infrared brazing: equipment costs range from $500 to $30,000.
- Diffusion brazing: equipment costs range from $50,000 to $300,000.

Soldering. The cost of soldering equipment depends on its complexity and on the level of automation. Costs range from less than $20 for manual soldering irons to more than $50,000 for automated equipment.

Summary

- Joining processes that do not rely on fusion or pressure at interfaces include brazing and soldering; instead they utilize filler materials that require some temperature rise in the joint. They can be used to join dissimilar metals of intricate shapes and a range of thicknesses.
- Adhesive bonding has gained increased acceptance in such major industries as aerospace, automotive, and sports. In addition to good bond strength, adhesives have other favorable characteristics, such as the ability to seal, insulate, prevent electrochemical corrosion between dissimilar metals, and reduce vibration and noise, by means of internal damping in the bond. Surface preparation and joint design are important factors.
- Mechanical fastening is one of the most common joining methods. Bolts, screws, and nuts are typical fasteners for machine components and structures that are likely to be taken apart for maintenance and for ease of transportation.
- Rivets and fasteners are semipermanent or permanent and used in a wide variety of applications.
- Thermoplastics can be joined by fusion-welding techniques, adhesive bonding, or mechanical fastening. Thermosets are usually joined by mechanical means, such as molded-in inserts and fasteners, or by solvent bonding. Ceramics can be joined by adhesive-bonding and metallizing techniques. Glasses are joined by heating the interfaces or by using adhesives.

Key Terms

Adhesive bonding
Braze welding
Brazing
Crimping
Electrically conducting adhesives
Fasteners
Filler metal
Flux
Hole preparation
Integrated snap fastener
Lead-free solders
Mechanical fastening
Press fitting
Reflow soldering
Rivet
Seaming
Shrink fitting
Snap-in fastener

Soldering
Solvent bonding
Stapling
Stitching
Threaded fasteners
Wave soldering

Bibliography

Adams, R.D. (ed.), **Adhesive Bonding: Science, Technology and Applications**, CRC Press, 2005.
ASM Handbook, Vol. 6A, **Welding Fundamentals and Proceses**, ASM International, 2011.
Bickford, J.H., **Introduction to the Design and Behavior of Bolted Joints**, 4th ed., CRC Press, 2007.
Brazing Handbook, 5th ed., American Welding Society, 2007.
Brockmann, W., Geiss, P.L., Klingen, J., and Schreoeder, K.B., **Adhesive Bonding: Materials, Applications and Technology**, Wiley, 2009.
Campbell, F.C., (ed.), **Joining: Understanding the Basics**, ASM International, 2011.
Dillard, D.A., (ed.), **Advances in Structural Adhesive Bonding**, Woodhead Publishing, 2016.
Ebnesajjad, S., **Adhesives Technology Handbook**, 3rd ed., William Andrew, 2014
Ebnesajjad, S., and Ebnesajjad, C., **Surface Treatment of Materials for Adhesive Bonding**, 2nd ed., William Andrew, 2013.
Gourley, R., and Walker, C. (eds.), **Brazing and Soldering 2012**, American Society for Metals, 2012.
Grewell, D., **Plastics and Composites Welding Handbook**, Hanser, 2003.
Humpston, G., and Jacobson, D.M., **Principles of Soldering**, ASM International, 2004.
Jacobson, D.M., and Humpston, G., **Principles of Brazing**, ASM International, 2005.
Pang, J.H.L., **Lead Free Solder: Mechanics and Reliability**, Springer, 2011.
Petrie, E.M., **Handbook of Adhesives and Sealants**, McGraw-Hill, 2nd ed., 2006.
Pizzi, A., and Mittal, K.L., **Handbook of Adhesive Technology**, 3rd ed., CRC Press, 2017.
Roberts, P., **Industrial Brazing Practice**, 2nd ed., CRC Press, 2013.
Rotheiser, J., **Joining of Plastics: Handbook for Designers and Engineers**, 3rd ed., Hanser, 2009.
Schwartz, M.M., **Soldering: Understanding the Basics**, ASM International, 2014.
Subramanian, K., **Lead-Free Solders**, Wiley, 2012.
Testing of Adhesive Joints: Best Practices, Wiley-VCH, 2012.
Troughton, M.J., (ed.), **Handbook of Plastics Joining**, 2nd ed., William Andrew, 2008.
Wegman, R.F., **Surface Preparation Techniques for Adhesive Bonding**, 2nd ed., William Andrew, 2012.

Review Questions

32.1. What is the difference between brazing and braze welding?

32.2. Are fluxes necessary in brazing? If so, why?

32.3. Why is surface preparation important in Adhesive-bonding?

32.4. What materials are typically used in solder?

32.5. Soldering is generally applied to thinner components. Explain why.

32.6. Explain the reasons that a variety of mechanical joining methods have been developed over the years.

32.7. List three brazing and three soldering techniques.

32.8. Describe the similarities and differences between the functions of a bolt and a rivet.

32.9. What precautions should be taken in the mechanical joining of dissimilar metals?

32.10. What difficulties are involved in joining plastics? Why?

32.11. What is the difference between a rivet and a bolt? What are the advantages of rivets?

32.12. What are the principles of (a) wave soldering and (b) reflow soldering?

32.13. What is a peel test? Why is it useful?

32.14. What is a combination joint?

32.15. What test methods are used to evaluate adhesives?

Qualitative Problems

32.16. Describe some applications in manufacturing for single-sided and double-sided adhesive tapes.

32.17. Explain how adhesives can be made to be electrically conductive.

32.18. Comment on your observations concerning the joints shown in Figs. 32.3, 32.6, 32.8, and 32.11.

32.19. Give examples of combination joints other than those shown in Fig. 32.13.

32.20. Discuss the need for fixtures for holding workpieces in the joining processes described in this chapter.

32.21. Explain why adhesively bonded joints tend to be weak in peeling.

32.22. It is common practice to tin-plate electrical terminals to facilitate soldering. Why is it tin that is used?

32.23. Give three applications where adhesive bonding is the best joining method.

32.24. How important is a close fit for two parts that are to be brazed? Explain.

32.25. If you are designing a joint that must be strong and also needs to be disassembled several times during the product's life, what kind of joint would you recommend? Explain.

32.26. Review Fig. 32.11 and explain why the examples under the "Poor," "Good," and "Very good" have these classifications.

32.27. Rate lap, butt, and scarf joints in terms of joint strength. Explain your answers.

32.28. What are the advantages of integrated snap fasteners?

32.29. List the advantages and disadvantages of mechanical fastening as compared with adhesive bonding.

32.30. List the joining methods that would be suitable for a joint that will encounter high stresses and will need to be disassembled several times during the product life, and rank the methods.

32.31. Loctite® is an adhesive used to keep metal bolts from vibrating loose; it basically glues the bolt to the nut once the nut is inserted in the bolt. Explain how this adhesive works.

Quantitative Problems

32.32. Refer to the simple butt and lap joints shown in Fig. 32.11. (a) Assuming the area of the butt joint is 3 mm $\times$ 20 mm and referring to the adhesive properties given in Table 32.3, estimate the minimum and maximum tensile force that this joint can withstand. (b) Estimate these forces for the lap joint assuming its area is 15 mm $\times$ 15 mm.

32.33. In Fig. 32.12a, assume that the cross-section of the lap joint is 20 mm $\times$ 20 mm, that the diameter of the solid rivet is 4 mm, and that the rivet is made of copper. Using the strongest adhesive shown in Table 32.3, estimate the maximum tensile force that this joint can withstand.

32.34. As shown in Fig. 32.15a, a rivet can buckle if it is too long. Referring to Chapter 14 on forging, determine the maximum length-to-diameter ratio of a rivet so that it would not buckle during riveting.

32.35. Figure 32.4 shows qualitatively the tensile and shear strength in brazing as a function of joint clearance. Search the technical literature, obtain data, and plot these curves quantitatively. Comment on your observations.

32.36. When manufacturing the fuselage of a commuter airplane, aluminum plates are glued together with lap joints. Because the elastic deformation for a single plate differs from the deformation for two plates glued together in a lap joint, the maximum shear stress in the glue is twice as high as the average shear stress. The shear strength of the glue is 20 MPa, the tensile strength of the aluminum plates is 95 MPa, and their thickness is 5 mm. Calculate the overlapping length needed to make the glue joint twice as strong as the aluminum plate.

Synthesis, Design, and Projects

32.37. Examine various household products and describe how their components are joined and assembled. Explain why those particular processes were used and not others.

32.38. Name several products that have been assembled by (a) seaming, (b) stitching, and (c) soldering.

32.39. Suggest methods of attaching a round bar (made of a thermosetting plastic) perpendicularly to a flat metal plate. Discuss their advantages and limitations.

32.40. Describe the tooling and equipment that would be necessary to perform the double-lock seaming operation shown in Fig. 32.17, starting with a thin, flat sheet.

32.41. Prepare a list of design guidelines for joining by the processes described in this chapter. Would these guidelines be common to most processes? Explain.

32.42. What joining methods would be suitable for assembling a thermoplastic cover over a metal frame? Assume that the cover is removed periodically, as is the top of a coffee can.

32.43. Repeat Problem 32.42, but for a cover made of (a) a thermoset, (b) a metal, and (c) a ceramic. Describe the factors involved in your selection of methods.

32.44. Comment on workpiece size and shape limitations, if any, for each of the processes described in this chapter.

32.45. Describe part shapes that cannot be joined by the processes covered in this chapter. Give specific examples.

32.46. Give examples of products in which rivets in a structure or in an assembly may have to be removed and later replaced by new rivets.

32.47. Visit a hardware store and investigate the geometry of the heads of screws that are permanent fasteners—that is, fasteners that can be screwed in, but not out.

32.48. Obtain a soldering iron and attempt to solder two wires together. First, try to apply the solder at the same time as you first put the soldering iron tip to the wires. Second, preheat the wires before applying the solder. Repeat the same procedure for a cool surface and a heated surface. Record your results and explain your findings.

32.49. Perform a literature search to determine the properties and types of adhesives used to affix artificial hips onto the human femur.

32.50. Review Fig. 32.9a and explain the shortcoming in using a constant thickness beam instead of a tapered double cantilever beam.

32.51. Review Fig. 32.9 and carefully sketch the stress distributions you expect in each geometry.

32.52. Design a joint to connect two 25 mm wide, 5 mm thick steel members. The overlap may be as much as 25 mm, and any one approach described in this chapter can be used.

32.53. For the same members in Problem 32.52, design a joint using threaded fasteners arranged in one row. Do you advise the use of one large fastener or many small fasteners? Explain.

32.54. For the same members in Problem 32.52, design a joint using a *combination* of joining techniques.

PART VII
Surface Technology

Our first visual or tactile contact with the objects around us is through their *surfaces*: surface roughness, waviness, reflectivity, and various other features. The preceding chapters described the properties of materials and manufactured components, basically in terms of their *bulk* characteristics, such as strength, ductility, hardness, and toughness. Also included were some descriptions of the influences of surfaces on these properties, such as the effect of surface preparation on fatigue life and on joining processes, and the sensitivity of brittle materials to surface roughness, scratches, and various defects.

Machinery and their various accessories typically have numerous members that slide against each other: bearings, slideways, pistons and cylinders, and tools and dies for casting, machining, and forming operations. Close examination will reveal that some of these surfaces are

- Smooth, while others are rough
- Slide against each other, some at high relative speeds while others move slowly
- Lubricated, while others are dry
- Subjected to heavy loads, while others support light loads
- Subjected to elevated temperatures, while others are at room temperature.

In addition to its geometric features, a **surface** is comprised of a very thin layer on the bulk material; their mechanical, physical, chemical, and metallurgical properties depend not only on the material and its processing history but also on the environment to which they been exposed. Consequently, the surface of a manufactured part typically possesses properties and behavior that can be significantly different from those of its bulk.

Although the bulk material generally determines a component's overall mechanical properties, the component's surfaces directly influence the part's performance in (see Fig. VII.1):

- Appearance and geometric features of the part and their role in subsequent operations, such as welding, soldering, adhesive bonding, painting, and coating

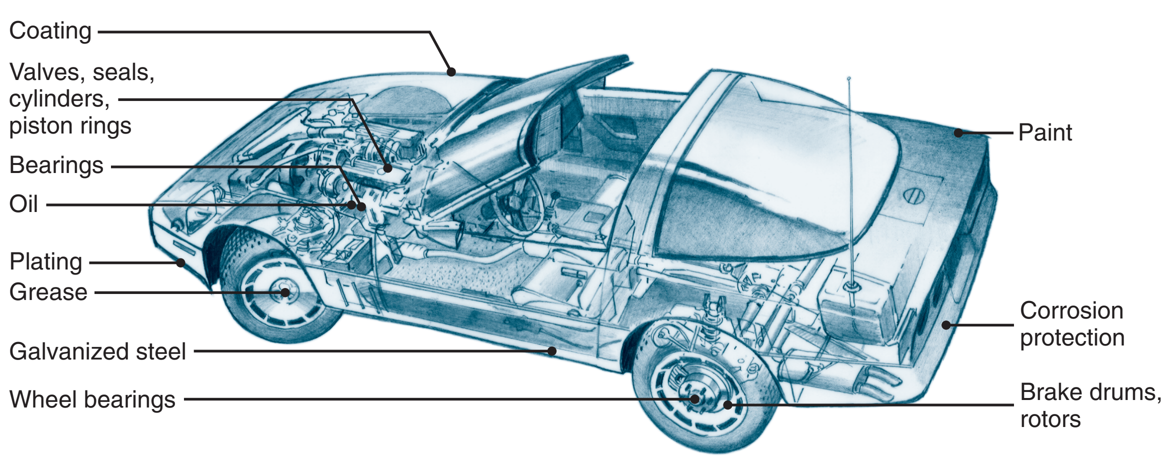


Figure VII.1: Components in a typical automobile that are related to the topics described in Part VII.

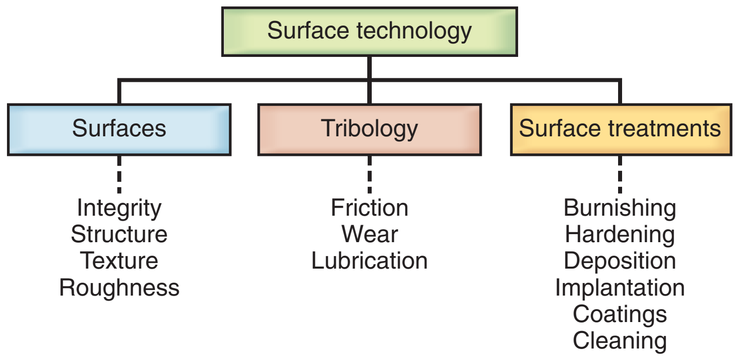


Figure VII.2: An outline of topics covered in Part VII.

- Resistance to corrosion
- Effectiveness of lubricants on parts while being made and throughout their service life
- Crack initiation and its propagation, due to surface defects, such as roughness, scratches, seams, and heat-affected zones, can lead to weakening and premature failure of a part
- Thermal and electrical conductivity of contacting bodies: rough surfaces have lower thermal and electrical conductivity than smooth surfaces
- Friction and wear of tools, molds, and dies used in manufacturing and of the products made.

Following the outline shown in Fig. VII.2, this chapter describes surface characteristics in terms of their structure and topography. The material and process variables that influence friction, wear, and lubrication are then described. Chapter 34 outlines the methods used to modify surfaces for better appearance, improved frictional behavior, effectiveness of lubricants, and resistance to wear and corrosion.

Chapter 33

Surface Roughness and Measurement; Friction, Wear, and Lubrication

Example:

- This chapter describes various features of surfaces that have a direct bearing on both the selection of manufacturing processes and the service life of the parts made.
- Surface features, such as roughness, texture, and lay, are explained, as well as the approaches used to quantitatively describe and measure surfaces.
- The chapter also examines the nature of friction, its role in manufacturing, and the factors influencing its magnitude.
- Wear and lubrication are then examined, along with various approaches to minimizing wear.
- The chapter ends with a summary of commonly used lubricants and their selection for a particular manufacturing process and for the materials involved.

33.1 Introduction

This chapter deals with surface phenomena. A surface is a distinct entity, with properties that can be significantly different from those of the bulk. Surfaces may have oxide layers, work-hardened layers, and contaminants of various types. Depending on the manner in which it was generated, a surface can have several physical defects that can have a major influence on the surface integrity of workpieces, tools, dies, and molds.

The chapter also describes those aspects of *friction*, *wear*, and *lubrication*, collectively known as **tribology**, that are relevant to manufacturing processes and operations and to the service life of products. Important topics are the nature of friction and wear for metallic and nonmetallic materials, and how they are influenced by various material and processing variables. Wear has a major economic impact, as it has been estimated that in the United States alone the total cost of replacing worn parts is more than $100 billion per year.

Finally, the chapter introduces the fundamentals of metalworking fluids, including the types, characteristics, and applications of commonly used liquid and solid lubricants and the lubrication practices employed, including the importance of biological and environmental considerations in their use, application, recycling, and ultimate disposal.

33.2 Surface Structure and Integrity

Upon close examination, it will be observed that the surface of a metal piece generally consists of several layers (Fig. 33.1):

1. The *bulk* metal, also known as **substrate**, has a structure that depends on the composition and processing history of the piece.

2. Above the bulk metal is a layer that usually has been deformed plastically and work hardened to a greater extent than the bulk. The depth and properties of this layer, called **surface structure**, depend on the processing method employed and the effects of frictional sliding on the surface. For example, if the surface has been produced by machining with a dull and worn tool (see Fig. 21.22), or it has been subjected to sliding against tools and dies, the *work-hardened layer* will be relatively thick, and usually will also develop residual stresses (Section 2.11).

3. Unless the metal is processed and kept in an inert (oxygen free) environment or it is a noble metal (gold or platinum), an **oxide layer** forms over the work-hardened layer. The oxide layer is generally

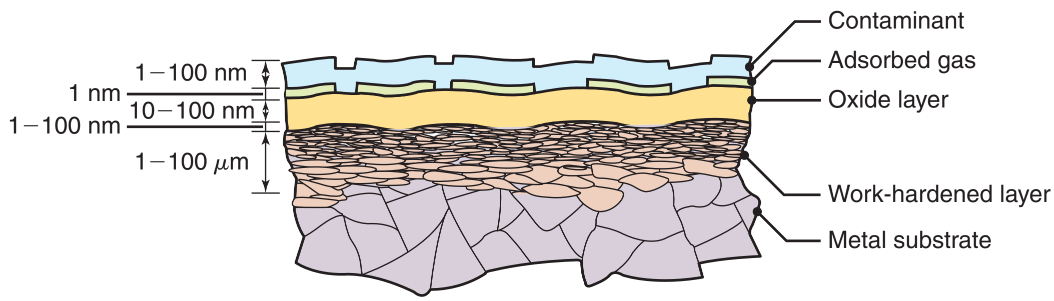


Figure 33.1: Schematic illustration of a cross-section of the surface structure of a metal. The thickness of the individual layers depends on both processing conditions and the processing environment. *Source:* After E. Rabinowicz and B. Bhushan.

much harder than the base metal, hence it is more abrasive. As a result, it has an important influence on friction, wear, and lubrication. Each metal forms its own unique oxide or oxides, and their behavior can be very complex. For example,

- *Iron* has a surface oxide structure, with FeO adjacent to the bulk metal, followed by a layer of Fe_3O_4, and then a layer of Fe_2O_3, exposed to the environment.
- *Aluminum* has a dense, amorphous surface layer of Al_2O_3, with a thick, porous, and hydrated aluminum-oxide layer over it.
- *Copper* has a bright, shiny surface when freshly scratched or machined. Soon after, however, it develops a Cu_2O layer, which is covered with a layer of CuO; the latter layer gives copper its somewhat dull color.
- *Stainless steels* are stainless because they develop a protective layer of chromium oxide by *passivation* (see Section 3.8).

4. Under normal environmental conditions, surface oxide layers are generally covered with *adsorbed* layers of gas and moisture.
5. Finally, the outermost surface of the metal may be covered with contaminants, such as dirt, dust, lubricant residues, cleaning-compound residues, and by pollutants from the environment.

The factors that pertain to the surface structures of the metals just described are also those in the surface structure of plastics and ceramics (Chapters 7 and 8). The surface texture of these materials also depends, as with metals, on their method of production.

Surface Integrity. *Surface integrity* describes not only the geometric (topological) features of surfaces and their physical and chemical properties, but also their mechanical and metallurgical properties and characteristics. Surface integrity is an important consideration in all design and manufacturing operations, because it influences such properties as fatigue strength, resistance to corrosion, and service life.

Several *surface defects*, caused by and produced during component manufacturing, can be responsible for inadequate surface integrity. These defects usually are caused by a combination of such factors as (a) existing defects in the raw or original material; (b) method or methods by which the surface has been produced; and (c) improper control of the processing parameters, which can result in excessive stresses, temperatures, or surface deformation.

The following list gives general definitions of major surface defects (in alphabetical order) found in practice:

- **Cracks** may be external or internal; those that require a magnification of $10\times$ or higher to be seen by the naked eye are called **microcracks**.
- **Craters** are shallow depressions.
- **Heat-affected zone** is that portion of a part that has been subjected to thermal cycling without melting, such as that shown in Fig. 30.19.
- **Inclusions** are nonmetallic, very small elements or compounds in the bulk of the material.
- **Intergranular attack** is the weakening of grain boundaries through liquid-metal embrittlement and corrosion (Section 1.5.2).
- **Laps, folds,** and **seams** are surface defects resulting from the overlapping of material during processing (see, for example, Fig. 14.17).
- **Metallurgical transformations** involve microstructural changes caused by temperature cycling of the material; these changes may consist of phase transformations, recrystallization, alloy depletion, decarburization, and molten and then recast, resolidified, or redeposited material.

- **Pits** are shallow surface depressions, usually the result of external chemical, electrical or physical attack, or fatigue wear.
- Surface **residual stresses** are caused by nonuniform deformation and nonuniform temperature distribution in the part (Section 2.11).
- **Splatter** is small, resolidified molten metal particles deposited on a surface, as during some welding operations.
- **Surface plastic deformation** is caused by high stresses due to factors such as friction, tool and die contact geometry with the workpiece, worn tools, and processing methods (see Fig. 21.21).

33.3 Surface Texture and Roughness

Regardless of the method of production, all surfaces have their own characteristics which, collectively, are referred to as *surface texture*. Although their description as a geometrical property is complex, several guidelines have been established for identifying surface texture in terms of well-defined and measurable quantities (Fig. 33.2):

- **Flaws** or **defects** are random irregularities, such as scratches, cracks, holes, depressions, seams, tears, and inclusions.
- **Lay** or **directionality** is the direction of the predominant surface pattern, usually visible to the naked eye.
- **Roughness** is defined as closely spaced, irregular deviations on a very small scale, expressed in terms of its height, width, and distance from each other along a surface.
- **Waviness** is a recurrent deviation from a flat surface; it is measured and described in terms of the distance between adjacent crests of the waves (*waviness width*) and the height between the crests and valleys of the waves (*waviness height*).

Surface roughness is generally characterized by two methods. The **arithmetic mean value** (**R_a**) is based on the schematic illustration of a rough surface, as shown in Fig. 33.3. It is defined as

$$R_a = \frac{a+b+c+d+\cdots}{n}, \tag{33.1}$$

where all coordinates a, b, c, ... are absolute values, and n is the number of readings. The *units* generally used for surface roughness are μm (microns).

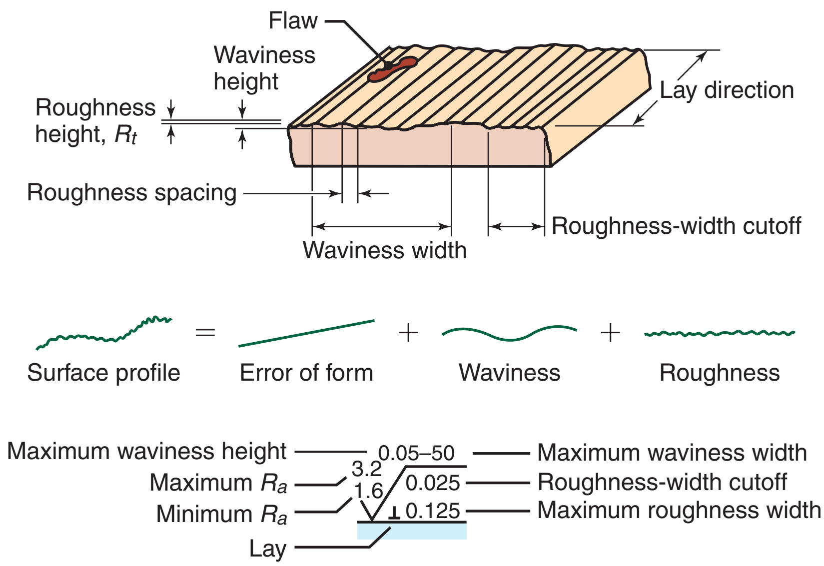
The **root-mean-square roughness** (**R_q**, formerly called RMS), is defined as

$$R_q = \sqrt{\frac{a^2+b^2+c^2+d^2+\cdots}{n}}. \tag{33.2}$$

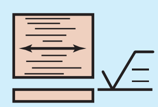
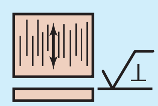
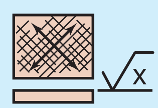
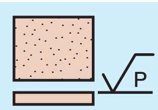
The datum line AB in Fig. 33.3 is located such that the sum of the areas above the line is equal to the sum of the areas below the line.

The **maximum roughness height** (**R_t**) is defined as the height from the deepest trough to the highest peak. It indicates how much material has to be removed in order to obtain a smooth surface, such as by polishing.

In general, a surface cannot be described by its R_a or R_q alone, because these values are averages. Two surfaces may have the same roughness value, but have actual topographies that are very different. For example, a few deep troughs on an otherwise smooth surface will not affect the roughness values



(a)

Lay symbol	Interpretation	Examples
=	Lay parallel to the line representing the surface to which the symbol is applied	
⊥	Lay perpendicular to the line representing the surface to which the symbol is applied	
X	Lay angular in both directions to line representing the surface to which symbol is applied	
P	Pitted, protuberant, porous, or particulate nondirectional lay	

(b)

Figure 33.2: (a) Standard terminology and symbols to describe surface finish. The quantities are given in μm. (b) Common surface lay symbols.

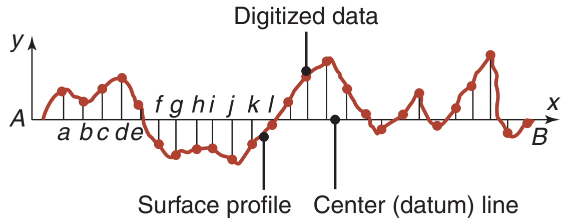


Figure 33.3: Coordinates used for surface-roughness measurement defined by Eqs. (33.1) and (33.2).

significantly. However, this type of surface profile can be significant in terms of friction, wear, and fatigue characteristics of a part. It is therefore important to analyze a surface in great detail, particularly for parts that are to be used in critical applications.

Measuring Surface Roughness. Instruments, called **surface profilometers**, typically are used to measure and record surface roughness. A profilometer has a *diamond stylus* that travels along a straight line over the surface (Fig. 33.4a), and periodically records height measurements. The distance the stylus travels is called **cutoff**, which generally ranges from 0.08 to 25 mm; for most engineering applications, a cutoff of 0.8 mm is typical. The rule of thumb is that the cutoff must be long enough to include 10 to 15 roughness irregularities.

In order to highlight roughness, profilometer traces are recorded on an exaggerated vertical scale (a few orders of magnitude larger than the horizontal scale (see Fig. 33.4c through f). The magnitude of the scale is called **gain** on the recording instrument. Thus, the recorded profile is distorted significantly, and the surface will appear to be much rougher than it actually is. The recording instrument also compensates for any surface waviness, and indicates only surface roughness.

Because the diamond stylus tip has a finite radius, the stylus path is different from the actual surface (note the path with the broken line in Fig. 33.4b), and the measured roughness is lower. The most common tip has a diameter of 10 μm. The smaller the stylus diameter and the smoother the surface, the closer is the path of the stylus to the actual surface profile.

Three-dimensional Surface Measurement. Because surface properties can vary significantly with the direction in which a profilometer trace is taken, there is often a need to measure *three-dimensional* surface profiles. In the simplest case, this can be done with a surface profilometer that has the capability of indexing a short distance between traces. A number of other alternatives have been developed, two of which are optical interferometers and atomic-force microscopes.

1. **Optical-interference microscopes** shine a light against a reflective surface and record the interference fringes that result from the incident and its reflected waves. This technique allows for a direct measurement of the surface slope over the area of interest. As the vertical distance between the sample and the interference objective is changed, the fringe patterns also change, thus allowing for a surface height measurement.

2. **Atomic-force microscopes** (AFMs) are used to measure extremely smooth surfaces and in some arrangements have the capability of distinguishing atoms on atomically smooth surfaces. In principle, an AFM is merely a very fine surface profilometer with a laser that is used to measure probe position. The surface profile can be measured with high accuracy and with vertical resolution on the atomic scale, and scan areas can be on the order of 100 μm square, although smaller areas are more common.

Surface Roughness in Engineering Practice. Requirements for surface roughness in typical engineering applications vary by as much as two orders of magnitude (Fig. 33.5). Some examples are:

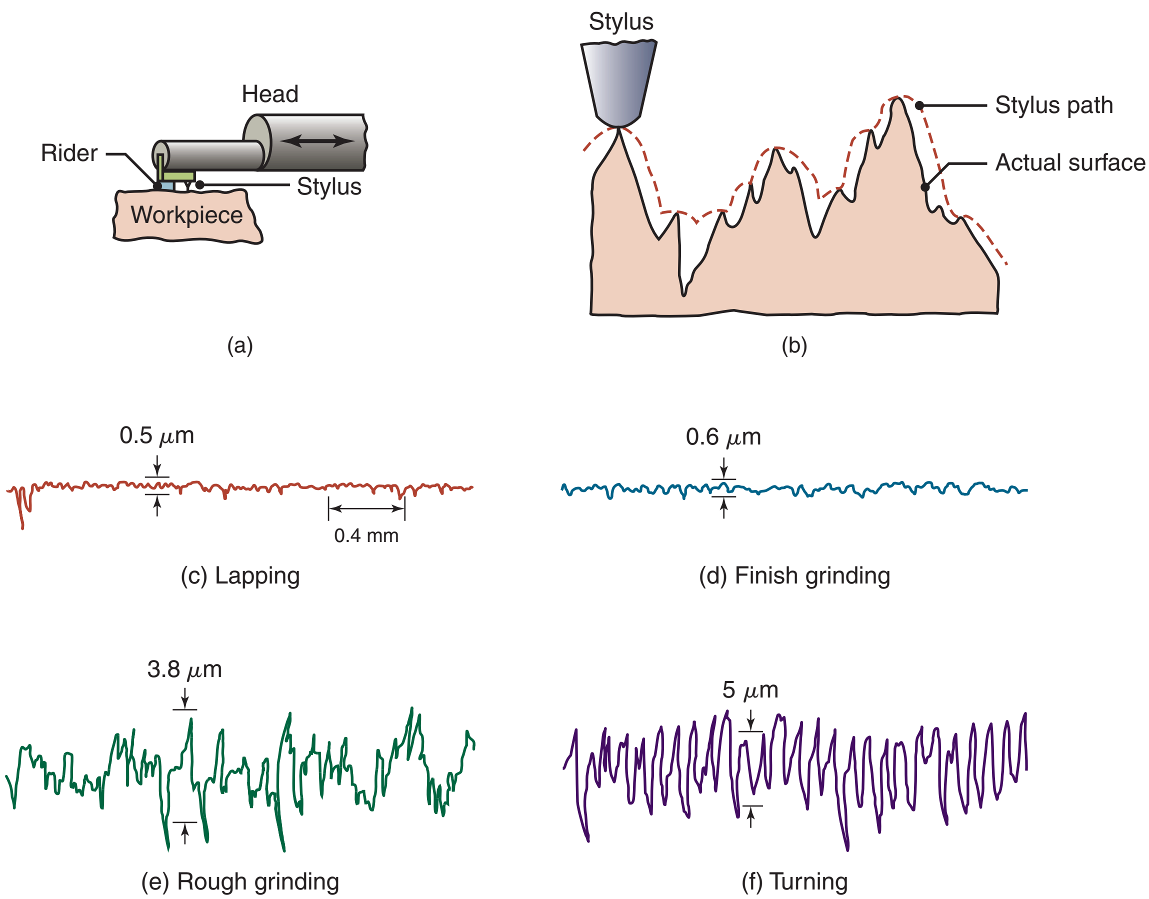


Figure 33.4: (a) Measuring surface roughness with a stylus. The rider supports the stylus and guards against damage. (b) Path of the stylus in surface-roughness measurements (broken line), compared with the actual roughness profile. Note that the profile of the stylus path is smoother than that of the actual surface. (c) through (f) Typical surface profiles produced by various machining and surface-finishing processes. Note the difference between the vertical and horizontal scales.

- Bearing balls 0.025 μm
- Crankshaft bearings 0.032 μm
- Brake drums 1.6 μm
- Clutch-disk faces 3.2 μm
- Gage blocks and precision instruments 0.02 μm.

33.4 Friction

Friction plays an important role in manufacturing processes, because of the relative motion and the friction forces that are always present at tool, die, and workpiece interfaces. Friction (a) *dissipates energy*, generating heat, which can have detrimental effects and (b) *impedes free movement* at interfaces, significantly influencing

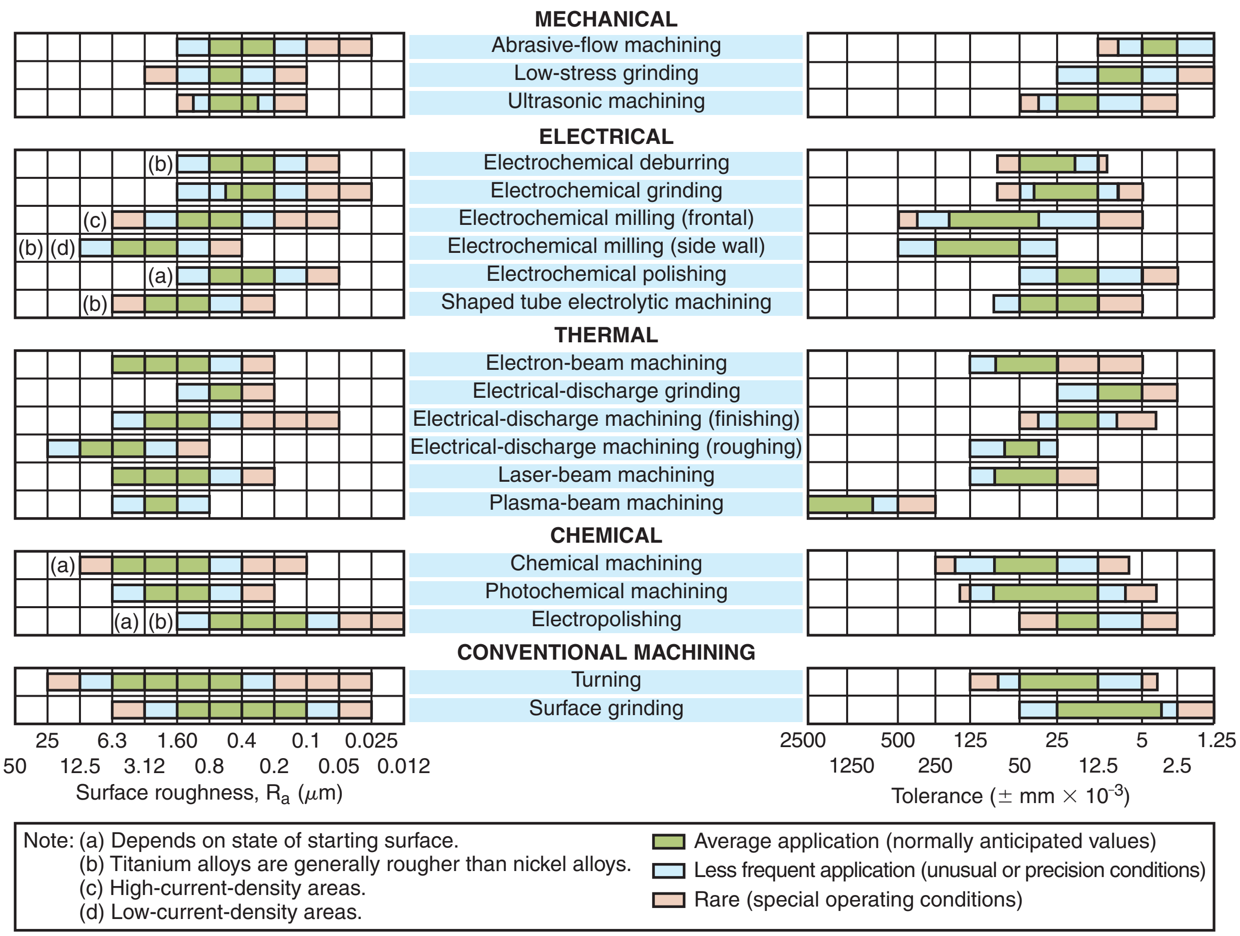


Figure 33.5: Surface roughness and tolerances obtained in various machining processes; note the wide range within each process (see also Fig. 23.14). *Source: Machining Data Handbook*, 3rd ed. Copyright 1980. Used by permission of Metcut Research Associates, Inc.

the flow and deformation of materials in metalworking processes. However, friction is not necessarily undesirable; for example, it would be impossible to roll metals, control material flow in forming and shaping operations, clamp workpieces on machines, or hold drill bits in chucks without friction.

There have been several theories to explain the phenomenon of friction. A commonly accepted theory of friction is the **adhesion theory**; it is based on the observation that two clean and dry surfaces, regardless of how smooth they are, contact each other at only a fraction of their apparent contact area (Fig. 33.6). The maximum slope of real surfaces range typically from 5° to 15°, unless they are purposely made to have high roughness. In such a situation, the normal (contact) load, N, is supported by minute **asperities**, very small projections from the two surfaces in contact with each other. The normal stresses at these asperities are therefore high, and thus can cause *plastic deformation* at the junctions, creating an *adhesive bond*. In other words, the asperities form *microwelds*, and it takes a certain force to shear the microweld. The cold pressure welding process (Section 31.2), for example, is based on this principle.

Another theory of friction is the **abrasion theory**: it is based on the notion that an asperity from a hard surface, such as a tool or a die, *penetrates* and *plows* a softer surface (see also Section 26.3). **Plowing** will cause displacement of the material and/or produce small slivers (chips), as in filing. In both situations described above, sliding between two bodies in contact will require a *tangential force*; this force is the **friction force**, F.

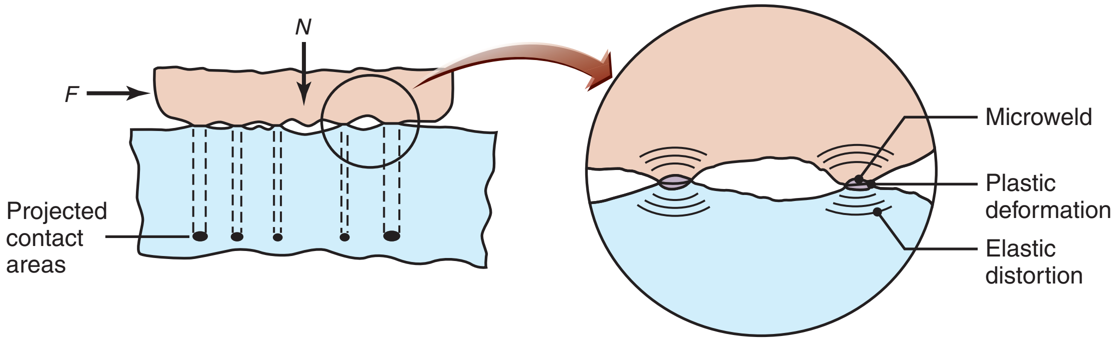


Figure 33.6: Schematic illustration of the interface of two bodies in contact showing real areas of contact at the asperities. In engineering surfaces, the ratio of the apparent-to-real areas of contact can be as high as 4 to 5 orders of magnitude.

The ratio F/N (Fig. 33.6a) is the **coefficient of friction**, μ. Depending on the materials and processes involved, μ in manufacturing varies significantly. For example, in metal-forming processes (Part III), it typically ranges from about 0.03 in cold working to about 0.7 in hot working, and from about 0.5 to as much as 2 in machining operations.

Friction of Plastics. Because their strength is low as compared to metals (Tables 2.2 and 7.1), hard *plastics* generally possess low frictional characteristics, including polyimides, polyesters, and fluorocarbons (*Teflon*). This property can make plastics better than metals for bearings, gears, seals, and general friction-reducing applications, provided that the loads are not high. Because of this characteristic, polymers are sometimes described as *self lubricating*. On the other hand, soft plastics and rubbers generally conform to surfaces, leading to high friction.

The factors involved in the friction of metals are generally also applicable to polymers.The plowing component of friction in thermoplastics and elastomers is significant, because of their viscoelastic behavior (i.e., they exhibit both viscous and elastic behavior) and subsequent hysteresis loss (Fig. 7.14). This condition can easily be simulated by dragging a dull nail across the surface of rubber, and observing how the rubber quickly recovers its shape.

An important factor in applications of plastics is the effect of temperature rise at the sliding interfaces caused by friction. As described in Section 7.3, thermoplastics rapidly lose their strength and become soft as temperature increases. Thus, if the temperature rise is not controlled, the sliding surfaces can undergo permanent deformation and thermal degradation. The frictional behavior of various polymers on metals is similar to that of metals on metals. The well-known low friction of teflon has been attributed to its molecular structure, which has no reactivity with metals; consequently, its adhesion and thus friction are low.

Friction of Ceramics. The mechanics of friction for ceramics is similar to that of metals; thus, adhesion and plowing at interfaces contribute to the friction force in ceramics as well. Usually, however, adhesion is less important with ceramics because of their chemical inertness and high hardness (Fig. 2.15), whereby the real area of contact at sliding interfaces is small. Abrasion and plowing can be significant, especially if ceramics interface with softer materials.

Reducing Friction. Friction can be reduced mainly through the (a) selection of materials that have low adhesion, such as carbides and ceramics; and (b) application of surface films and coatings. *Lubricants*, such as oils, or solid films, such as graphite, interpose an adherent layer between the tool, die, and workpiece, which minimizes adhesion and the interactions between two sliding bodies. Friction also can be reduced significantly by mechanical means, by subjecting the tool- or die-workpiece interface to *ultrasonic vibrations*,

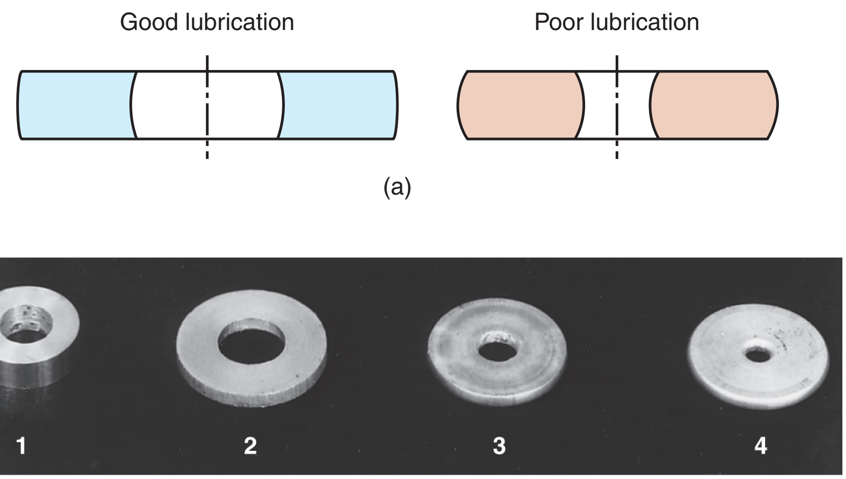


Figure 33.7: Ring-compression test between flat dies. (a) Effect of lubrication on type of ring-specimen barreling. (b) Test results: (1) original specimen and (2) to (4) increasing friction. *Source:* After A.T. Male and M.G. Cockcroft.

typically at 20 kHz. The amplitude of the vibrations periodically separates the two surfaces, allowing the lubricant to flow more freely into the interface.

Friction Measurement. The coefficient of friction usually is determined experimentally, either during an actual manufacturing operation or in simulated laboratory tests, using small-scale specimens of various shapes. A test that has gained wide acceptance, particularly for bulk-deformation processes (Chapters 13 to 15), is the **ring-compression test**. A flat ring is upset plastically between two flat platens (Fig. 33.7a). As its height is reduced, the ring expands radially outward (volume constancy). If friction at the interfaces is zero, both the inner and the outer diameters of the ring expand as if it were a solid disk. With increasing friction, however, the internal diameter becomes smaller. For a particular reduction in height, there is a critical friction at which the internal diameter increases from its original diameter if μ is lower, and it decreases if μ is higher (Fig. 33.7b).

By measuring the change in the specimen's internal diameter and using the curves shown in Fig. 33.8 (obtained through theoretical analyses), the coefficient of friction can be determined. Note that each ring geometry and each material has its own specific set of curves. The most common geometry of a specimen has an outer diameter/inner diameter/height proportion of 6:3:2. The actual size of the specimen is usually not relevant in these tests. Thus, once the percentage of reduction in internal diameter and height is known, the magnitude of μ can be determined using the appropriate chart.

Example 33.1 Determination of Coefficient of Friction

Given: In a ring-compression test, a specimen 10 mm in height and with an outside diameter (OD) of 30 mm and an inner diameter (ID) of 15 mm is reduced in thickness by 50%.

Find: Determine the coefficient of friction, μ, if the OD is 38.9 mm after deformation.

Solution: First it is necessary to determine the new ID (which is obtained from volume constancy) as follows:

$$\text{Volume} = \frac{\pi}{4}\left(30^2 - 15^2\right)(10) = \frac{\pi}{4}\left(38.9^2 - \text{ ID}^2\right)(5).$$

From this equation, the new ID is calculated as 12.77 mm. Thus, the change in internal diameter is

$$\Delta\text{ID} = \frac{12.77 - 15}{15} = -0.1487 \text{ or } 14.87\% \text{ (decrease)}$$

With a 50% reduction in height and a 14.87% reduction in internal diameter, the friction coefficient can be obtained from Fig. 33.8 as μ=0.09.

33.5 Wear

The importance of *wear* is evident in the number of parts and components that continually have to be replaced or repaired in a wide variety of consumer and commercial products. *Wear plates*, placed in dies and sliding mechanisms where the loads are high, are an important component in some metalworking machinery. These plates, also known as *wear parts*, are expected to wear, but they can easily be replaced, and thus prevent more costly repairs.

Although wear generally alters a part's surface topography, its surface finish, and it may result in severe surface damage, it can also have a beneficial effect. The *running-in* period for engines produces

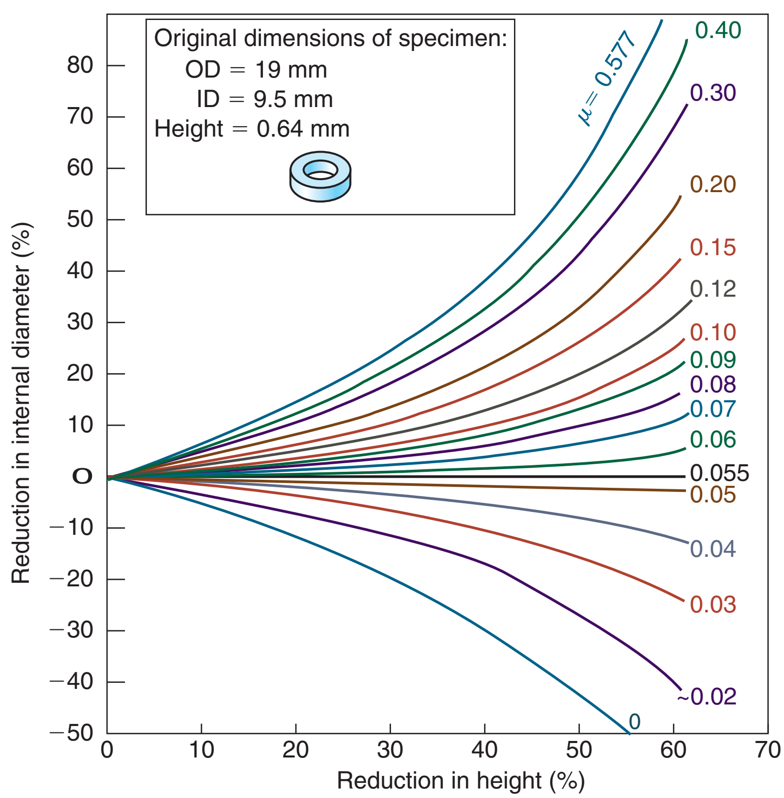


Figure 33.8: Chart to determine friction coefficient from a ring-compression test. Reduction in height and change in internal diameter of the ring are measured; then μ is read directly from this chart. For example, if the ring specimen is reduced in height by 40% and its internal diameter decreases by 10%, the coefficient of friction is 0.10.

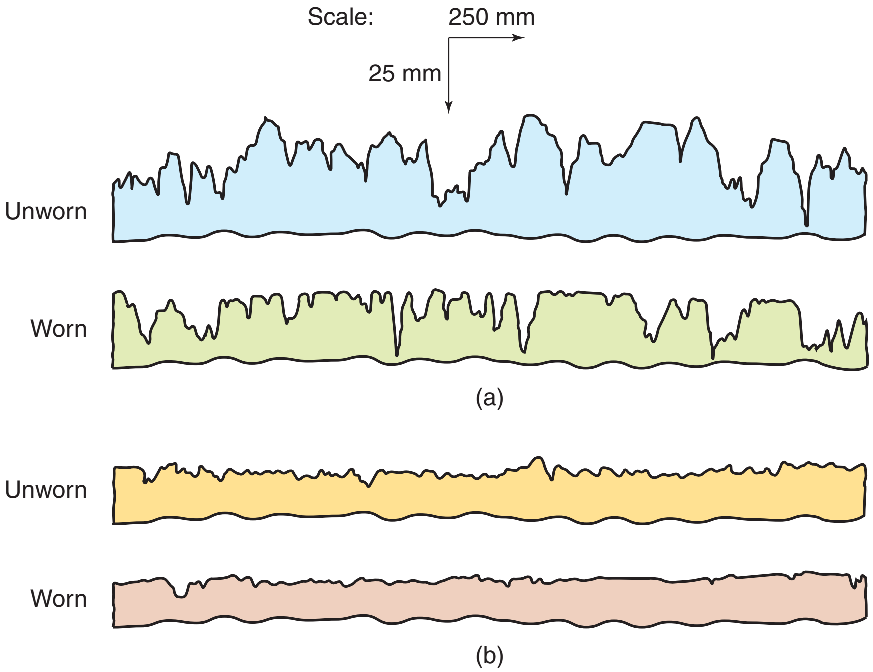


Figure 33.9: Changes in original (a) wire-brushed and (b) ground-surface profiles after wear. Note the difference in the vertical and horizontal scales. *Source:* After E. Wild and K.J. Mack.

small particles of wear while removing the peaks from asperities, as can seen in Fig. 33.9. Under controlled conditions, wear can thus be regarded as a type of smoothing or polishing process.

Described below are basic wear mechanisms relevant to manufacturing operations.

Adhesive Wear. If a tangential force is applied to the model shown in Fig. 33.10, shearing can take place either (a) at the original interface of the two bodies or (b) along a path below or above the interface. Sliding causes *adhesive wear*, also called *sliding wear*. Because of such factors as strain hardening at the asperity contacts, diffusion between the two bodies, and mutual solid solubility (Section 4.3) of the materials in contact, the adhesive bonds formed at the asperity junctions often are stronger than the base metals themselves. During sliding, fracture usually follows a path in the weaker or softer component, generating a wear fragment. Although this fragment is typically attached to the harder component (the upper surface in Fig. 33.10c), it eventually becomes detached during further rubbing at the interface and develops into a loose **wear particle**.

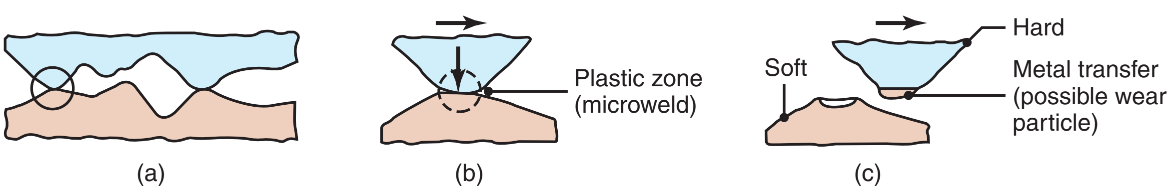


Figure 33.10: Schematic illustration of (a) two contacting asperities, (b) adhesion between two asperities, and (c) the formation of a wear particle.

In more severe conditions, such as ones with high normal loads and strongly bonded asperities on clean surfaces, adhesive wear is described as *scuffing, smearing, tearing, galling,* or *seizure*; these are called **severe wear**. Oxide layers on surfaces have a major influence on adhesive wear, at times acting as a *protective film*, resulting in **mild wear**, consisting of small wear particles.

Adhesive wear can be reduced by one or more of the following methods:

1. Selecting materials that do not develop strong adhesive bonds
2. Using a harder material as one member of the pair
3. Using materials that oxidize more readily
4. Applying hard coatings (see Chapter 34)
5. Coating one surface with a softer, hence with low shear strength, material, such as tin, silver, lead, or cadmium
6. Using an appropriate lubricant.

Abrasive Wear. This type of wear is caused by a hard, rough surface, or a surface containing protruding hard particles, sliding across another surface. As a result, *microchips* or *slivers* are produced, as wear particles, and leave grooves or scratches on the softer surface (Fig. 33.11). Such processes as filing, grinding, ultrasonic machining, and abrasive-jet and abrasive water-jet machining act in this manner.

There are two basic types of abrasive wear. In **two-body wear**, abrasive action takes place between two sliding surfaces or between loose abrasive particles and a solid body. This type of wear is the basis of **erosive wear**, such as occurs from the movement of slurries through pipes or sand particles impacting a ship's propeller. In **three-body wear**, abrasive particles are present between two sliding solid bodies, including wear particle or a hard contaminant carried by a lubricant. This situation indicates the importance of periodic filtering lubricants in metalworking operations, machinery, and in automotive, aircraft, and helicopter engines.

The **abrasive-wear resistance** of pure metals and ceramics has been found to be directly proportional to their hardness. Abrasive wear can therefore be reduced by increasing the hardness of the materials involved (usually by heat treating, Chapter 4) or by reducing the normal load. Elastomers and rubbers resist abrasive wear well, because they *deform elastically*, then recover when abrasive particles cross past over their surfaces. The best example is an automobile tire, constantly in contact with paved or unpaved road surfaces, which typically are rough and abrasive; even a highly hardened steel wheel would not last long under such severe conditions.

Corrosive Wear. Also known as *oxidation wear* or *chemical wear*, this type of wear is caused by chemical and electrochemical reactions between metal surfaces and the environment. Among corrosive media are water, seawater, oxygen, acids, chemicals, and atmospheric hydrogen sulfide and sulfur dioxide. The fine corrosive products on the surface become the wear particles. When the corrosive layer is destroyed or

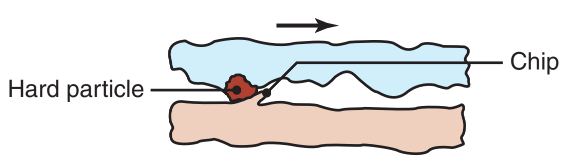


Figure 33.11: Schematic illustration of abrasive wear in sliding. Longitudinal scratches on a surface usually indicate abrasive wear.

removed through sliding, another layer begins to form, and the process of removal and corrosive-layer formation is repeated. Corrosive wear can be reduced by:

- Selecting materials that will resist environmental attack
- Applying a coating (Chapter 34)
- Controlling the environment
- Reducing operating temperatures in order to lower the rate of chemical reaction.

Fatigue Wear. Also called *surface fatigue* or *surface-fracture wear*, this type of wear is caused when surfaces are subjected to cyclic loading, such as rolling contact in bearings and gears or in forging operations (see also Section 2.7). The wear particles are usually formed through the mechanism of *spalling* or *pitting*. **Thermal fatigue** is another type of fatigue wear, whereby surface cracks are first generated by thermal stresses from thermal cycling, as when a cool die is repeatedly brought in contact with hot workpieces. The individual cracks then join each other, and the surface begins to spall, a phenomenon similar to the development of potholes on roads. Thermal fatigue results in *heat checking* of molds and dies in die casting and hot working operations.

Fatigue wear can be reduced by:

- Lowering contact stresses
- Reducing thermal cycling
- Improving the quality of materials by removing impurities, inclusions, and various other flaws that can act as local points for crack initiation and propagation.

Several other types of wear also can be observed in manufacturing operations:

- **Erosion**, caused by loose particles abrading a surface
- **Fretting corrosion**, when interfaces are subjected to very small reciprocal movements
- **Impact wear**, removal of very small amounts of material from a surface, through the impacting action of particles (similar to the mechanism of ultrasonic machining (Section 26.6)).

In many situations in manufacturing, component wear is the result of a *combination* of different types of wear. Note in Fig. 33.12, for example, that even in the same forging die, various types of wear take place in different locations of the die cavity. A similar situation also can exist in cutting tools, as shown in Fig. 21.18.

Wear of Thermoplastics. Wear mechanisms of thermoplastics are similar to that of metals. Their abrasive-wear behavior depends partly on the ability of the polymer to deform and recover elastically, as in rubber and elastomers. Typical polymers with good wear resistance are polyimides, nylons, polycarbonate, polypropylene, acetals, and high-density polyethylene. These polymers are either molded or machined to make gears, pulleys, sprockets, and similar mechanical components. Thermoplastics can be blended with internal lubricants (such as polytetrafluoroethylene, silicon, graphite, molybdenum disulfide, and rubber particles) that are interspersed within the polymer matrix (see Section 7.5).

Wear of Reinforced Plastics. The wear resistance of reinforced plastics depends on the type, amount, and orientation of fiber reinforcements in the polymer matrix (see Chapter 9). Carbon, glass, and aramid fibers all improve wear resistance. Wear in these materials usually takes place when the fibers are pulled away from the matrix, called *fiber pullout*. Wear is highest when the sliding direction is parallel to the fibers, because they can then be pulled out more easily. Long fibers increase the wear resistance of composites, because they (a) are more difficult to pull out and (b) prevent cracks in the matrix from propagating to the surface.

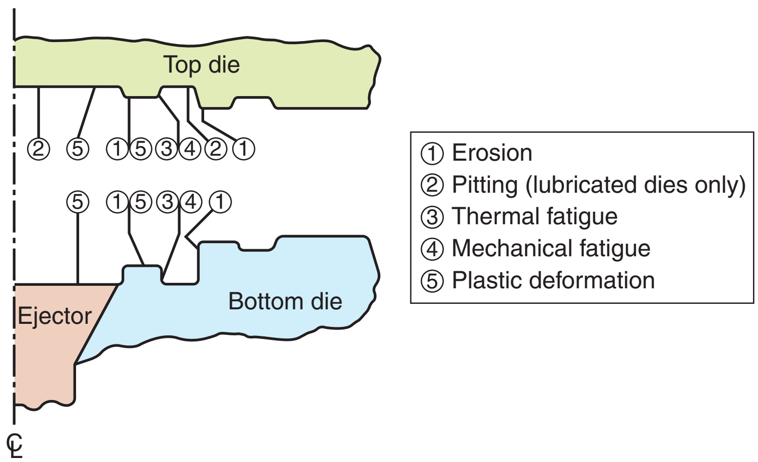


Figure 33.12: Types of wear observed in the cavity of a single pair of dies used for hot forging. *Source:* After T.A. Dean.

Wear of Ceramics. When ceramics slide against metals, wear is typically caused by (a) small-scale deformation and surface fracture, (b) plowing, (c) fatigue, and (d) surface chemical reactions. While sliding along each other, material from the surface of a metal body can be transferred to the oxide-type ceramic surface, forming metal oxides. Thus, sliding actually takes place between the metal and the metal-oxide surface.

33.6 Lubrication

Lubrication to reduce friction and wear dates back four millennia; Egyptian chariot wheels, for example, were lubricated with beef tallow in 1400 B.C. A variety of oils were used for lubrication in metalworking operations, beginning in about 600 A.D. (see Table I.2).

As noted in various chapters, the surfaces of tools, dies, molds, and workpieces typically are subjected to (a) *force* and *contact pressure*, both ranging from very low to multiples of the yield stress of the workpiece material; (b) *relative speed*, from very low to very high; and (c) *temperature*, which generally ranges from ambient to melting point. In addition to selecting appropriate materials and controlling process parameters to reduce friction and wear, **lubricants**, also called **metalworking fluids**, are widely applied.

Regimes of Lubrication. There are basically four regimes of lubrication of interest in manufacturing operations (Fig. 33.13):

1. **Thick-film lubrication:** The two surfaces are separated *completely* by a film of lubricant, thus lubricant viscosity is a major factor. Such films can develop in some regions of the workpiece in high-speed operations and from using high-viscosity lubricants that become *trapped* at dieworkpiece interfaces. The result is a dull, grainy surface appearance on the workpiece after a forming operation, the degree of roughness varying with grain size of the workpiece. Also, in such operations as coining and precision forging (Section 14.4), trapped lubricants are undesirable because they prevent the blank from completely filling the die cavity.

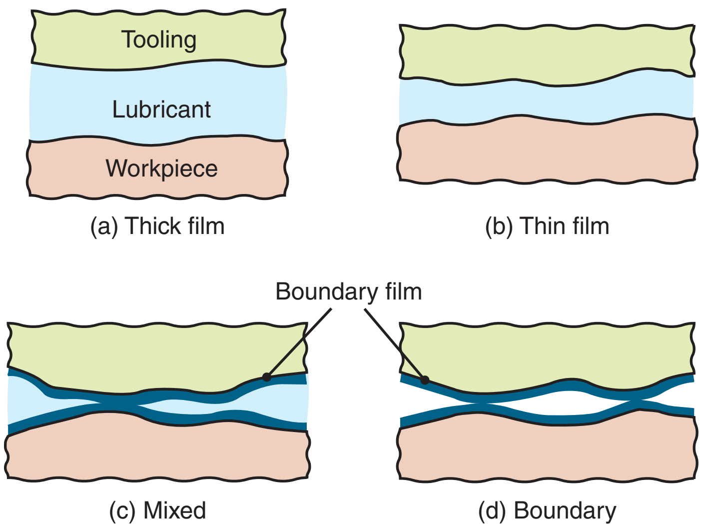


Figure 33.13: Regimes of lubrication generally occurring in metalworking operations. *Source:* After W.R.D. Wilson.

2. **Thin-film lubrication:** As the load between the die and the workpiece increases, or as the speed and viscosity of the metalworking fluid decrease, the lubricant film becomes thinner, known as **thin-film lubrication**. This condition raises the friction at interfaces and results in slight wear.

3. **Mixed lubrication:** In *mixed* or *partial lubrication*, a significant portion of the load is carried by the physical contact between the asperities of the two contacting surfaces. The rest of the load is carried by the pressurized fluid film trapped in pockets, such as the valleys between asperities.

4. **Boundary lubrication:** The load is supported by the contacting surfaces that are covered with a *boundary film* of lubricant, which is a lubricant layer that can be as thin as a single molecule (Fig. 33.13d). The film is attracted to the metal surfaces, preventing direct metal-to-metal contact of the two bodies, thus reducing wear. Boundary lubricants typically are natural oils, fats, fatty acids, esters, or soaps. The films can *break down* as a result of (a) *desorption*, caused by high temperatures developed at the sliding interfaces; or (b) by being rubbed off during sliding. Deprived of this protective film, the sliding metal surfaces then begin to wear and possibly severely score the surface.

Various Considerations. Note that the valleys in the surface of the contacting bodies (see Figs. 33.2a, 33.4, and 33.6) can serve as local *reservoirs* or *pockets* for retaining lubricants, thereby supporting a substantial portion of the load. The workpiece, but not the die, should have the rougher surface; as the asperities plastically deform, the workpiece is flattened by the tooling and the lubricant is released, or *percolated* from the surface. If the harder die surface is rough, there is no percolation effect, and the asperities, acting like a file, may damage the workpiece surface.

The recommended surface roughness on most dies is about 0.4 μm. The overall *geometry* of the interacting bodies also is an important consideration in ensuring proper lubrication. The movement of the workpiece into the deformation zone, as occurs during wire drawing, extrusion, and rolling, should allow a supply of lubricant to be carried into the dieworkpiece interface.

33.7 Metalworking Fluids and Their Selection

The functions of a *metalworking fluid* are to:

- *Reduce friction*, thus reducing force and energy requirements and any rise in temperature
- *Reduce wear*, thus reducing seizure and galling
- *Improve material flow* in tools, dies, and molds
- Act as a *thermal barrier* between the workpiece and its tool and die surfaces, thus preventing workpiece cooling in hot-working processes
- Act as a *release* or *parting agent*, a substance that helps in the removal or ejection of parts from dies and molds.

Several types of metalworking fluids are available, with diverse chemistries, properties, and characteristics that could fulfill these requirements (see also Section 22.12).

33.7.1 Oils

Oils maintain high film strength on surfaces, as can be observed when trying to clean an oily surface. Although they are very effective in reducing friction and wear, oils have low thermal conductivity and low specific heat. Consequently, they do not effectively conduct away the heat generated by friction and plastic deformation during processing. Moreover, it is difficult and costly to remove oils from component surfaces that are later to be painted or welded, and it is difficult to dispose of them (see Section 34.16).

The sources of oils are: (a) **mineral** (*petroleum* or *hydrocarbon*), (b) **animal**, or (c) **vegetable**. Oils may be **compounded** with additives or with other oils. Compounding changes such properties as viscosity-temperature behavior, surface tension, heat resistance, and boundary-layer characteristics.

33.7.2 Emulsions

An *emulsion* is a mixture of two immiscible liquids, usually oil and water, in various proportions, along with additives. *Emulsifiers* are substances that prevent the dispersed droplets in a mixture from joining each other, hence the term immiscible. Milky in appearance, emulsions are also known as **water-soluble oils** or **water-based coolants**. They are of two types: (a) *Indirect emulsion*, where water droplets are dispersed in the oil; and (b) *direct emulsion*, where mineral oil is dispersed in water, in the form of very small droplets. Direct emulsions are important metalworking fluids, because the presence of water gives them high *cooling* capacity. They are effective particularly in high-speed machining (Section 25.5), where a severe temperature rise can have detrimental effects on tool life, surface integrity of workpieces, and dimensional accuracy of parts machined.

33.7.3 Synthetic and Semisynthetic Solutions

Synthetic solutions are *chemical* fluids that contain inorganic and other chemicals dissolved in water; they do not contain mineral oils. The chemical agents are added to impart various properties. Semisynthetic solutions are basically *synthetic* solutions, to which small amounts of emulsifiable oils have been added.

33.7.4 Soaps, Greases, and Waxes

Soaps typically are reaction products of sodium or potassium salts with fatty acids. Alkali soaps are soluble in water; other metal soaps generally are insoluble. Soaps are effective *boundary lubricants*; they can form thick film layers at die-workpiece interfaces, particularly when applied on conversion coatings for cold metalworking applications (Section 34.10).

Greases are solid or semisolid lubricants, generally consisting of soaps, mineral oil, and various additives. They are highly viscous and adhere well to metal surfaces. Although used extensively in machinery, greases are of limited use in manufacturing processes. *Waxes* may be of animal or plant (*paraffin*) origin. Compared with greases, they are less greasy and are more brittle. Waxes are of limited use in metalworking operations, except as lubricants for copper and, in the form of a chlorinated paraffin, for stainless steels and high-temperature alloys.

33.7.5 Additives

Metalworking fluids usually are *blended* with additives, including oxidation inhibitors, rust-preventing agents, foam inhibitors, wetting agents, and antiseptics.

Sulfur, *chlorine*, and *phosphorus* are important additives to oils. Known as **extreme-pressure** (EP) **additives**, and used either singly or in combination, they react chemically with metal surfaces, forming adherent surface films of metallic sulfides and chlorides. These films have low shear strength and good antiweld properties and thus can effectively reduce friction and wear. However, they also may preferentially attack the cobalt binder in tungsten-carbide tools and dies (through *selective leaching*), causing changes in the surface roughness and surface integrity of those tools (Section 22.4).

33.7.6 Solid Lubricants

Because of their unique properties and characteristics, several solid materials are used as lubricants in manufacturing operations.

Graphite. Graphite (Section 8.6) is weak in shear along its *basal planes* (see Fig. 1.4), thus it has low coefficient of friction in that direction. It can be an effective solid lubricant, particularly at elevated temperatures; however, friction is low only in the presence of air or moisture; otherwise, friction is very high; in fact, graphite can become abrasive. Graphite can be applied either by rubbing it on surfaces or by making it part of a *colloidal* (dispersion of small particles) suspension in a liquid carrier, such as water, oil, or alcohol.

Molybdenum Disulfide. A widely used lamellar solid lubricant, molybdenum disulfide (MoS_2) is somewhat similar in appearance to graphite. Unlike graphite, however, it has high friction coefficient in an ambient environment. It is used as a carrier for oils, typically applied by rubbing it on the workpiece surface.

Metallic and Polymeric Films. Because of their low strength, thin layers of soft metals and polymer coatings also are used as solid lubricants. Suitable metals include lead, indium, cadmium, tin, and silver; polytetrafluoroethylene, polyethylene, and methacrylates also are used. However, these coatings have limited applications, because of their lack of strength under high contact stresses, especially at elevated temperatures.

Soft metals are used to coat high-strength metals, such as steels, stainless steels, and high-temperature alloys. For example, copper or tin is chemically deposited on the surface of a metal before it is processed further. If the oxide of a particular metal has low friction, and is sufficiently thin, the oxide layer can serve as a solid lubricant, particularly at elevated temperatures (see also Section 15.3).

Glasses. Although it is a solid material, glass becomes viscous at elevated temperatures and thus serves as a liquid lubricant. Its viscosity is a function of temperature, but not of pressure (Section 8.4). Poor thermal conductivity also makes glass attractive, because it acts as a thermal barrier between hot workpieces and relatively cool dies. As a lubricant, it is typically used in such applications as hot extrusion and hot forging.

Conversion Coatings. Lubricants may not always adhere well to workpiece surfaces, particularly under high normal and shearing stresses. Failure to adhere has detrimental effects, especially in forging, extrusion, and the wire drawing of steels, stainless steels, and high-temperature alloys. For these applications, the workpiece surfaces are first transformed through a chemical reaction with acids, hence the term *conversion* (see also Section 34.10).

Although this reaction leaves a somewhat rough and spongy surface, this surface acts as a carrier for the lubricant. After treatment, any excess acid from the surface is removed, using borax or lime. A liquid lubricant, such as soap, is then applied to the surface; it adheres well to the surface and cannot be scraped off easily. *Zinc-phosphate* conversion coatings are often used on carbon and low-alloy steels. *Oxalate coatings* are used for stainless steels and high-temperature alloys.

33.7.7 Selection of Metalworking Fluids

Selecting a metalworking fluid for a particular application and workpiece material involves a consideration of several factors:

1. Specific manufacturing process
2. Workpiece material
3. Tool or die material
4. Processing parameters
5. Compatibility of the fluid with the tool and die materials and the workpiece
6. Surface preparation required
7. Method of applying the fluid
8. Removal of the fluid and cleaning of the workpiece after processing
9. Contamination of the fluid by other lubricants, such as those used to lubricate machinery
10. Storage and maintenance of fluids
11. Treatment of waste lubricant
12. Biological and environmental considerations
13. Costs involved in all of the factors listed above.

The specific function of a metalworking fluid, whether it is primarily a *lubricant* or a *coolant*, also must be taken into account. Water-based fluids are very effective coolants; however, as lubricants, they are not as effective as oils. It is estimated that water-based fluids are used in 80–90% of all machining operations.

Specific requirements for metalworking fluids are:

- They should not leave any harmful residues that could interfere with production operations
- Fluids should not stain or corrode the workpiece or the equipment
- Periodic inspection is necessary, to detect any deterioration caused by accumulation of oxides, chips, wear debris, bacterial growth, and general degradation, and breakdown due to temperature and time. Wear particles are particularly important, because they cause damage to the system; proper inspection and filtering are thus essential.

After completion of manufacturing operations, workpiece surfaces typically have residual lubricants; they should be removed prior to further processing, such as welding or painting. Oil-based lubricants are more difficult and expensive to remove than water-based fluids. Various cleaning solutions and techniques are described in Section 34.16.

Biological and Environmental Considerations. These considerations are important factors in the selection and use of metalworking fluids. Hazards include contacting or inhaling some of these fluids, such as *dermatitis*, inflammation of the skin, and long-term exposure to *carcinogens*. Improper disposal of metalworking fluids cause adverse effects on the environment as well. To prevent or restrict the growth of microorganisms, such as bacteria, yeasts, molds, algae, and viruses, chemicals (*biocides*) are added to metalworking fluids.

Much progress has been made in developing environmentally safe (*green*) fluids and the technology and the equipment for their proper treatment, recycling, and disposal. In the United States, for example, laws and regulations concerning the manufacture, transportation, use, and disposal of metalworking fluids are promulgated by the U.S. Occupational Safety and Health Administration (OSHA), the National Institute for Occupational Safety and Health (NIOSH), and the Environmental Protection Agency (EPA).

Summary

- In many applications, surfaces and their properties are as important as the bulk properties of materials. A surface not only has a particular shape, roughness, and appearance, but also has properties that can differ significantly from those of the bulk material.
- Surfaces are exposed to the environment, and thus are subject to environmental attack. They also may come into contact with tools and dies (during processing) or with other components (during their service life).
- Geometric and material properties of surfaces can affect the properties of friction, wear, fatigue, corrosion, and electrical and thermal conductivity.
- Measurement and description of surface features and their characteristics are important aspects of manufacturing. The most common surface-roughness measurement is the arithmetic mean value. The instruments usually used to measure surface roughness include profilometers, optical interferometers, and atomic force microscopes.
- Friction and wear are among the most significant factors in processing materials. Much progress has been made in understanding these phenomena and identifying the factors that govern them.
- Affinity and solid solubility of the two materials in contact, the nature of surface films, the presence of contaminants, and process parameters such as load, speed, and temperature are among important factors.
- A wide variety of metalworking fluids, such as oils, emulsions, synthetic solutions, and solid lubricants, is available for specific applications. Their selection and use requires a thorough consideration of several factors regarding the workpiece and die materials and the particular manufacturing process. Biological and environmental considerations also are important factors in their selection.

Key Terms

Abrasive wear
Additives
Adhesion
Adhesive wear
Arithmetic mean value
Asperities
Boundary lubrication
Coefficient of friction
Compounded oils
Conversion coatings
Coolant
Emulsion

Extreme-pressure additives
Fatigue wear
Flaw
Fretting corrosion
Friction force
Greases
Impact wear
Lay
Lubricant
Lubrication
Maximum roughness height
Metalworking fluids
Microwelds
Mixed lubrication
Oils
Oxide layer
Pit
Plowing
Ring-compression test
Root-mean-square roughness
Running-in
Selective leaching
Self lubricating
Severe wear
Soaps
Solid lubricants
Substrate
Surface defects
Surface integrity
Surface profilometer
Surface roughness
Surface structure
Surface texture
Thick-film lubrication
Thin-film lubrication
Tribology
Ultrasonic vibrations
Water-soluble oils
Waviness
Waxes
Wear
Wear parts

Bibliography

Abdelbari, A., **Wear of Polymers and Composites**, Woodhead Publishing, 2015.
ASM Handbook, Vol. 18, **Friction, Lubrication, and Wear Technology**, ASM International, 2017.
Astakhov, V.P., **Tribology of Metal Cutting**, Elsevier, 2012.
Barber, J.R., **Contact Mechanics**, Springer, 2018.
Basu, B., and Kalin, M., **Tribology of Ceramics and Composites**, Wiley-American Ceramic Society, 2011.
Bhushan, B., **Introduction to Tribology**, 2nd ed., Wiley, 2013.
Byers, J.P., **Metalworking Fluids**, 3rd ed., Marcel Dekker, 2017.
Davim, J.P., **Surface Integrity in Machining**, Springer, 2010.
Gohar, R., and Rahnejat, H. (eds.), **Fundamentals of Tribology**, 3rd ed., World Series Pub. Co., 2018.
Griffiths, B., **Manufacturing Surface Technology: Surface Integrity and Functional Performance**, Butterworth-Heinemann, 2001.
Hirani, H., **Fundamentals of Engineering Tribology, with Applications**, Cambridge, 2016.
Hutchings, I., and Shipway, P., **Tribology**, 2nd ed., Butterworth-Heinemann, 2017.
Leach, R., **Optical Measurements of Surface Topography**, Springer, 2011.
Rabinowicz, E., **Friction and Wear of Materials**, 2nd ed., Wiley, 1995.
Stachowiak, G.W., **Wear: Materials, Mechanisms and Practice**, Wiley, 2005.
Stachowiak, G.W., and Batchelor, A.W., **Engineering Tribology**, 4th ed., Butterworth-Heinemann, 2013.
Straffelini, G., **Friction and Wear**, Springer, 2015.
Yang, S., and Wenhui, L., **Surface Finishing Theory and New Technology**, Springer, 2017.

Review Questions

33.1. What is tribology?

33.2. Explain what is meant by (a) surface texture and (b) surface integrity.

33.3. List and explain the types of defects typically found on surfaces.

33.4. Define the terms (a) *roughness* and (b) *waviness*.

33.5. Explain why the results from a profilometer are not a true depiction of the actual surface.

33.6. Describe the features of the ring-compression test. Does it require the measurement of forces?

33.7. List the types of wear generally observed in engineering practice.

33.8. Define the terms *wear*, *friction*, and *lubricant*.

33.9. How can adhesive wear be reduced? Abrasive wear?

33.10. Explain the mechanisms through which a wear particle is formed from adhesive wear, and two- and three-body abrasive wear.

33.11. Explain the functions of a lubricant in manufacturing processes.

33.12. What is grease? What is an emulsion?

33.13. What is the role of additives in metalworking fluids?

33.14. Describe the factors involved in lubricant selection.

Qualitative Problems

33.15. Give several examples that show the importance of friction in manufacturing processes as described in Parts III and IV.

33.16. Explain the significance of the fact that the hardness of metal oxides is generally much higher than that of the base metals themselves. Give some examples.

33.17. What factors would you consider in specifying the lay of a surface for a part? Explain.

33.18. Explain why identical surface-roughness values do not necessarily represent the same type of surface.

33.19. Why are the requirements for surface-roughness design in engineering applications so broad? Explain with specific examples.

33.20. What is the significance of a surface-temperature rise resulting from friction? Give some examples based on topics covered in the preceding chapters.

33.21. Explain the causes of lay on surfaces.

33.22. Give several examples of how wear on molds, tools, and dies affects a manufacturing operation.

33.23. Comment on the surface roughness of various parts and components with which you are familiar. What types of parts exhibit the coarsest surface? What types exhibit the finest? Explain.

33.24. Give two examples in which waviness on a surface would be desirable. (b) Give two examples in which it would be undesirable.

33.25. Do the same as for Problem 33.20, but for surface roughness.

33.26. Describe your observations regarding Fig. 33.8.

33.27. Give the reasons that an originally round specimen in a ring-compression test may become oval after it is upset.

33.28. Explain why graphite and molybdenum disulfide are effective solid lubricants.

33.29. Explain the reason that the abrasive-wear resistance of a material is a function of its hardness.

33.30. On the basis of your own experience, make a list of parts and components that have to be replaced because of wear.

33.31. Explain why the types of wear shown in Fig. 33.12 occur in those particular locations in the forging die.

33.32. List the similarities and differences between adhesive and abrasive wear.

33.33. List the requirements of a lubricant.

33.34. List manufacturing operations in which high friction is desirable and those in which low friction is desirable.

33.35. List manufacturing operations in which high wear is desirable and those in which low wear is desirable.

33.36. Does the presence of a lubricant affect abrasive wear? Explain.

33.37. It is observed that the coefficient of friction between a carriage and the ways on a lathe is 0.35. To reduce friction and wet the surfaces, kerosene (a very low viscosity fluid) is applied to the interface. Instead of reducing the friction, it is now measured to be 0.38. Provide an explanation for these measurements.

Quantitative Problems

33.38. Refer to the profile shown in Fig. 33.3, and offer some reasonable numerical values for the vertical distances from the centerline. Calculate the R_a and R_q values. Then give another set of values for the same general profile and calculate the same two quantities. Comment on your observations.

33.39. Obtain several different parts made of various materials, inspect their surfaces under an optical microscope at different magnifications, and make an educated guess as to what manufacturing process or finishing process was likely used to produce each of these parts. Explain your reasoning.

33.40. A surface with a triangular sawtooth roughness pattern has a peak-to-valley height of 4 μm. Find the R_a and R_q values.

33.41. Refer to Fig. 33.7b, and make measurements of the external and internal diameters (in the horizontal direction in the photograph) of the four specimens shown. Remembering that in plastic deformation the volume of the rings remains constant, estimate (a) the reduction in height and (b) the coefficient of friction for each of the three compressed specimens.

33.42. Using Fig. 33.8, make a plot of the coefficient of friction versus the change in internal diameter for a constant reduction in height of 35%.

33.43. Assume that in Example 33.1 the coefficient of friction is 0.16. If all other parameters remain the same, what is the new internal diameter of the specimen?

Synthesis, Design, and Projects

33.44. List the steps you would follow if you wished to reduce friction in a manufacturing process.

33.45. Discuss the tribological differences between ordinary machine elements (such as gears, cams, and bearings) and metalworking processes using tools, molds, and dies. Consider such factors as load, speed, and temperature.

33.46. Section 33.2 listed major surface defects. How would you go about determining whether or not each of these defects is a significant factor in a particular application?

33.47. Describe your own thoughts regarding biological and environmental considerations in the use of metalworking fluids.

33.48. Wear can have detrimental effects in manufacturing operations. Can you visualize situations in which wear could be beneficial? Explain, and give some examples.

33.49. Many parts in various appliances and automobiles have to be replaced because they were worn. Describe the methodology you would follow in determining the type(s) of wear these components have undergone.

33.50. In the second paragraph of the introduction to Part VII, five different sets of interfacial conditions were outlined, from (a) to (e). For each of these, give several examples from the manufacturing processes described in this book.

33.51. Describe your thoughts on the desirability of integrating surface-roughness measuring instruments into the machine tools described in Parts III and IV? How would you go about doing so, giving special consideration to the factory environment in which they are to be used? Make some preliminary sketches of such a system.

33.52. On the basis of the topics discussed in this chapter, do you think there is a direct correlation between friction and wear of materials? Explain.

33.53. A current interest is the development of carbon-free lubricants, also known as *white lubricants*. Prepare a two-page paper on the current developments in white lubricants.

Chapter 34

Surface Treatments, Coatings, and Cleaning

Examples:

- As described throughout the preceding chapters, material and process selection are critical aspects of manufacturing; often, however, the surface properties of a part also determine its performance in service.
- This chapter describes the surface-finishing operations that can be performed on parts for technical and aesthetic reasons.
- The chapter presents surface treatment, cleaning, and coating processes commonly performed, and includes an outline of mechanical surface treatments, such as shot peening, laser peening, and roller burnishing, for imparting compressive residual stresses onto metal surfaces of parts made.
- Coating operations are then examined, including cladding, thermal spray operations, physical and chemical vapor deposition, ion implantation, and electroplating; the benefits of diamond and diamondlike carbon coatings are also introduced.

34.1 Introduction

After a part is made, some of its surfaces may have to be processed further to ensure that they have certain specific properties and characteristics. **Surface treatments** may be necessary in order to:

- *Improve resistance to wear, erosion, and indentation*, such as for machine-tool slideways (Figs. 23.2 and 35.1), shafts, rolls, cams, and gears
- *Reduce friction*, especially on sliding surfaces of tools, dies, bearings, and machine ways
- *Reduce adhesion*, such as for electrical contacts
- *Improve resistance to corrosion and oxidation* on sheet metals for appliances, gas-turbine components, food packaging, and medical devices
- *Improve fatigue resistance* of bearings and shafts with fillets
- *Rebuild surfaces* on worn tools, dies, molds, and machine components
- *Modify surfaces*, their appearance, dimensional accuracy, and frictional characteristics
- *Impart decorative features*, such as texture and color.

A wide variety of techniques are employed to impart these characteristics to metallic, nonmetallic, and ceramic materials. The mechanisms involved include (a) plastic deformation of the workpiece surfaces, (b) chemical reactions, (c) thermal treatment, (d) deposition, (e) implantation, and (f) organic coatings and paints. Some of these techniques also are used in making semiconductor devices (Chapters 28 and 29).

34.2 Mechanical Surface Treatments

Several approaches are used to mechanically improve the surface properties of manufactured parts and components; the more common methods are described below.

Shot Peening. In this process, the workpiece surface is impacted repeatedly with cast steel, glass, or ceramic balls (called *shot*), which make overlapping indentations on the surface. Using shot sizes that range from 0.125 to 5 mm in diameter, this action causes plastic deformation of surfaces, to depths up to 1.25 mm. Because the plastic deformation is not uniform throughout the part's thickness (see also Fig. 2.14c), shot peening causes compressive residual stresses on the surface, thus improving the fatigue life of the component. Unless the processing parameters are well controlled, the deformation can be so severe as to cause damage to the surface being worked. The extent of surface deformation can be reduced by *gravity peening*, which involves larger shot sizes, but fewer number of impacts.

Shot peening is used extensively on shafts, gears, springs, oil-well drilling equipment, and turbine and compressor blades. Note, however, that if these parts are later subjected to high temperatures, such as gas turbine blades, the residual stresses will begin to relax (*thermal relaxation*) and their beneficial effects will be diminished.

Laser Shot Peening. In this process, also called *laser shock peening*, first developed in the mid-1960s but not commercialized until much later, the workpiece surface is subjected to *pulses* (planar laser shocks) from high-power lasers. This peening process produces compressive residual-stress layers that are typically 1 mm deep, with less than 1% of cold working taking place on the surface.

Laser shot peening has been applied successfully and reliably to jet-engine fan blades and to materials such as titanium, nickel alloys, and steels, for improved fatigue resistance and some corrosion resistance. Laser intensities are on the order of 100–300 J/cm^2 and have a pulse duration of 10–50 ns. Because they are now solid state, their cost is much lower.

Water-jet Peening. In this process, a water jet, at pressures as high as 400 MPa, impinges on the workpiece surface, inducing compressive residual stresses and surface and subsurface hardening at the same level as in shot peening. Water-jet peening has been used successfully on steels and aluminum alloys. The control of processing variables, such as jet pressure, jet velocity, nozzle design, and its distance from the surface, is important to avoid development of excessive surface roughness or surface damage.

Ultrasonic Peening. This process uses a hand tool that vibrates by a piezoelectric transducer, at a frequency of 22 kHz. A variety of heads can be used for different applications.

Roller Burnishing. Also called *surface rolling*, the surface of the component is cold worked by the action of a hard and highly polished roller or set of rollers. The process is used on flat, cylindrical, or conical surfaces (Fig. 34.1); it improves surface finish by removing scratches, tool marks, and pits, and induces compressive surface residual stresses. Consequently, corrosion resistance is improved, since corrosive products and residues cannot be entrapped. In a variation of this process, called *low-plasticity burnishing*, the roller travels only once over the surface, inducing minimal plastic deformation.

Internal cylindrical surfaces of holes also can be burnished by a process called **ballizing** or **ball burnishing**. In this process, a smooth ball, slightly larger than the bore diameter, is pushed through the length of the hole.

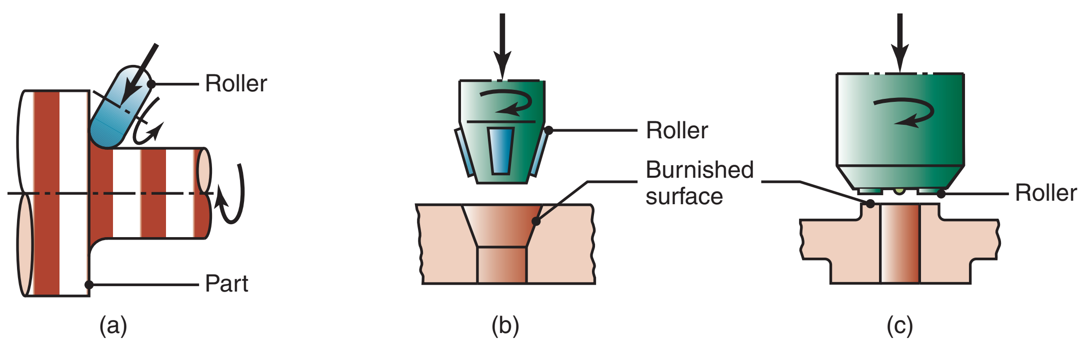


Figure 34.1: Burnishing tools and roller burnishing of (a) the fillet of a stepped shaft to induce compressive surface residual stresses for improved fatigue life; (b) a conical surface; and (c) a flat surface.

Typically used on hydraulic-system components, seals, valves, spindles, and fillets on shafts, roller burnishing improves mechanical properties as well as surface finish. It can be used either by itself or in combination with other finishing processes, such as grinding, honing, and lapping. The finishing operation is done after burnishing the part, in order to produce a smooth surface. The equipment can be mounted on CNC machine tools for improved productivity and consistency of performance. All types of soft or hard metals can be roller burnished.

Explosive Hardening. In this process, the surfaces are subjected to high pressures through detonating a layer of an explosive sheet placed directly on the workpiece surface. Contact pressures developed can be as high as 35 GPa, lasting about 2–3 μs. Significant increases in surface hardness can be achieved, with very little change (less than 5%) in the shape of the component. Railroad rail surfaces, for example, are often explosively hardened.

34.3 Mechanical Plating and Cladding

Mechanical Plating. In this process, also called *mechanical coating*, *impact plating*, or *peen plating*, fine metal particles are compacted over workpiece surfaces by glass, ceramic, or porcelain beads, propelled by rotary means, such as tumbling. This process, which is basically cold welding particles onto a surface, is typically used for hardened-steel parts, with plating thickness typically less than 25 μm.

Cladding. Also called *clad bonding*, parts are bonded with a thin layer of corrosion-resistant metal, through the application of pressure by rolls or other means (see Fig. 31.1). A typical example is cladding of aluminum (*Alclad*), in which a pure or corrosion-resistant layer of aluminum alloy is clad over an aluminum-alloy body (core). The cladding layer is anodic to the core, and usually has a thickness less than 10% of the total thickness of the part.

Examples of cladding are 2024 aluminum clad with 1230 aluminum, and 3003, 6061, and 7178 aluminum clad with 7072 aluminum; other applications include steels clad with stainless-steel or nickel alloys. The cladding material may also be applied using dies, as in cladding steel wire with copper, or with explosives. Multiple-layer cladding is also utilized in special applications.

Laser cladding involves fusion of a wire or powder material over a substrate. It has been successfully applied to metals and ceramics, particularly for enhanced friction and wear behavior of the components.

34.4 Case Hardening and Hard Facing

Surfaces also may be hardened by thermal means in order to improve their friction and wear properties, as well as their resistance to indentation, erosion, abrasion, and corrosion. The most common methods are:

Case Hardening. Traditional methods of case hardening (*carburizing, carbonitriding, cyaniding, nitriding, flame hardening*, and *induction hardening*) are described in Section 4.10 and summarized in Table 4.1. In addition to common heat sources, such as gas or electricity, an electron beam or a laser beam also can be used as a heat source, for both metals and ceramics. Case hardening, as well as several other surface-treatment processes described in this chapter, induces compressive residual stresses on surfaces, such as by the formation of martensite.

Hard Facing. In this process, a relatively thick layer, edge, or point of wear-resistant hard metal is deposited on a surface by fusion-welding techniques (Chapter 30). Several layers, known as *weld overlay*, can be deposited. Hard facing enhances the wear resistance of the materials; thus it is used in making tools, dies, and various industrial components.

Spark Hardening. Hard coatings of tungsten, chromium, or molybdenum carbides can be deposited by an electric arc, in a process variously called *spark hardening*, *electric spark hardening*, or *electrospark deposition*. The deposited layer is typically 250 μm thick. Hard-facing alloys can be used as electrodes, rods, wires, or powder. Typical applications include valve seats, oil-well drilling tools, and dies for hot metalworking.

34.5 Thermal Spraying

Thermal spraying is a series of processes in which coatings of various metals, alloys, carbides, ceramics, and polymers are deposited on metal surfaces by a spray gun, with a stream heated by an oxyfuel flame, an electric arc, or a plasma arc. The earliest applications of thermal spraying, in the 1910s, involved metals, hence the term **metallizing**. The surfaces to be sprayed are first cleaned of oil and dirt, then roughened by, for example, grit blasting, to improve their bond strength (Section 26.8). The coating material can be in the shape of wire, rod, or powder; when the droplets or particles impact the workpiece, they solidify and bond to the surface.

Particle velocities typically range from 150 to 1000 m/s, but can be higher for special applications. Temperatures are in the range of 3000°–8000°C. The sprayed coating is hard and wear resistant, with a layered structure of deposited material; however, the coating can have porosity as high as 20% due to entrapped air and oxide particles. Bond strength depends on the particular process and techniques used; it is mostly mechanical in nature, hence the importance of roughening the surface prior to spraying, but can also be metallurgical. Bond strength generally ranges from 7 to 80 MPa, depending on the particular process used.

Typical applications of thermal spraying include aircraft engine components (such as in rebuilding worn parts), storage tanks, tank cars, rocket motor nozzles, and components that require resistance to wear and corrosion. In an automobile, thermal spraying is often applied to crankshafts, valves, fuel-injection nozzles, piston rings, and engine blocks. The process is also used in gas and petrochemical industries for repairing worn parts and restoring dimensional accuracy to parts that may have not been machined or shaped correctly.

The source of energy in thermal-spraying processes is of two types: combustion and electrical.

1. **Combustion Spraying**
   - **Thermal wire spraying** (Fig. 34.2a): The oxyfuel flame melts the wire and deposits it on the surface. The bond is of medium strength, and the process is relatively inexpensive.
   - **Thermal metal-powder spraying** (Fig. 34.2b): This process is similar to thermal wire spraying, but uses a metal powder instead of wire.
   - **Detonation gun**: Controlled and repeated explosions take place by means of an oxyfuel-gas mixture. The detonation gun has a performance similar to that of plasma.
   - **High-velocity oxyfuel-gas spraying** (HVOF): This process has characteristics that are similar to that of the detonation gun, but is less expensive.
2. **Electrical Spraying**
   - **Twin-wire arc:** An arc is formed between two consumable wire electrodes. The resulting bond has good strength, and the process is the least expensive.
   - **Plasma:** Either conventional, high-energy, or vacuum (Fig. 34.2c) plasma produces temperatures on the order of 8300°C and results in good bond strength with very low oxide content. **Low-pressure plasma spray** (LPPS) and **vacuum plasma spray** both produce coatings with high bond strength and with very low levels of porosity and surface oxides.

Cold Spraying. The particles to be sprayed are at a lower temperature and are not melted; thus, oxidation is minimal. The spray jet in cold spraying is narrow and highly focused; it has very high impact velocities, thereby improving the bond strength of the particles on the surface.

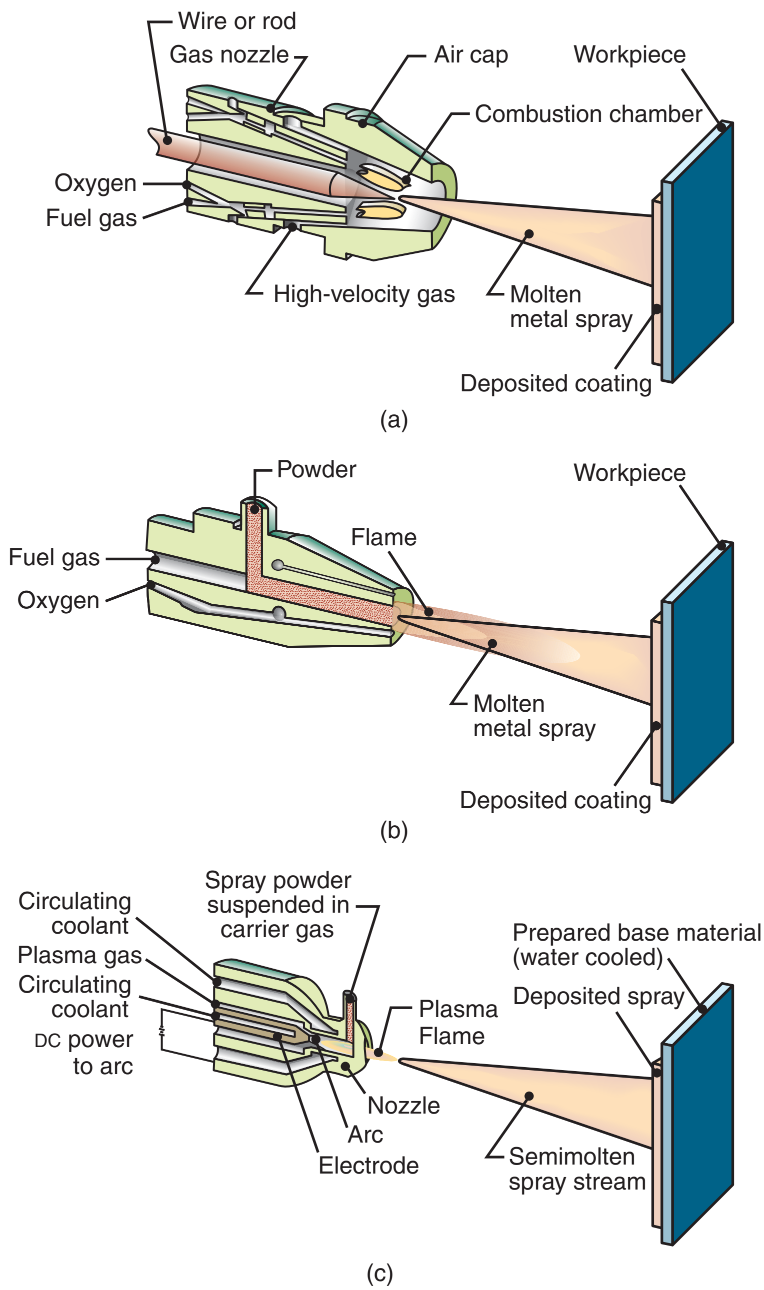


Figure 34.2: Schematic illustrations of thermal-spray operations: (a) thermal wire spray, (b) thermal metal-powder spray, and (c) plasma spray.

34.6 Vapor Deposition

Vapor deposition is a process in which a workpiece surface (substrate) is subjected to chemical reactions, by gases containing chemical compounds of the material to be deposited. The coating thickness is usually a few microns, much less than the thicknesses that result from the techniques described in Sections 34.2 and 34.3. The substrate may be metal, plastic, glass, or paper, and the deposited material may consist of metals, alloys, carbides, nitrides, borides, ceramics, or oxides. Control of coating composition, its thickness, and porosity is important. Typical applications for vapor deposition are coating cutting tools, drills, reamers, milling cutters, punches, dies, and wear surfaces.

There are two major vapor-deposition processes:

34.6.1 Physical Vapor Deposition

The three basic types of *physical vapor deposition* (PVD) processes are (a) vacuum deposition, or arc evaporation; (b) sputtering; and (c) ion plating. These processes are performed in a high vacuum and at temperatures in the range from 200° to 500°C. In PVD, the particles to be deposited are carried physically to the workpiece, rather than by chemical reactions, as in chemical vapor deposition.

Vacuum Deposition. In vacuum deposition, or evaporation, the metal is evaporated at a high temperature in a vacuum and is deposited on the substrate, which usually is at room temperature, or slightly higher for improved bonding. Coatings with uniform thickness can be deposited, even on parts with complex shapes. In **arc deposition** (PV/ARC), the coating material (cathode) is evaporated by several arc evaporators (Fig. 34.3), using highly localized electric arcs. The arcs produce a highly reactive plasma, consisting of the ionized vapor of the coating material; the vapor condenses on the substrate (anode), coating it. Applications of this process are both functional (oxidation-resistant coatings for high-temperature applications, electronics, and optics) and decorative (hardware, appliances, and jewelry). In **pulsed-laser and electron-beam deposition**, the energy beams heat the target into a vapor.

Sputtering. In this process, an electric field ionizes an inert gas (usually argon); the positive ions then bombard the coating material (cathode), causing sputtering (ejection) of its atoms. The atoms condense on the workpiece, which is heated to improve bonding (Fig. 34.4). In **reactive sputtering**, the inert gas is replaced by a reactive gas (such as oxygen), in which case the atoms are oxidized and are deposited. Carbides and nitrides also are deposited by this process. Alternatively, very thin polymer coatings can be deposited on metal and polymeric substrates with a reactive gas, causing polymerization of the plasma. **Radio-frequency** (RF) sputtering is used for nonconductive materials, such as electrical insulators and semiconductor devices.

Ion Plating. *Ion plating* is a generic term, describing a variety of combined processes of sputtering and vacuum evaporation. Basically, an electric field causes a glow, generating a plasma (Fig. 34.5); the vaporized

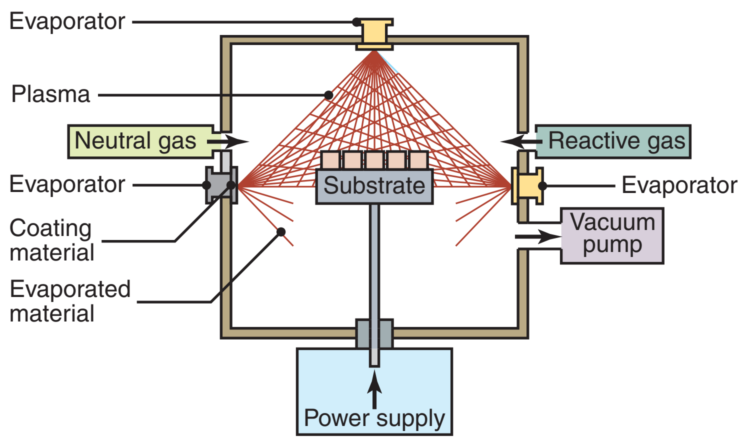


Figure 34.3: Schematic illustration of the physical-vapor-deposition process. Note that there are three arc evaporators and the parts to be coated are placed on a tray inside the chamber.

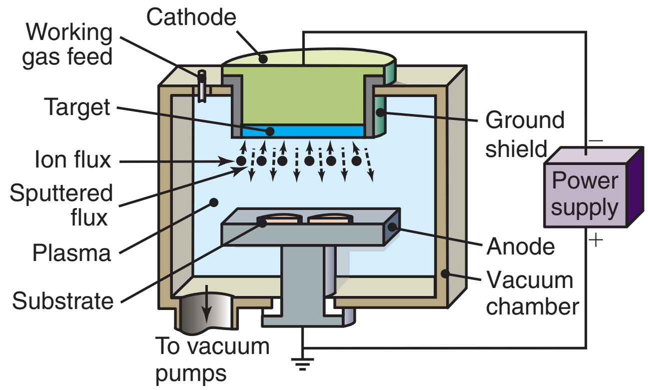


Figure 34.4: Schematic illustration of the sputtering process.

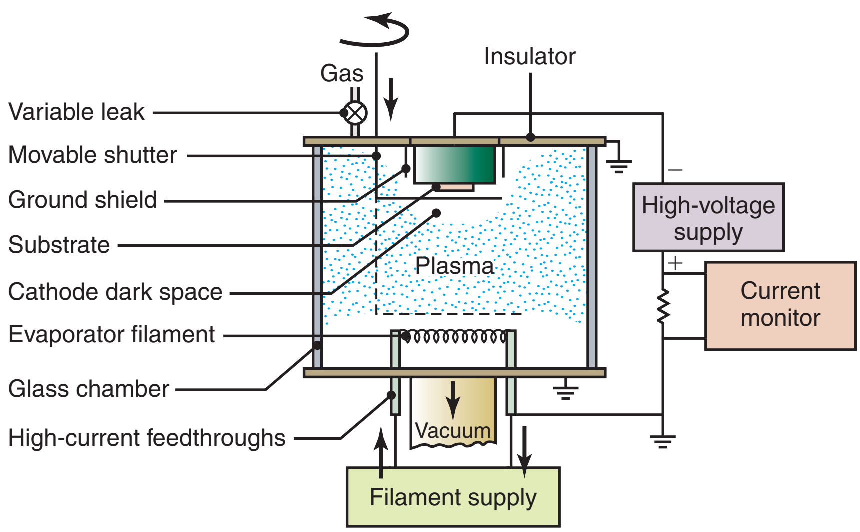


Figure 34.5: Schematic illustration of an ion-plating apparatus.

atoms are ionized only partially. **Ion-beam-enhanced (assisted) deposition** is capable of producing thin films, as coatings for semiconductors and tribological, and optical applications. Bulky parts can be coated in large chambers, using high-current power supplies of 15 kW and at voltages of 100,000 DC. **Dual ion-beam deposition** is a hybrid coating technique, combining PVD and simultaneous ion-beam bombardment, resulting in good adhesion on metals, ceramics, and polymers. Ceramic bearings and dental instruments are examples of its applications.

34.6.2 Chemical Vapor Deposition

Chemical vapor deposition (CVD) is a *thermochemical* process (Fig. 34.6). In a typical application, such as coating cutting tools with titanium nitride (Section 22.5), the tools are first placed on a graphite tray and heated to 950° to 1050°C, at atmospheric pressure and in inert atmosphere. Titanium tetrachloride (a gas), hydrogen, and nitrogen are then introduced into the chamber. The chemical reactions deposit titanium nitride on tool surfaces, with hydrogen chloride that is produced exhausted from the reaction

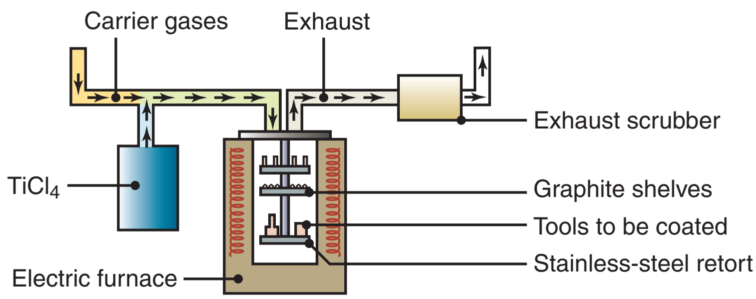


Figure 34.6: Schematic illustration of the chemical-vapor-deposition process; note that parts and tools to be coated are placed on trays inside the chamber.

chamber. Because of its toxicity, however, the exhaust gas must be cleaned, using exhaust scrubbers, before being vented to the atmosphere. For a coating of titanium carbide, methane is substituted for the other gases.

CVD coatings usually are thicker than those obtained with PVD. A typical cycle is long, consisting of (a) three hours of heating, (b) four hours of coating, and (c) six to eight hours of cooling to room temperature. The thickness of the coating depends on temperature, time, and the flow rates of the gases used. Almost any material can be coated and any material can serve as a substrate, although bond strength will vary. This process is also used to produce diamond coatings without binders, unlike polycrystalline diamond films which use 1 to 10% binder materials. The **medium-temperature CVD** (MTCVD) technique results in a higher resistance of the coating to crack propagation than CVD.

34.7 Ion Implantation and Diffusion Coating

In *ion implantation*, ions (charged atoms) are introduced onto the workpiece surface. The ions are accelerated in a vacuum to such an extent that they penetrate the substrate to a depth of a few microns. Ion implantation (not to be confused with ion plating, Section 34.6.1) modifies surface properties, by increasing surface hardness and improving resistance to friction, wear, and corrosion. The process can be controlled accurately, and the surface can be masked to prevent ion implantation in unwanted locations in a part.

This process is particularly effective on such materials as aluminum, titanium, stainless steels, tool and die steels, carbides, and chromium coatings. The process is typically used for cutting and forming tools, dies and molds, and metal prostheses, such as artificial hips and knees. For specific applications, such as semiconductors (Section 28.3), ion implantation is called **doping**, that is alloying with small amounts of various elements.

Diffusion Coating. This is a process in which an alloying element is diffused into the surface of the substrate (usually steel), altering its surface properties. The alloying elements can be supplied in solid, liquid, or gaseous states. This process has acquired different names, depending on the diffused element, as shown in Table 4.1, listing various diffusion processes such as *carburizing*, *nitriding*, and *boronizing*.

34.8 Laser Treatments

As described in various chapters of this book, lasers are having increasingly wider use, such as in machining, forming, joining, additive manufacturing, and metrology, as well as in surface engineering (laser peening, alloying, surface treatments, and texturing). Powerful, efficient, reliable, and less expensive lasers are widely available for cost-effective surface treatments, as outlined in Fig. 34.7.

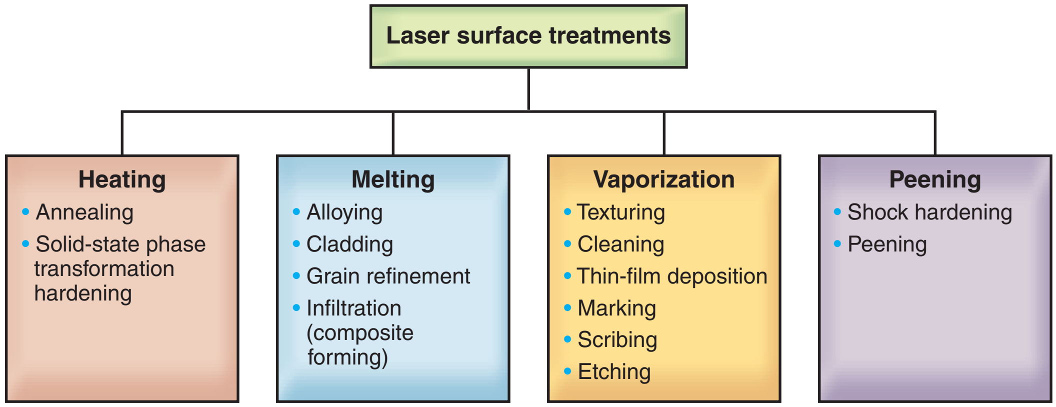


Figure 34.7: An outline of laser surface-engineering processes. *Source:* After N.B. Dahotre.

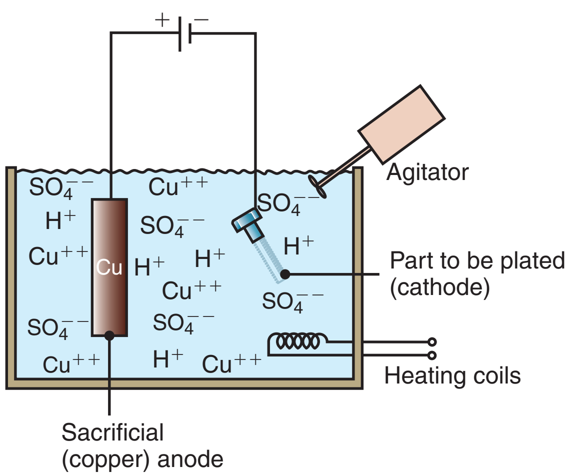
Example 34.1 Applications of Laser Surface Engineering

Several applications of lasers in engineering practice are given in this example. The most commonly used lasers are Nd:YAG and CO_2; excimer lasers are generally used for surface texturing (see also Table 27.2).

1. *Localized surface hardening*
   - Cast irons: diesel-engine cylinder liners, automobile steering assemblies, and camshafts
   - Carbon steels: gears and electromechanical parts
2. *Surface alloying*
   - Alloy steels: bearing components
   - Stainless steels: diesel-engine valves and seat inserts
   - Tool and die steels: dies for forming and die casting
3. *Cladding*
   - Alloy steels: automotive valves and valve seats
   - Superalloys: turbine blades
4. *Ceramic coating*
   - Aluminum-silicon alloys: automotive-engine bore
5. *Surface texturing and laser polishing*
   - Metals, plastics, ceramics, and wood: all types of products

34.9 Electroplating, Electroless Plating, and Electroforming

Plating imparts resistance to wear, resistance to corrosion, high electrical conductivity, better appearance, and reflectivity.



(a)

(b)

Figure 34.8: (a) Schematic illustration of the electroplating process. (b) Examples of electroplated parts. *Source:* Courtesy of Shutterstock/Jarous.

Electroplating. In electroplating, the workpiece (cathode) is plated with a different metal (anode), transferred through a water-based electrolytic solution (Fig. 34.8). Although the plating process involves a number of reactions, the process basically consists of the following sequence:

1. Metal ions from the anode are discharged by means of the potential energy from an external source of electricity or are delivered in the form of metal salts.

2. Metal ions are dissolved into the solution and are deposited over the cathode.

The volume of the plated metal can be calculated from the equation

$$\text{Volume} = cIt, \tag{34.1}$$

where I is the current in amperes, t is time, and c is a constant that depends on the plated metal, the electrolyte, and the efficiency of the system; typically, it is in the range of 0.03–0.1 mm^3/amp-s. It can be noted that for the same volume of material deposited, the deposited thickness is inversely proportional to the surface area. The deposition rate is typically on the order of 75 μm/h, thus electroplating is a slow process. Thin-plated layers are typically on the order of 1 μm; for thick layers, the plating can be as much as 500 μm.

The *plating solutions* are either strong acids or cyanide solutions. As the metal is being plated from the solution, it has to be periodically replenished. This is accomplished through two principal methods: (a) salts of metals are occasionally added to the solution or (b) a *sacrificial anode* of the metal to be plated is used in the electroplating tank and dissolves at the same rate that the metal is deposited.

There are three basic methods of electroplating:

1. **Rack plating:** The parts to be plated are placed in a rack, which is then conveyed through a series of processing tanks.

2. **Barrel plating:** Small parts are placed inside a permeable barrel, which is then placed inside the processing tank(s). This operation is commonly performed on small parts, such as bolts, nuts, gears, and fittings. The electrolytic fluid can freely penetrate through the barrel and provide the metal for plating; electrical contact is provided through the barrel and through contact with other parts.

3. **Brush processing:** the electrolytic fluid is pumped through a handheld brush with metal bristles. The workpiece can be very large, and the process is suitable for field plating; it can also be used to apply coatings on large equipment without disassembling them.

Simple electroplating can be done in a single-process bath or tank, but more commonly, a sequence of operations is involved in a plating line. The rate of film deposition depends on the local current density, which is not necessarily uniform on a part. Workpieces with complex shapes may require a modified geometry because of varying plating thicknesses, as can be seen in Fig. 34.9. The following equipment and processes may be part of an electroplating operation:

- Chemical cleaning and degreasing are used to remove surface contaminants, enhancing surface adhesion of the plated coating.

- Parts may be exposed to a strong acid bath (*pickling solution*) to eliminate or reduce the thickness of the oxide coating on the workpiece.

- A base coating may be applied; this may involve the same or a different metal; if the desired metal coating will not adhere well to the substrate, an intermediate coating can be applied.

- A separate tank is used for final electroplating.

- Rinse tanks will be used throughout the sequence.

Common plating metals are chromium, nickel (for corrosion protection), cadmium, copper (corrosion resistance and electrical conductivity), and tin and zinc (corrosion protection, especially for sheet steel). **Chromium plating** involves first plating the metal with copper, then with nickel, and finally with chromium. **Hard chromium plating** is done directly on the base metal, and results in a surface hardness of up to 70 HRC (see Fig. 2.15) and a thickness of about 0.05 mm or higher. This method is used to improve the resistance to wear and corrosion of tools, valve stems, hydraulic shafts, and diesel- and aircraft-engine cylinder liners.

Examples of electroplating include copper-plating aluminum wire and phenolic boards for printed circuits, chrome-plating hardware, tin-plating copper electrical terminals (for ease of soldering), galvanizing sheet metal (see also Section 34.11), and plating components such as metalworking dies that require resistance to wear and galling (cold welding of small pieces from the workpiece surface). Metals such as

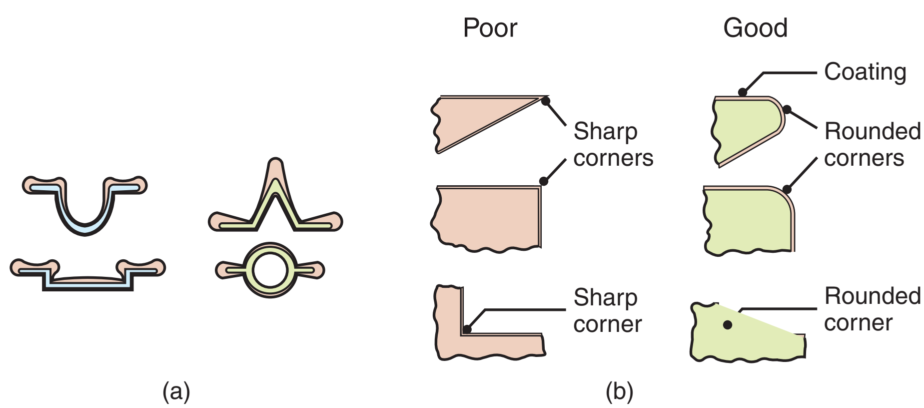


Figure 34.9: (a) Schematic illustration of nonuniform coatings (exaggerated) in electroplated parts. (b) Design guidelines for electroplating. Note that sharp external and internal corners should be avoided for uniform plating thickness.

gold, silver, and platinum are important electroplating materials in the electronics and jewelry industries for electrical contact and for decorative purposes, respectively.

Plastics, such as ABS, polypropylene, polysulfone, polycarbonate, polyester, and nylon, also can be electroplated. Because they are not electrically conductive, plastics must first be preplated, by a process such as electroless nickel plating. Parts to be coated may be simple or complex, and size is not a limitation.

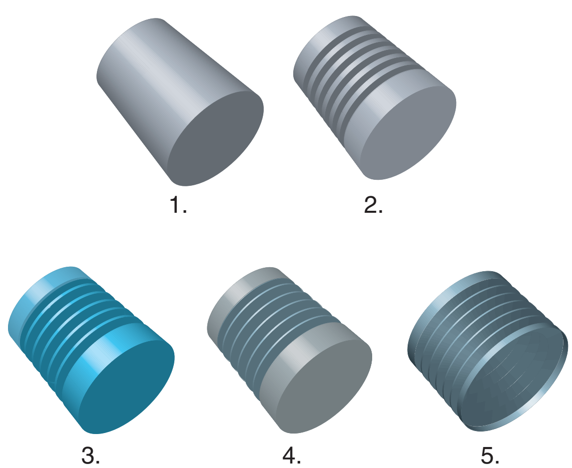
Electroless Plating. This process is carried out by a chemical reaction, without using an external source of electricity. The most common application utilizes nickel as the plating material, although copper also is used. In *electroless nickel plating*, nickel chloride (a metallic salt) is reduced, with sodium hypophosphite as the reducing agent, to nickel metal, which is then deposited on the part. The hardness of nickel plating ranges between 425 and 575 HV; the plating can subsequently be heat treated to 1000 HV. The coating has excellent wear and corrosion resistance.

Cavities, recesses, and the inner surfaces of tubes can be plated successfully. Electroless plating also can be used with nonconductive materials, such as plastics and ceramics. The process is more expensive than electroplating, but unlike electroplating, the coating thickness of electroless plating is always uniform.

Electroforming. A variation of electroplating, electroforming is a metal-fabricating process. Metal is electrodeposited on a *mandrel* (also called a *mold* or a *matrix*), which is then removed. The coating itself thus becomes the product (Fig. 34.10). Both simple and complex shapes can be made by electroforming, with wall thicknesses as small as 0.025 mm. Parts may weigh from a few grams to as much as 270 kg.

Mandrels are made from a variety of materials: including (a) metals, such as zinc or aluminum; (b) nonmetals, which can be made electrically conductive with the appropriate coatings; and (c) low-melting alloys, wax, or plastics, all of which can be melted away or dissolved with suitable chemicals. Mandrels should be physically removable from the electroformed part without damaging it.

This process is particularly suitable for low production quantities or intricate parts, such as molds, dies, waveguides, nozzles, and bellows, made of nickel, copper, gold, and silver. The process is also suitable for aerospace, electronics, and electro-optics applications.



(a)

(b)

Figure 34.10: (a) Typical sequence in electroforming. (1) A mandrel is selected with the correct nominal size. (2) The desired geometry (in this case, that of a bellows) is machined into the mandrel. (3) The desired metal is electroplated onto the mandrel. (4) The plated material is trimmed if necessary. (5) The mandrel is dissolved through chemical machining. (b) A collection of electroformed parts. *Source:* Courtesy of Servometer®, Cedar Grove, NJ.

34.10 Conversion Coatings

Conversion coating, also called *chemical-reaction priming*, is the process of producing a coating that forms on metal surfaces as a result of chemical or electrochemical reactions. Oxides that naturally form on their surfaces (see Section 33.2) are a form of conversion coating. Various metals, particularly steel, aluminum, and zinc, can be conversion coated.

Phosphates, chromates, and *oxalates* are used to produce conversion coatings, for such purposes as providing corrosion protection, prepainting, and decorative finishing. An important application is the conversion coating of workpieces to serve as lubricant carriers in cold-forming operations, particularly zinc-phosphate and oxalate coatings (see Section 33.7.6). Two common methods of coating are *immersion* and *spraying*.

Anodizing. This is an oxidation process (*anodic oxidation*), in which the part surfaces are converted to a hard and porous oxide layer which provides corrosion resistance and a decorative finish. The part is the anode in an electrolytic cell immersed in an acid bath, which results in rapid oxidation of the workpiece. Organic dyes of various colors, usually black, red, bronze, gold, or gray, can be used to produce stable and durable surface films. Typical applications include aluminum furniture and utensils, picture frames, keys, sporting goods, and architectural shapes. Anodized surfaces also serve as a good base for painting, especially on aluminum, which otherwise is difficult to paint.

Coloring. As the name implies, coloring involves processes that alter the color of metals, alloys, and ceramics. This change is caused by the conversion of surfaces, by chemical, electrochemical, or thermal processes, into such chemical compounds as oxides, chromates, and phosphates. A common example is *blackening* of iron and steels, a process that utilizes solutions of hot, caustic soda, resulting in chemical reactions that produce a lustrous, black oxide film on surfaces.

34.11 Hot Dipping

In *hot dipping*, the part, usually steel or iron, is dipped into a bath of molten metal, such as (a) zinc, for galvanized-steel sheet and plumbing supplies; (b) tin, for tinplate and tin cans for food containers; (c) aluminum (aluminizing); and (d) *terne*, an alloy of lead with 10 to 20% tin. Hot-dipped coatings on discrete parts provide long-term corrosion resistance to galvanized pipes, plumbing supplies, and other similar products.

A typical continuous *hot-dipped galvanizing line* for sheet steel is shown in Fig. 34.11. The rolled sheet is first cleaned electrolytically, then scrubbed by brushing. The sheet is then annealed in a continuous furnace with controlled atmosphere and temperature, and dipped in molten zinc at about 450°C. The thickness of the zinc coating is controlled by a wiping action from a stream of air or steam, called an *air knife*, also used in wave soldering (see Fig. 32.7b).

34.12 Porcelain Enameling; Ceramic and Organic Coatings

Metals can be coated with a variety *vitreous* (glassy) coatings, to provide corrosion and electrical resistance, and for protection at elevated temperatures. These coatings usually are classified as **porcelain enamels**, and generally include enamels and ceramics. The root of the word porcelain is *porcellana*, in Italian meaning marine shell. Note that the word *enamel* also is used as a term for *glossy paints*, indicating a smooth, hard coating.

Enamels. Porcelain enamels are glassy inorganic coatings that consist of various metal oxides and are available in various colors and transparencies. *Enameling*, which was a fully developed art by the Middle Ages, involves fusing the coating material to the substrate at temperatures of 425° to 1000°C to

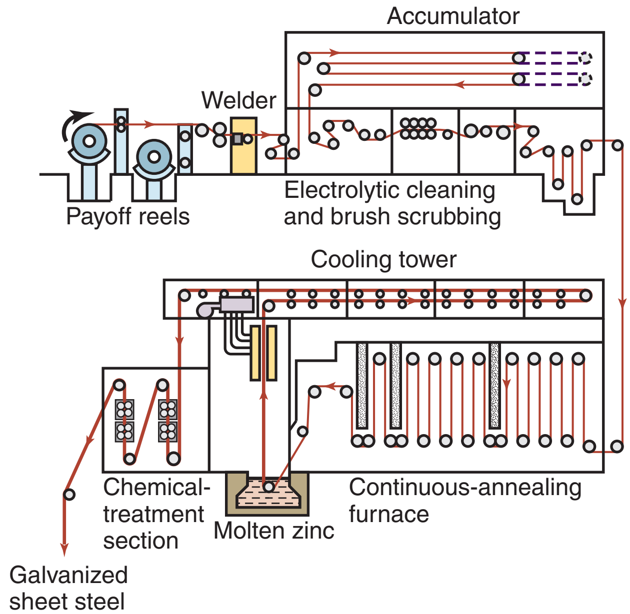


Figure 34.11: Flow line for the continuous hot-dipped galvanizing of sheet steel. The welder (upper left) is used to weld the ends of coils to maintain continuous material flow. *Source:* Courtesy of the American Iron and Steel Institute.

liquefy the oxides. The coating may be applied by dipping, spraying, or electrodeposition, and thicknesses are usually in the range of 0.05–0.6 mm. The viscosity of the material can be controlled using binders so that the coating adheres also to vertical surfaces during application. Depending on their composition, enamels have varying resistances to alkali, acids, detergents, cleansers, and water.

Typical applications for porcelain enameling are household appliances, plumbing fixtures, chemical-processing equipment, signs, cookware, and jewelry; they are also used as protective coatings on jet-engine components. Metals that are coated are typically steels, cast iron, and aluminum. For chemical resistance, glasses are used as a lining material, where its thickness is much greater than that of enamel. **Glazing** is the application of glassy coatings onto ceramic wares, to give them decorative finishes and to make them impervious to moisture.

Ceramic Coatings. Ceramics, such as aluminum oxide and zirconium oxide (Sections 22.6 to 22.8), are applied to a substrate at room temperature by means of binders and then fired in a furnace to fuse the coating material. Usually applied using thermal spraying techniques, the coatings act as thermal barriers, for turbine blades, diesel-engine components, hot-extrusion dies, and nozzles for rocket motors. The coatings extend the life of these components and also are used for electrical-resistance applications to withstand repeated arcing.

Organic Coatings. Metal surfaces can be coated or precoated with a variety of organic coatings, films, and laminates to improve appearance and corrosion resistance. Coatings are applied to coil stock on continuous lines (see Fig. 13.11), with thicknesses generally in the range of 0.0025–0.2 mm. Organic coatings have a wide range of characteristics , such as flexibility, durability, hardness, resistance to abrasion and chemicals, color, texture, and gloss. Coated sheet metals are subsequently shaped into various products, such as TV cabinets, appliance housings, paneling, shelving, residential-building siding, gutters, and metal furniture.

Example 34.2 Ceramic Coatings for High-temperature Applications

Table 34.1 shows various ceramic coatings and their typical applications at elevated temperatures. These coatings may be applied either singly or in layers, each layer with its own special properties, as is done in multiple-layer coated cutting tools (Fig. 22.8).

Table 34.1: Ceramic Coatings Used for High-temperature Applications.

Property	Type of ceramic	Applications
Wear resistance	Chromium oxide, aluminum oxide, aluminum titania	Pumps, turbine shafts, seals, and compressor rods for the petroleum industry; plastics extruder barrels; extrusion dies
Thermal insulation	Zirconium oxide (yttria stabilized), zirconium oxide (calcia stabilized), magnesium zirconate	Fan blades, compressor blades, and seals for gas turbines; valves, pistons, and combustion heads for automotive engines
Electrical insulation	Magnesium aluminate, aluminum oxide	Induction coils, brazing fixtures, general electrical applications

Critical applications of organic coatings involve, for example, the protection of naval aircraft, as they are constantly subjected to rain, seawater, pollutants, high humidity, aviation fuel, and deicing fluids, as well as being impacted by particles such as dust, gravel, and stones. For aluminum structures, organic coatings consist typically of an epoxy primer and a polyurethane topcoat.

34.13 Diamond Coating and Diamondlike Carbon

The properties of *diamond* that are relevant to manufacturing engineering are described in Section 8.7. Important advances continue to be been made in *diamond coating* over metals, glass, ceramics, and plastics. The techniques employed are chemical vapor deposition, plasma-assisted vapor deposition, and ion-beam-enhanced deposition.

Examples of diamond-coated products are scratchproof windows, such as those used in aircraft and military vehicles for protection in sandstorms; turbine blades; fuel-injection nozzles;.cutting tools, such as inserts, drills, and end mills; wear faces of micrometers and calipers; surgical knives; razors; electronic and infrared heat seekers and sensors; light-emitting diodes; and speakers for stereo systems.

Techniques also have been developed to produce **freestanding diamond films**, on the order of 1 mm thick and up to 125 mm in diameter. These films include smooth, optically clear diamond film, which is then laser cut to desired shapes and brazed onto workpieces.

Growth of diamond films on crystalline-copper substrate are being done by implantation of carbon ions. An important application is in making computer chips (Chapter 28). Diamond can be doped to form p- and n-type ends on semiconductors for transistors. Its high thermal conductivity allows closer packing of chips than would be possible with silicon or gallium-arsenide chips, significantly increasing the speed of computers. Diamond is also an important material for MEMS devices (Chapter 29), because of its favorable friction and wear characteristics.

Diamondlike Carbon. *Diamondlike carbon* (DLC) coatings, a few nanometers in thickness, are produced by a low-temperature, ion-beam-assisted deposition process. The structure of DLC is between that of diamond and graphite (Section 8.6). Less expensive than diamond films but with similar properties, it has low friction, high hardness, and chemical inertness, as well as having a smooth surface. DLC has applications in such areas as tools and dies, engine components, gears, bearings, MEMS devices, and microscale probes. As a coating on cutting tools, it has a hardness of about 5000 HV, as compared with about twice that for diamond.

34.14 Surface Texturing

Manufactured surfaces can be modified further by secondary operations for functional, optical, or aesthetic reasons. Called *surface texturing*, the secondary operations generally consist of the following techniques:

- **Etching**, using chemicals or sputtering techniques
- **Electric arcs**
- **Lasers**, using pulsed beams
- **Atomic oxygen**, reacting with surfaces to produce a fine, cone-like surface texture

34.15 Painting

Paints have been widely used for thousands of years as a surface coating and for decoration. Paints are generally classified as

- **Enamels,** producing a smooth coat with a glossy or semiglossy appearance
- **Lacquers,** forming an adherent film by evaporation of a solvent
- **Water-based paints,** applied easily, but have a porous surface and absorb water, making them more difficult to clean.

Paints are available with good resistance to abrasion, high temperatures, and fading. Their selection depends on specific requirements, such as resistance to abrasion, marring, impact, flexing, acids, solvents, detergents, alkali, fuels, staining, and general environmental attack.

Common methods of applying paint are dipping, brushing, rolling, and spraying (Fig. 34.12). In **electrocoating** or **electrostatic spraying**, paint particles are charged *electrostatically* and are attracted to surfaces, producing a uniformly adherent coating. Unlike paint losses in conventional spraying, which may be as much as 70% of the paint, the loss in electrostatic spraying can be as little as 10%. However, deep recesses and corners can be difficult to coat by this method. Using *robotic controls* for guiding the spray nozzles is now a common practice (Section 37.6.3).

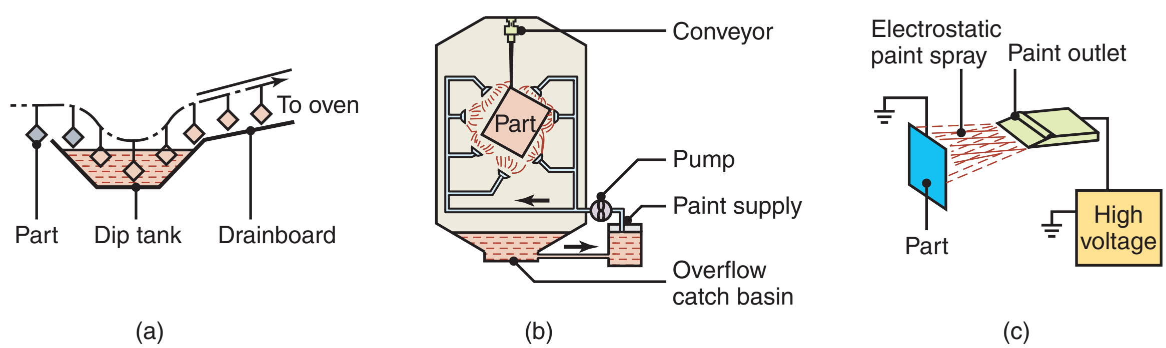


Figure 34.12: Methods of paint application: (a) dip coating, (b) flow coating, and (c) electrostatic spraying (used particularly for automotive bodies).

34.16 Cleaning of Surfaces

The word **clean** or *degree of cleanliness* of a surface is somewhat difficult to define; two common and simple tests are:

1. Observing whether water continuously and uniformly coats a surface, known as the *waterbreak* test. If water collects as individual droplets, the surface is not clean, a phenomenon that can easily be demonstrated by wetting dinner plates that have been washed to different degrees of cleanliness.
2. Wiping the surface with a clean white cloth and observing any residues on a clean white cloth.

A clean surface can have both beneficial and detrimental effects. Although a surface that is not clean may reduce the tendency for adhesion in sliding, thus reducing friction, cleanliness is generally essential for more effective application of coatings, paints, adhesive bonding, brazing, soldering, moving parts in machinery, and assembly operations. Also, aluminum cans must have clean outer surfaces, as otherwise labels cannot be printed on them.

In manufacturing operations, the type of *cleaning* process required depends on the type of *metalworking-fluid residues* and *contaminants* to be removed. Water-based fluids, for example, are easier and less expensive to remove than oil-based fluids. Contaminants, also called *soils*, may consist of rust, scale, chips, various metallic and nonmetallic debris, metalworking fluids, solid lubricants, pigments, polishing and lapping compounds, and general environmental elements.

Basically, there are three types of cleaning methods:

Mechanical Cleaning. This operation consists of physically removing the contaminants, often with wire or fiber brushing, abrasive blasting, tumbling, or with steam jets. Many of these operations, including *ultrasonic cleaning*, are particularly effective in removing rust, scale, and other solid contaminants from surfaces.

Electrolytic Cleaning. In this process, a charge is applied to the part to be cleaned in an aqueous and often alkaline cleaning solution. The charge develops bubbles of hydrogen or oxygen, depending on polarity, being released at the surface. The bubbles are abrasive and help remove contaminants.

Chemical Cleaning. This process usually involves removal of oil and grease from surfaces, and consists of one or more of the following:

- **Solution:** The soil is dissolved in the cleaning solution.
- **Saponification:** A chemical reaction converts the animal or vegetable oils into a soap which is soluble in water.
- **Emulsification:** The cleaning solution reacts with the soil or lubricant residues and forms an emulsion; the soil and the emulsifier then become suspended in the emulsion.
- **Dispersion:** The concentration of soil on the surface is decreased by the action of surface-active elements in the cleaning solution.
- **Aggregation:** Lubricant residues are removed from a surface by the agents in the cleanser, and are then collected as large dirt particles.

Cleaning Fluids. Common cleaning fluids used in plants in conjunction with electrochemical processes for more effective cleaning include:

- **Alkaline solutions:** A complex combination of water-soluble chemicals, alkaline solutions are the least expensive and most widely used cleaning fluids in manufacturing. Small parts may be cleaned in rotating drums or barrels. Most parts are cleaned on continuous conveyors, by spraying them with the solution and rinsing them with water.

- **Emulsions:** Emulsions used generally consist of kerosene and oil-in-water and various types of emulsifiers.
- **Solvents:** Petroleum solvents, chlorinated hydrocarbons, and mineral spirits are generally used, especially for short runs; fire and toxicity are major hazards.
- **Hot vapors:** Chlorinated solvents can be used to remove oil, grease, and wax by this process, also known as **vapor degreasing**; it is simple and the cleaned parts are dry.
- **Acids, salts, and mixtures of organic compounds**: Effective in cleaning parts covered with heavy paste or oily deposits and rust.

Design Guidelines for Cleaning. Cleaning discrete parts with complex shapes can be difficult. Some basic design guidelines include (a) avoiding deep, blind holes, (b) making several smaller components instead of one large component, and (c) providing appropriate drain holes in the parts to be cleaned.

The *treatment* and *disposal* of cleaning fluids, as well as of various fluids and waste materials from the processes described in this chapter, are among the most important considerations for environmentally safe manufacturing operations (see also Section I.4).

Summary

- Surface treatments are an important aspect of all manufacturing operations. They are used to impart specific mechanical, chemical, and physical properties, such as appearance, and resistance to corrosion, friction, wear, and fatigue.
- Processes used include mechanical working and such surface treatments as heat treatment, deposition, and plating. Surface coatings include enamels, nonmetallic materials, and paints.
- Clean surfaces often are important for further processing of parts, such as coating, painting, and welding, and in the use of the products. Cleaning costs can have a significant economic impact on manufacturing operations.

Key Terms

Anodizing
Ballizing
Blackening
Case hardening
Chemical cleaning
Chemical vapor deposition
Cladding
Cleaning fluids
Coloring
Conversion coating
Diamond coating
Diamondlike carbon
Diffusion coating
Electroforming
Electroless plating
Electroplating
Enamel
Explosive hardening
Freestanding diamond film
Glazing
Hard-chromium plating
Hard facing
Hot dipping
Ion implantation
Ion plating
Laser peening

Mechanical plating
Metallizing
Painting
Physical vapor deposition
Porcelain enamel
Roller burnishing
Shot peening
Spraying
Sputtering
Surface texturing
Thermal spraying
Vapor deposition
Waterbreak test
Water-jet peening

Bibliography

ASM Handbook, Vol. 5, **Surface Engineering**, ASM International, 1994.
ASM Handbook, Vol. 5A, **Thermal Spray Technology**, ASM International, 2013.
ASM Handbook, Vol. 5B, **Protective Organic Coatings**, ASM International, 2015.
Boronenkov, V., and Korobov, Y., **Fundamentals of Arc Spraying**, Springer, 2016.
Bunshah, R.F. (ed.), **Handbook of Hard Coatings: Deposition Technologies, Properties and Applications**, Elsevier, 2014.
Cavaliere, P., **Cold-Spray Coatings**, Springer, 2017.
Davis, J.R. (ed.), **Surface Engineering for Corrosion and Wear Resistance**, ASM International and IOM Communications, 2001.
Djokic, S.S., **Electrodeposition and Surface Finishing: Fundamentals and Applications**, Springer, 2014.
Fauchais, P.L., Heberlein, J.V.R., and Boulos, M.I., **Thermal Spray Fundamentals**, Springer, 2014,
Kay, C.M., and Karthikeyan, J., **High Pressure Cold Spray: Principles and Applications**, ASM International, 2016.
Momber, **Blast Cleaning Technology**, Springer, 2008.
Pawlowski, L., **The Science and Engineering of Thermal Spray Coatings**, Wiley, 2008.
Pierson, H.O., **Handbook of Chemical Vapor Deposition**, 2nd ed., Noyes Publications, 2000.
Polak, T.A. (ed.), **Handbook of Surface Treatments and Coatings**, ASME Press, 2003.
Roberge, P.R., **Handbook of Corrosion Engineering**, 2nd ed., McGraw-Hill, 2012.
Ruzbarsky, J., and Panda, A., **Plasma and Thermal Spraying**. Springer, 2016.
Schulze, V., **Modern Mechanical Surface Treatment**, Wiley-VCH, 2006.
Tracton, A.A. (ed.), **Coatings Technology Handbook**, 3rd ed., CRC Press, 2005.
Villafuerte, J., **Modern Cold Spray: Materials, Process, and Applications**, Springer, 2015.

Review Questions

34.1. Explain why surface treatments may be necessary for various parts made by one or more processes.

34.2. What is shot peening? Why is it performed?

34.3. What are the advantages of roller burnishing?

34.4. Explain the difference between case hardening and hard facing.

34.5. Describe the principles of physical and chemical vapor deposition. What applications do these processes have?

34.6. Review Fig. 34.2 and describe the principles behind thermal spray.

34.7. What is electroplating? Why is it dangerous?

34.8. What is the principle of electroforming? What are the advantages of electroforming?

34.9. Explain the difference between electroplating and electroless plating.

34.10. How is hot dipping performed?

34.11. What is an air knife? How does it function?

34.12. Describe the common painting systems presently in use in industry.

34.13. What is a conversion coating? Why is it so called?

34.14. Describe the difference between thermal spraying and plasma spraying.

34.15. What is cladding, and why is it performed?

34.16. How are diamond coatings produced?

Qualitative Problems

34.17. Describe how roller-burnishing processes induce compressive residual stresses on the surfaces of parts.

34.18. Explain why some parts may be coated with ceramics. Give some examples.

34.19. List and briefly describe five surface treatment techniques that use lasers.

34.20. Give examples of part designs that are suitable for hot-dip galvanizing.

34.21. Comment on your observations regarding Fig. 34.9.

34.22. It is well known that coatings may be removed or depleted during the service life of components, particularly at elevated temperatures. Describe the factors involved in the strength and durability of coatings.

34.23. Make a list of the coating processes described in this chapter and classify them in relative terms as "thick" or "thin."

34.24. Sort the coating processes described in this chapter according to (a) maximum thickness generally achieved; (b) typical coating time.

34.25. Why is galvanizing important for automotive-body sheet metals?

34.26. Explain the principles involved in various techniques for applying paints.

34.27. List several applications for coated sheet metal, including galvanized steel.

Quantitative Problems

34.28. Taking a simple example, such as the process shown in Fig. 34.1, estimate the force required for roller burnishing. (*Hint:* See Sections 2.6 and 14.4.)

34.29. Estimate the plating thickness in electroplating a 30-mm solid-metal ball using a current of 12 A and a plating time of 1.6 hours. Assume that $c = 0.08$ in Eq. (34.1).

Synthesis, Design, and Projects

34.30. Which surface treatments are functional, and which are decorative? Are there any treatments that serve both functions? Explain.

34.31. Explain the role of conversion coatings. Based on Fig. 33.13, what lubrication regime is most suitable for application of conversion coatings?

34.32. An artificial implant has a porous surface area where it is expected that the bone will attach and grow into the implant. Without consulting the literature, make recommendations for producing a porous surface; then review the literature and describe the actual processes used.

34.33. If one is interested in obtaining a textured surface on a coated piece of metal, should one apply the coating first or apply the texture first? Explain.

34.34. It is known that a mirrorlike surface finish can be obtained by plating workpieces that are ground; that is, the surface finish improves after coating. Explain how this occurs.

34.35. It has been observed in practice that a thin layer of chrome plating, such as that on older model automobile bumpers, is better than a thick layer. Explain why, considering the effect of thickness on the tendency for cracking.

34.36. Outline the reasons that the topics described in this chapter are important in manufacturing processes and operations.

34.37. Shiny, metallic balloons have festive printed patterns that are produced by printing screens and then plated onto the balloons. How can metallic coatings be plated onto a rubber sheet?

34.38. Because they evaporate, solvents and similar cleaning solutions have adverse environmental effects. Describe your thoughts on what modifications could be made to render cleaning solutions more environmentally friendly.

34.39. A roller-burnishing operation is performed on a shaft shoulder to increase fatigue life. It is noted that the resultant surface finish is poor, and a proposal is made to machine the surface layer to further improve fatigue life. Will this be advisable? Explain.

34.40. The shot-peening process can be demonstrated with a ball-peen hammer (in which one of the heads is round). Using such a hammer, make numerous indentations on the surface of a piece of aluminum sheet (a) 2 mm and (b) 10 mm thick, respectively, placed on a hard flat surface such as an anvil. Note that both pieces develop curvatures, but one becomes concave and the other convex. Describe your observations and explain why this happens. (*Hint:* See Fig. 2.14.)

34.41. Obtain several pieces of small metal parts (such as bolts, rods, and sheet metal) and perform the waterbreak test on them. Then clean the surfaces with various cleaning fluids and repeat the test. Describe your observations.

34.42. Inspect various products, such as small and large appliances, silverware, metal vases and boxes, kitchen utensils, and hand tools, and comment on the type of coatings they may have and the reasons they are coated.

PART VIII
Engineering Metrology, Instrumentation, and Quality Assurance

The vast majority of manufactured parts are components of a product's subassemblies, and they must fit and be assembled properly so that the product performs its intended function during its service life. Dimensions and other surface features of a part are measured to ensure that it is manufactured consistently and within the specified range of dimensional tolerances. Note, for example, that (a) a piston must fit into a cylinder within specified tolerances, (b) a computer or cellphone housing must be assembled without overhang, and with good mechanical integrity, and (c) the slideways of a machine tool must be made with a certain accuracy so that the parts produced on that machine are accurate within their specifications.

Measurement of the dimensions and features of parts is an integral aspect of **interchangeable parts manufacturing**, the basic concept behind standardization and mass production. For example, if a ball bearing in an appliance is worn and must be be replaced years after its purchase, all one has to do is purchase a replacement with the same specification or part number.

The first of the next two chapters describes the principles involved in and various instruments and modern machines used for measuring dimensional features, such as length, angle, flatness, and roundness. Testing and inspecting parts are equally important aspects of manufacturing operations; thus, the methods for the nondestructive and destructive testing of parts also are described. **Product quality** (Chapter 36) is one of the most important aspects of manufacturing. It is essential to recognize the technological and economic importance of *building quality into a product*, rather than inspecting it *after* it is made.

Chapter 35

Surface Treatments, Coatings, and Cleaning

- This chapter describes the importance of measurement, noting that measurement of parts and their certification to a certain standard is essential to ensuring that parts fit and operate properly in products.
- A wide variety of measurement techniques, gages, instruments, and machines is now available, as described and illustrated in this chapter.
- The chapter also describes features of automated measurement, concluding with an introduction to the principles of dimensioning and tolerancing.

35.1 Introduction

Engineering metrology is defined as the measurement of length, thickness, diameter, taper, angle, flatness, and profile. Note that these are *geometric* measures; mechanical and physical property measurements are not considered in metrology. Consider, for example, the slideways for machine tools (Fig. 35.1; see also Figs. 23.2 and 24.18). These components must have specific dimensions, angles, flatness, and tolerances in order for the machine tool to function properly,

Traditionally, measurements have been made *after* the part has been made, an approach known as **postprocess inspection**. Here, the term *inspection* means checking the dimensions of what has been just produced, or is being produced, and determining whether those dimensions comply with the specified dimensional tolerances and other specifications. Today, measurements are being made *while* the part is produced on the machine, an approach known as **in-process, in-situ, online, or real-time inspection**.

An important aspect of metrology is **dimensional tolerance** (after the Latin *tolerare* meaning to tolerate). Tolerances are important because of their major role in part interchangeability, proper functioning of the product, and manufacturing costs. Generally, the smaller the tolerance, the higher are the production costs, because of the additional operations required to reach it.

35.2 Measurement Standards

The earliest experience with measurement is usually with a simple *ruler*, also called a *rule*, to measure **length**. Rulers are used as a *standard* against which dimensions are measured. Traditionally, in most English-speaking countries, the units *inch* and *foot* have been and continue to be used.

In most of the world, however, the *meter* is used as a length standard. Originally, 1 m was defined as one ten-millionth of the distance between the North Pole and the equator. Later, the original meter length was standardized as the distance between two scratches on a platinum-iridium bar, kept under controlled conditions in a building outside Paris. In 1960, the meter was officially defined as 1,650,763.73 wavelengths, in a vacuum, of the orange light given off by electrically excited krypton 86, a rare gas. The precision of this measurement was set as 1 part in 10^9. The meter is now an international standard of length in the Système International d'Unités (SI).

Numerous measuring instruments and devices are used in engineering metrology, each of which has its own resolution, precision, and other features. Two terms commonly used to describe the type and quality of an instrument are:

1. **Resolution:** The smallest difference in dimensions that the measuring instrument can detect or distinguish. A wooden yardstick, for example, has far less resolution than a micrometer.

2. **Precision:** Sometimes incorrectly called accuracy, precision is the degree to which an instrument gives repeated measurements of the same standard. For example, an aluminum ruler will expand or contract, depending on temperature variations in the environment in which it is used; thus, its precision can be affected even while being held by hand.

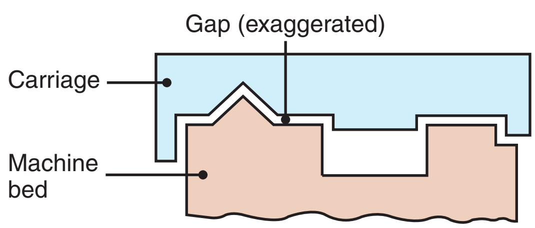


Figure 35.1: Cross-section of a machine-tool slideway. The width, depth, angles, and other dimensions all must be produced and measured accurately for the machine tool to function as designed.

35.3 Geometric Features of Parts: Analog and Digital Measurements

The most commonly used quantities and geometric features in engineering practice are listed and described in this section.

- **Length,** includes all linear dimensions of parts
- **Diameter,** includes outside and inside diameters
- **Roundness,** includes out-of-roundness, concentricity, and eccentricity
- **Depth,** such as that of drilled holes and cavities in dies and molds
- **Straightness,** shafts, bars, and tubing
- **Flatness,** machined, ground, and polished surfaces
- **Parallelism,** two shafts or slideways in a machine
- **Perpendicularity,** a threaded rod inserted into a flat plate
- **Angle,** internal and external angles
- **Profile,** curvatures of parts made by various processes.

A wide variety of instruments and machines is available to accurately and rapidly measure these quantities, either on stationary parts or on parts that are in continuous production. In engineering metrology, the words **instrument** and **gage** often are used interchangeably. Because of major global trends in automation and computer control of manufacturing operations (Part IX), modern measuring equipment and instrumentation have become an *integral part of production machinery*. Implementation of **digital** instrumentation and developments in computer-integrated manufacturing have, together, led to the total integration of measurement technologies within manufacturing systems.

Temperature control is very important, particularly for making measurements with precision instruments. The standard measuring temperature is 20°C, and all gages are calibrated at this temperature. Measurements must be taken in controlled environments maintaining the standard temperature, usually within ±0.3°C, although this can be lower for extremely precise measurements.

An **analog** instrument, such as a vernier caliper or micrometer (Fig. 35.2a), relies on the skill of the operator to properly interpolate and read the graduated scales. In contrast, measurements in a digital micrometer are indicated directly (Fig. 35.2b). More importantly, digital equipment can easily be integrated into other equipment (Fig. 35.2c), including production machinery and systems for statistical process control (SPC), as described in Chapter 36.

35.4 Traditional Measuring Methods and Instruments

35.4.1 Line-graduated Instruments

Graduated means marked to indicate a certain quantity. These instruments are used for measuring lengths or angles.

Linear Measurement (Direct Reading).

- **Rules:** The simplest and most commonly used instrument for making linear measurements is a *steel rule* (*machinist's rule*), bar, or tape, with fractional or decimal graduations. Lengths are measured directly, to an accuracy that is limited to the nearest division, usually 1 mm.

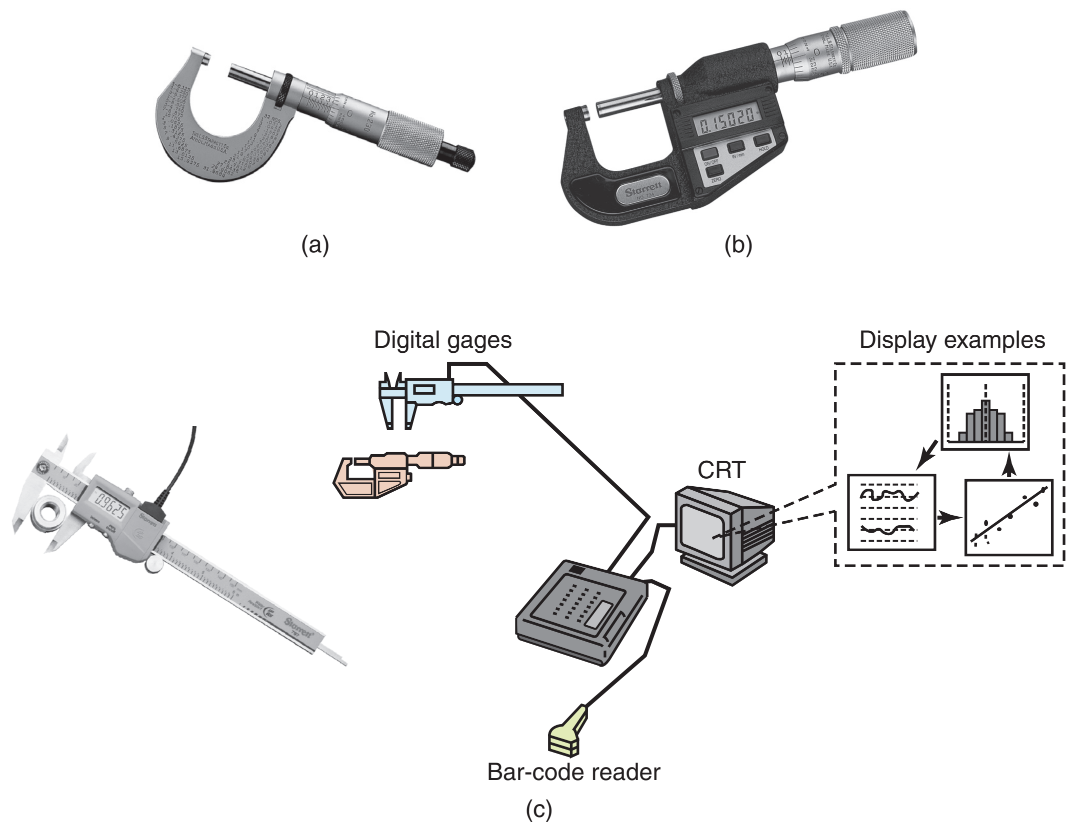


Figure 35.2: (a) A vernier (analog) micrometer. (b) A digital micrometer with a range of 0 to 25 mm and a resolution of 1.25 μm; note that it is much easier to read dimensions on this instrument than on analog micrometers. (c) Schematic illustration showing the integration of digital gages with microprocessors for real-time data acquisition for statistical process control. *Source:* (a) Courtesy of L.C. Starrett Co. (b) Courtesy of Mitutoyo America Corp.

- **Calipers:** Also called *caliper gages* and *vernier calipers* (named for P. Vernier, in the 1600s), they have a graduated beam and a sliding *jaw*. *Digital calipers* are in wide use.
- **Micrometers:** These instruments are commonly used for measuring the thickness and inside or outside dimensions of parts. *Digital micrometers* are equipped with readouts (Fig. 35.2b). They are also available for measuring internal diameters (*inside micrometer*) and depths (*micrometer depth* gage, Fig. 35.3). The anvils on micrometers can be equipped with conical or ball contacts, to measure recesses on parts, threaded-rod diameters, and wall thicknesses of tubes and curved sheets or plates.

Linear Measurement (Indirect Reading). These instruments typically are calipers and *dividers*, without any graduated scales. They are used to transfer the measured size to a direct-reading instrument, such as a ruler. The accuracy of indirect-measurement tools is limited. *Telescoping gages* can be used for indirect measurement of holes or cavities.

Angle Measurement

- **Bevel protractor:** This is a direct-reading instrument, similar to a common protractor, except that it has a movable element. The two blades of the protractor are placed in contact with the part being measured; the angle is then read directly on the vernier scale. A bevel protractor is a *combination square*, a steel rule equipped with devices for measuring 45° and 90° angles.

Figure 35.3: A digital micrometer depth gage. *Source:* Courtesy of Starrett Co.

- **Sine bar:** This method consists of a plate mounted on two solid cylinders (Fig. 35.4). One cylinder is a pivot for the plate, and the other is located a fixed distance away, commonly 254 mm, from the first bar. The angle of the plate can be varied by inserting gage blocks (Section 35.4.4) under the second bar; the relationship is given by

$$\sin\theta = \frac{h}{l} \tag{35.1}$$

- **Surface plates:** These plates are used to support both the part to be measured and the measuring instrument. The plates are made to high flatness, with variations as little as 0.25 μm across the plate. They typically are made of cast iron or natural stones, such as granite. Granite plates have the desirable properties of being resistant to corrosion, are nonmagnetic, and have low thermal expansion, thereby minimizing thermal distortion.

Comparative Length Measurement. Instruments used for measuring comparative lengths, also called *deviation-type* instruments, amplify and measure variations in the distance between two or more surfaces. These instruments, the most common example being a **dial indicator** (Fig. 35.5), compare dimensions, hence the term *comparative*. The indicator is first set to zero at a certain reference surface, then the instrument or the surface to be measured, either external or internal, is brought into contact with the pointer. The movement of the indicator is read directly on the circular dial, to accuracies as high as 1 μm. Dial indicators with electrical and fluidic amplification mechanisms and with a digital readout also are available.

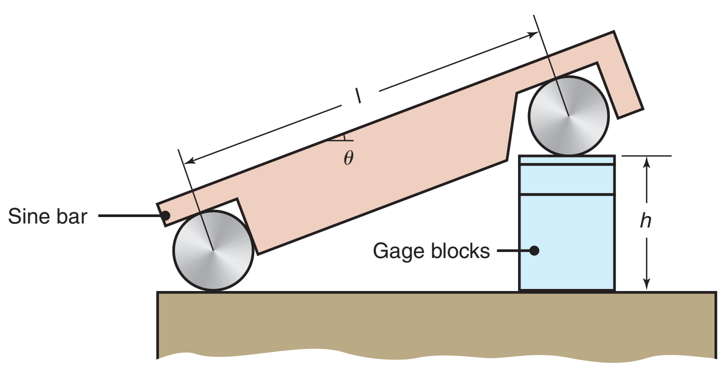


Figure 35.4: Schematic illustration of a sine bar.

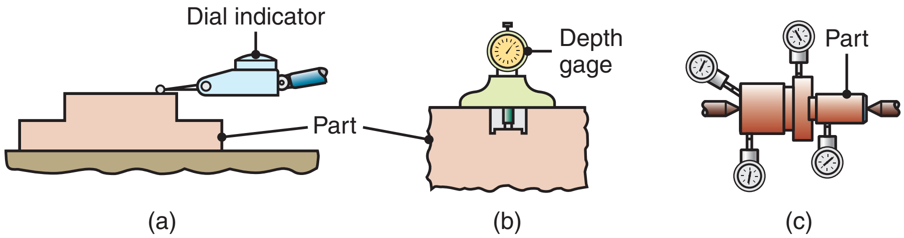


Figure 35.5: Three uses of dial indicators to measure (a) roundness and (b) depth, and (c) for multiple-dimension gaging of a part.

35.4.2 Measuring Geometric Features

Straightness. *Straightness* can be checked with a straightedge or a dial indicator (Fig. 35.6). An *autocollimator*, which resembles a telescope with a light beam that bounces back from the object, is used to accurately measure small angular deviations on a flat surface. *Laser beams* are used to align individual machine elements in the assembly of machine components.

Flatness. *Flatness* can be measured by mechanical means, using a *surface plate* and a *dial indicator*. This method can be used to measure *perpendicularity*, which also can be measured by precision-steel squares. Flatness can also be measured by **interferometry**, using an *optical flat*. The device consists of a glass or fused-quartz disk, with parallel flat surfaces, placed on the workpiece surface (Fig. 35.7a). When a *monochromatic* (one wavelength) light beam is aimed at the surface at an angle, the optical flat splits the light beam into two beams, appearing as light and dark fringes or bands, as shown in Fig. 35.7b.

The number of fringes that appear is related to the distance between the surface of the part and the bottom surface of the optical flat (Fig. 35.7c). A perfectly flat surface (i.e., one in which the angle between the two surfaces is zero) will not split the light beam, thus fringes will not appear. However, when surfaces are not flat, the fringes are curved (Fig. 35.7d). The interferometry method is also used for observing textures and scratches on surfaces (Fig. 35.7e).

Diffraction gratings consist of two optical flat glasses with different lengths and closely spaced parallel lines that are scribed on their surfaces. The grating on the shorter glass is inclined slightly, whereby *interference fringes* develop when viewed over the longer glass. The position of these fringes depends on the relative position of the two sets of glasses. With modern equipment and using electronic counters and photoelectric sensors, a resolution of 2.5 μm can be obtained, with gratings having 40 lines/mm.

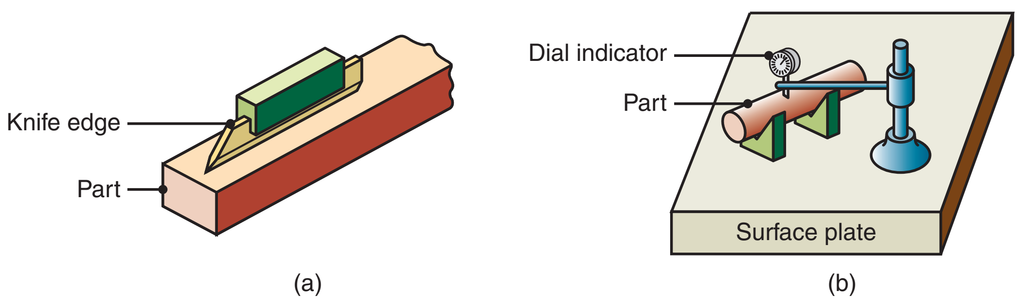


Figure 35.6: Measuring straightness manually with (a) a knife-edge rule and (b) a dial indicator. *Source:* After F.T. Farago.

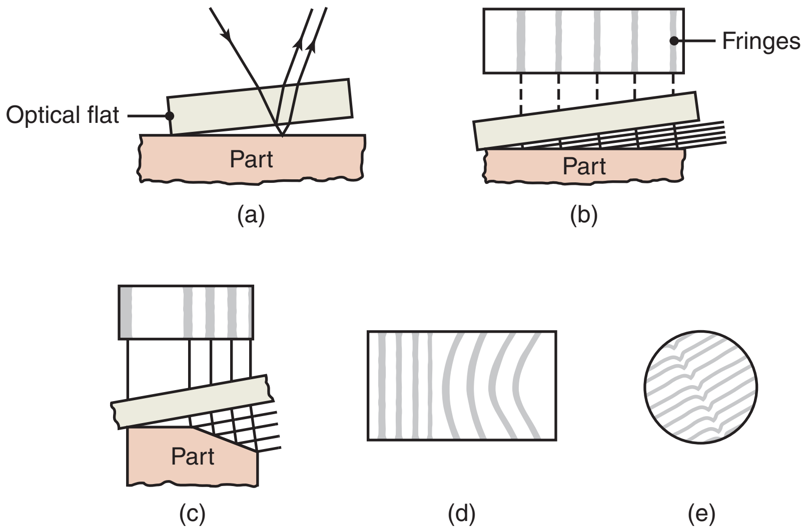


Figure 35.7: (a) Interferometry method for measuring flatness with an optical flat. (b) Fringes on a flat, inclined surface. An optical flat resting on a flat workpiece surface will not split the light beam, and no fringes will be present. (c) Fringes on a surface with two inclinations. Note that the greater the incline, the closer together are the fringes. (d) Curved fringe patterns, indicating curvatures on the workpiece surface. (e) Fringe pattern indicating a scratch on the surface.

Roundness. This feature is described as a deviation from true roundness, which, mathematically, is a circle. The term *out-of-roundness* (ovality) is actually more descriptive of the shape of the part (Fig. 35.8a) than the word *roundness*. True roundness is essential to the proper functioning of rotating shafts, bearing races, pistons, cylinders, and steel balls in bearings.

Methods of measuring roundness generally fall into two categories:

1. The round part is first placed on a *V-block* or between centers (Fig. 35.8b and c, respectively), then rotated while the point of a dial indicator is in continuous contact with the parts surface. After one full rotation of the part, the difference between the maximum and the minimum readings on the dial is noted; the difference is called **total indicator reading** (TIR) or **full indicator movement**. This method can also be used to measure the straightness (squareness) of the machined end faces of shafts, such as in a facing operation (Fig. 23.1e).

2. In *circular tracing*, the part is placed on a platform and its roundness is measured by rotating the platform (Fig. 35.8d); alternatively, the probe can be rotated around a stationary part.

Profile. *Profile* may be measured by such means as (a) comparing the surface with a *template* or a *profile gage* for checking conformity, as in measuring radii and fillets, and (b) using several dial indicators or similar instruments. For *advanced measuring machines* see Section 35.5.

Measuring Screw Threads and Gear Teeth. Threads can be measured by means of *thread gages*, of various designs, that compare the thread produced against a standard thread. Some of the commonly used gages are *threaded plug gages*, *screw-pitch gages*, micrometers with cone-shaped points, and *snap gages* (see Section 35.4.4), with anvils in the shape of threads. Gear teeth are measured with (a) instruments that are similar to dial indicators, (b) calipers (Fig. 35.9a), and (c) micrometers, using round pins or balls of various diameters (Fig. 35.9b). Advanced methods include using optical projectors and coordinate-measuring machines (Section 35.5).

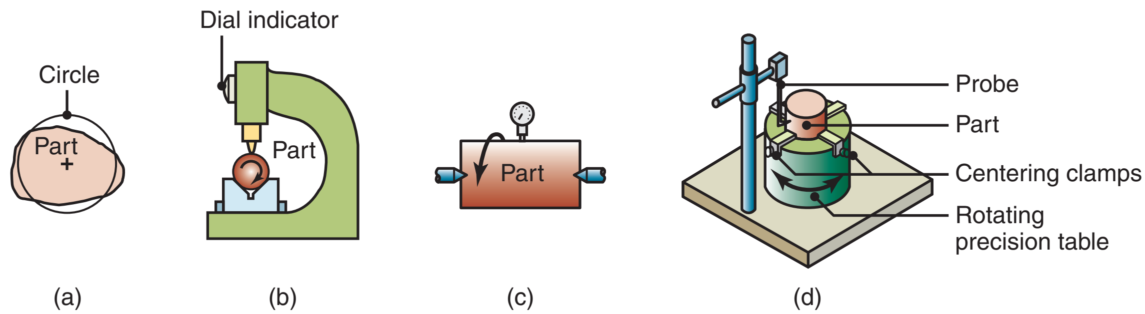


Figure 35.8: (a) Schematic illustration of out-of-roundness (exaggerated). Measuring roundness with (b) a V-block and dial indicator, (c) a round part supported on centers and rotated, and (d) circular tracing. *Source:* After F.T. Farago.

35.4.3 Optical Contour Projectors

Coordinate measuring machines, also called **optical comparators**, were developed in the 1940s to check the geometry of cutting tools for machining screw threads, now used for checking all profiles (Fig. 35.10). The part is mounted on a table or between centers, and the image is projected onto a screen at magnifications of 100× or higher. Linear and angular measurements are made directly on the screen, which is marked with reference lines and circles. For angular measurements, the screen can be rotated.

35.4.4 Gages

Gage Blocks. *Gage blocks* are individual square, rectangular, or round blocks of various sizes. For general use, they are made from heat-treated and stress-relieved alloy steels. The better gage blocks are made of ceramics (often zirconia) and chromium carbide; unlike steels, these materials do not rust, but they are brittle and must be handled accordingly. *Angle blocks* are used for angular gaging. Gage blocks have a flatness within 1.25 μm. Environmental temperature control is essential when gages are used for high-precision measurements.

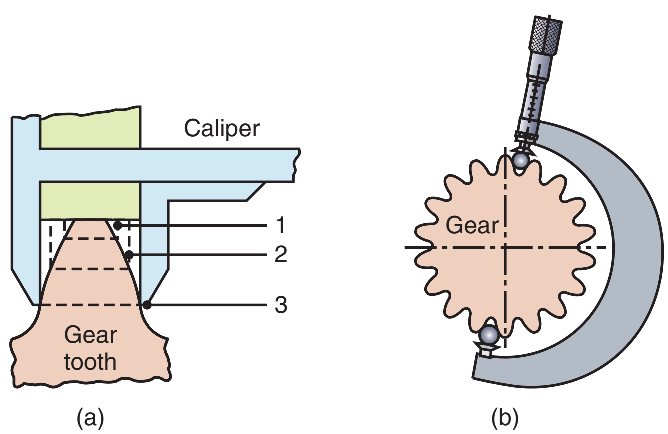


Figure 35.9: Measuring gear-tooth thickness and profile with (a) a gear-tooth caliper and (b) pins or balls and a micrometer.

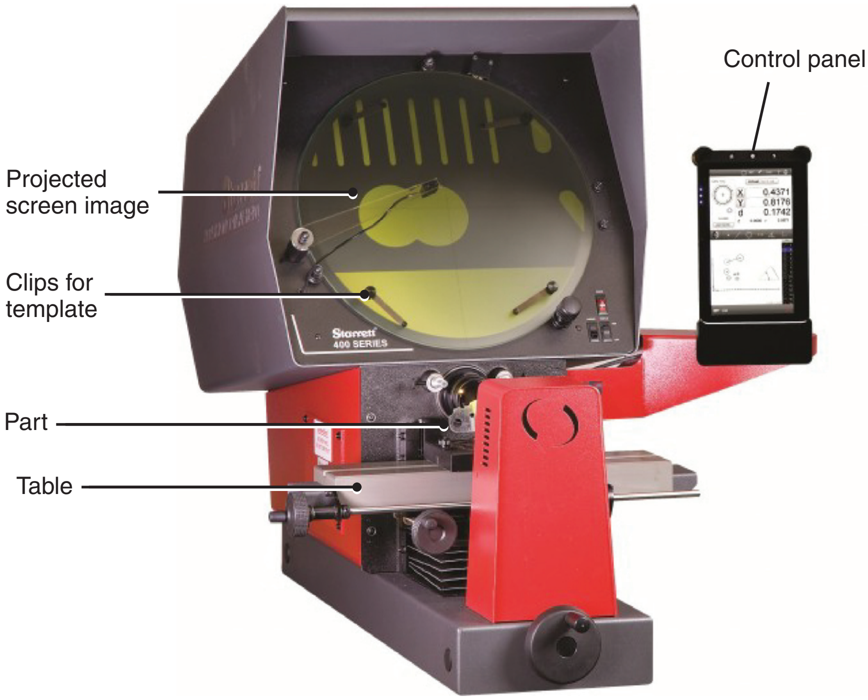


Figure 35.10: A horizontal-beam contour projector with a 0.4 m-diameter screen with 150-W tungsten halogen illumination. *Source:* Courtesy of L.S. Starrett Company, Precision Optical Division.

Fixed Gages. These gages are replicas of the shapes of the parts to be measured. Although *fixed* gages are easy to use and inexpensive, they indicate only whether a part is too small or too large as compared with an established standard; their use has become less common. Examples of fixed gages include the following:

- **Plug gages** are commonly used for holes (Fig. 35.11a and b). The *GO gage* is smaller than the *NOT GO* (or *NO GO*) gage and slides into any hole that has a dimension larger than the diameter of the gage. The *NOT GO* gage must not go into the same hole. Two gages are required for such measurements, although both may be on the same device, either at opposite ends or in two steps at one end (*step-type gage*). Plug gages are available for measuring internal tapers, splines, and threads (in which the *GO* gage must screw into the threaded hole).

- **Ring gages** (Fig. 35.11c) are used to measure round parts. Ring *thread gages* are used to measure external threads.

- **Snap gages** (Fig. 35.11d) are commonly used to measure external dimensions. They are made with adjustable gaging surfaces, for use with parts that have different dimensions. One of the gaging surfaces can be set at a different gap from the other, thus making the device a one-unit *GO*-and-*NOT-GO* gage.

Air Gages. The basic operation of an *air gage*, also called a **pneumatic gage**, is shown in Fig. 35.12a. The gage head (*air plug*) has two or more holes, typically 1.25 mm in diameter, through which pressurized air, supplied by a constant-pressure line, escapes. The smaller the gap between the gage and the hole, the more difficult it is for the air to escape, and hence, the higher is the back pressure. The back pressure, which is indicated by a pressure gage, is calibrated to measure the dimensional variations of holes.

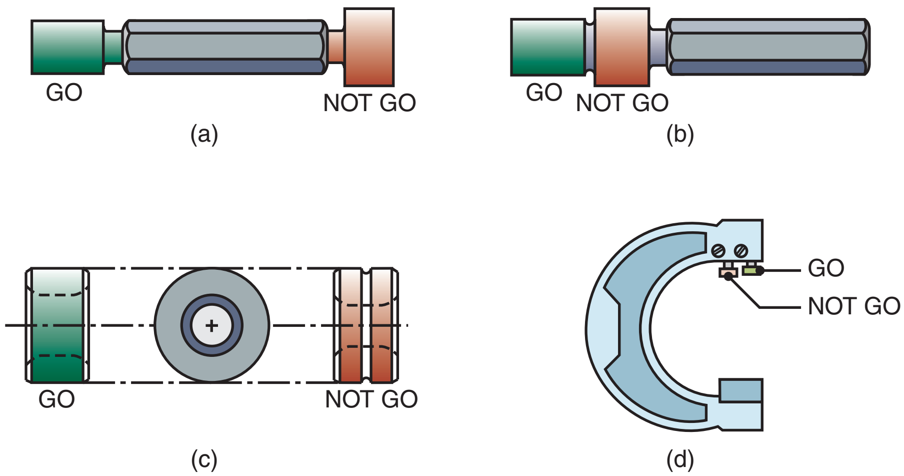


Figure 35.11: (a) Plug gage for holes, with *GO* and *NOT GO* on opposite ends of the gage. (b) Plug gage with *GO* and *NOT GO* on one end. (c) Plain ring gages for gaging round rods; note the difference in knurled surfaces to identify the two gages. (d) Snap gage with adjustable anvils.

Air gages can be rotated during use to indicate and measure any out-of-roundness of a hole. The outside diameters of such parts as pins and shafts also can be measured when the air plug is in the shape of a ring slipped over the part. In cases where a ring is not suitable, a fork-shaped gage head, with the airholes at the tips can be used (Fig. 35.12b). Various shapes of air heads, such as the conical head shown in Fig. 35.12c, can be made for use in specialized applications.

Air gages are easy to use, and have a resolution as fine as 0.125 μm. If the surface roughness of the part is too high, however, the readings may be unreliable. Their noncontacting nature and the low pressure involved has the benefit of not distorting or damaging the measured part.

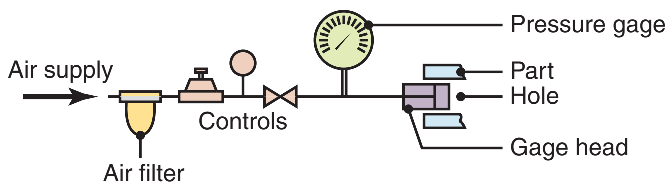
35.5 Modern Measuring Instruments and Machines

A wide variety of measuring instruments and gages has been developed, ranging from simple, hand-operated devices to computer-controlled machines with very large workspaces.

Electronic Gages. *Electronic gages* sense the movement of the contacting pointer through changes in the electrical resistance of a strain gage, inductance, or capacitance. The electrical signals are then converted and digitally displayed as linear dimensions. A handheld electronic gage for measuring bore diameters is shown in Fig. 35.13. When its handle is squeezed slightly, the gage can be inserted into the bore, and the diameter is read directly. A microprocessor-assisted electronic gage for measuring vertical length is shown in Fig. 35.14.

A commonly used electronic gage is the *linear-variable differential transformer* (LVDT), for measuring small displacements. Although more expensive than other types, electronic gages have such advantages as rapid response, digital readout, lower possibility of human error, versatility, flexibility, and the capability to be integrated into automated systems through microprocessors and computers.

Laser Micrometers. In this instrument, a laser beam scans the workpiece (Fig. 35.15), typically at a rate of 350/s. Laser micrometers are capable of resolutions as high as 0.125 μm. They are suitable not only for stationary parts, but also for in-line measurement of stationary, rotating, or vibrating parts, as well as parts in high-speed continuous production facilities. Moreover, because there is no physical contact, they can measure parts that are at elevated temperatures or are too flexible to be measured by other means. The laser beams can be of various types (such as scanning or rastoring for stationary parts), yielding **point cloud** descriptions of part surfaces. In a *point cloud*, a large number of surface points are measured and their coordinates are stored; through interpolation between points, a surface is then defined. Laser micrometers are of the shadow type or are charge-coupled device (CCD), based for in-line measurement while a part is in production.



(a)

(b)

Figure 35.12: (a) Schematic illustration of the principle of an air gage. (b) A conical head for air gaging; note the three small airholes on the conical surface. *Source:* (b) Courtesy of Stotz Gaging Co.

Figure 35.13: The TESA Brown & Sharpe INTRIMIK internal micrometer, with its high-precision machined measuring cone thread, features measuring bolts which are aligned to provide three-line contact. This original, patented design respects the Abbe principle. *Source:* Courtesy of TESA SA.

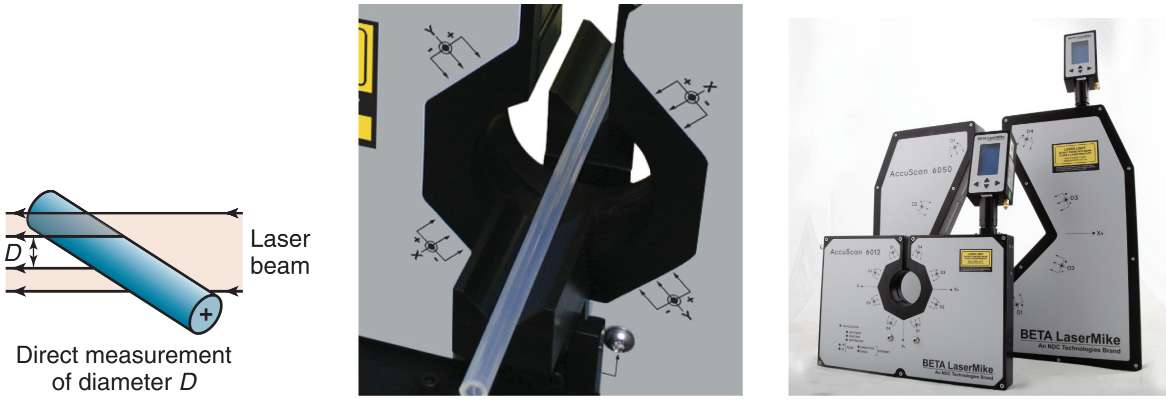
Laser Interferometry. This technique is used to check and calibrate machine tools for various geometric features during their assembly. The method has better accuracies than those of gages or indicators. Laser interferometers are also used to automatically compensate for positioning errors in coordinate-measuring machines and computer-numerical control machines.

35.5.1 Coordinate-measuring Machines

As schematically shown in Fig. 35.16a, a *coordinate-measuring machine* (CMM) basically consists of a platform on which the workpiece being measured is placed and then moved linearly or rotated. A *probe* (Fig. 35.16b; see also Fig. 25.6) is attached to a head that is capable of various movements, and records all measurements. In addition to the tactile probe shown in the figure, other types are *scanning, laser* (Fig. 35.16c), and *vision probes*, all of which are nontactile. A CMM for measuring a typical part is shown in Fig. 35.16d.

CMMs are very versatile and capable of recording measurements of complex profiles with high resolution (0.25 μm) and at high speed. They are built rigidly and ruggedly to resist environmental effects in plants, such as temperature variations and vibration. They can be placed close to machine tools for

Figure 35.14: A TESA MICRO-HITE electronic vertical-length measuring instrument with an accuracy of 1.8 μm. *Source:* Courtesy of TESA SA.



(a) (b) (c)

Figure 35.15: (a) and (b) Measurement of diameter using a laser scan micrometer; (b) Two types of laser micrometer. Such devices are available in a variety of sizes and can scan in multiple directions if needed. *Source:* (b) and (c) Courtesy of BETA Lasermike.

more efficient measurement and rapid feedback, so that processing parameters are corrected before the next part is made. Although large machines can be expensive, most machines with a touch probe and computer-controlled three-dimensional movement are inexpensive and suitable for use in small shops.

Example 35.1 A Coordinate Measurement Machine for Car Bodies

A large, horizontal CNC coordinate-measuring machine used for measuring all dimensions of a car body is shown in Fig. 35.17. This machine has a measuring range of 6 m $\times$ 1.6 m $\times$ 2.4 m high, a resolution of 0.1 μm, and a measuring speed of 5 mm/s. The system has temperature compensation, within a range from 16° to 26°C in order to maintain good measurement accuracy. For efficient measurements, the machine has two heads, with touch-trigger probes that are controlled simultaneously and have full three-dimensional movements. The probes are software controlled, and the machine is equipped with safety devices to prevent the probes from inadvertently hitting any part of the car body during their movements. The equipment shown around the base of the machine includes supporting hardware and software controlling all movements and it records all measurements.

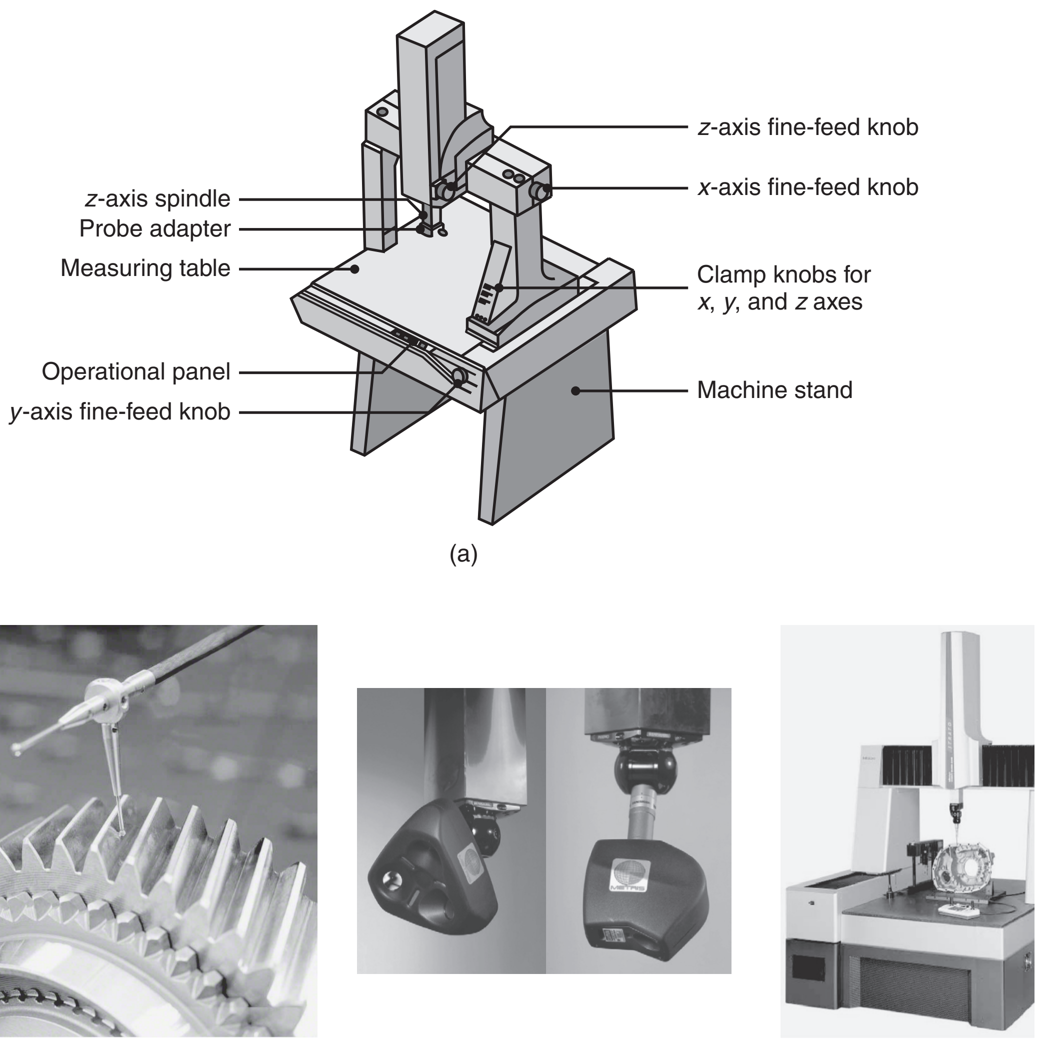
Source: Courtesy of Mitutoyo America Corporation.

35.6 Automated Measurement

Automated measurement is based on various **online sensor systems** that continuously monitor the dimensions of parts while they are being made and, if necessary, uses these measurements as input to make adjustments (see also Sections 36.12 and 37.7). Manufacturing cells and flexible manufacturing systems (Chapter 39) all have led to the adoption of advanced measuring techniques and systems.

To appreciate the importance of online monitoring of dimensions, consider the following: If a machine tool has been producing a certain part, consistently and with acceptable dimensions, what factors contribute to the subsequent deviation in the dimensions of the same part produced by the same machine? There are several technical and human factors that have to be considered:

- *Variations* in the properties and dimensions of the incoming material
- *Distortion* of the machine, because of thermal effects caused by such factors as changes in the ambient temperature, deterioration of metalworking fluids, and changes in machine tool bearings and various components
- *Wear* of tools, dies, and molds
- *Human errors*.



(a)

(b) (c) (d)

Figure 35.16: (a) Schematic illustration of a coordinate-measuring machine. (b) A touch signal probe. (c) Examples of laser probes. (d) A coordinate-measuring machine with a complex part being measured. *Source:* (b) through (d) Courtesy of Mitutoyo America Corp.

Figure 35.17: A large coordinate-measuring machine with two heads measuring various dimensions on a car body.

Due to one or more of these factors, the dimensions of parts made will vary randomly, thus making continuous monitoring during production necessary.

35.7 General Characteristics and Selection of Measuring Instruments

The characteristics and quality of measuring instruments are generally described by various specific terms, defined as follows (in alphabetical order):

- **Accuracy:** The degree of agreement of the measured dimension with its true magnitude
- **Amplification**, also called **magnification**: The ratio of an instruments output to the input dimension
- **Calibration**: The adjustment or setting of an instrument to give readings that are accurate within a reference standard
- **Drift**, also called **stability**: An instrument's capability to maintain its calibration over time
- **Linearity**: The accuracy of the readings of an instrument over its full working range
- **Precision**: Degree to which an instrument gives repeated measurement of the same standard
- **Repeat accuracy**: Same as accuracy, but repeated several times
- **Resolution**: Smallest dimension that can be read on an instrument
- **Rule of 10** (*gage maker's rule*): An instrument or gage should be 10 times more accurate than the dimensional tolerances of the part being measured; a factor of 4 is known as the *mil standard rule*
- **Sensitivity**: Smallest difference in dimension that an instrument can distinguish or detect
- **Speed of response**: How rapidly an instrument indicates a measurement, particularly when a number of parts are measured in rapid succession.

Selection of an appropriate measuring instrument for a particular application also depends on (a) the size and type of parts to be measured, (b) the environment, such as temperature, humidity, and dust, (c) operator skills required, and (d) cost of equipment.

35.8 Geometric Dimensioning and Tolerancing

Individually made parts eventually are assembled into subassemblies, assemblies, and products. For example, it is often taken for granted that when a thousand lawn mowers are manufactured and assembled, each part of the mower will fit properly with its intended components. Likewise, when a broken bolt on an old machine is to be replaced, an identical bolt is purchased. This is done with confidence, because the bolt is made according to certain standards and specifications. Thus, the dimensions of all similar bolts will vary by only a small, specified amount, called *tolerance*, and that it will not affect their function. In other words, all similar bolts are **interchangeable**.

Dimensional Tolerance. *Dimensional tolerance* is defined as the permissible or acceptable variation in the dimensions, such as height, width, depth, diameter, and angles, of a part. Tolerances are unavoidable, because it is virtually impossible and unnecessary to manufacture two parts that have precisely the same dimensions. Moreover, because close dimensional tolerances can significantly increase the product cost (see Fig. 40.2), an unnecessarily narrow tolerance range is economically undesirable. For many parts, however, close tolerances *are* necessary for their proper functioning. Examples are precision measuring instruments, hydraulic pistons, rolling-element bearings, and turbine blades for jet aircraft engines.

Measuring dimensions and features of parts rapidly and reliably can be a challenging task. Each of the six million parts of a Boeing 747-400 aircraft, for example, required measuring about 25 features, representing a total of 150 million measurements. Surveys have indicated that the dimensional tolerances on state-of-the-art parts made have been shrinking by a factor of three every 10 years, and that this trend will continue (see Fig. 25.17).

Importance of Tolerance Control. Surfaces that are not functional do not need close tolerance control; dimensional tolerances become important only when a part is to fit and be assembled with another part. For example, note that the accuracy of the holes and the distance between the holes for a connecting rod (Fig. 14.8a) are far more critical than the rod's width and thickness at various locations along its length.

To recognize the importance of dimensional tolerances, consider the assembly of a simple round shaft (axle) and a wheel with a round hole. Assume that the shaft's diameter is 25 mm (Fig. 35.18). The wheel is a casting, with a 25 mm diameter hole drilled into it. Will the rod fit into the hole without forcing it or will it be loose in the hole? The 25 mm dimension is the **nominal size** of the shaft. If a second shaft is made, at a different time or is selected randomly from a large lot, each shaft will likely have a slightly different diameter (see also Chapter 36). Machines with the same setup may produce rods of slightly different diameters, depending on various factors, such as speed of operation, temperature, lubrication, and variations in the properties of the incoming material.

Certain terminology has been established over the years to specifically define geometric quantities. One such system is the International Organization for Standardization (ISO) system, shown in Fig. 35.18. Note that both the shaft and the hole have minimum and maximum diameters, the difference being the tolerance for each member. A proper CAD representation or engineering drawing should specify these parameters with numerical values, as shown in Fig. 35.19.

The range of dimensional tolerances achievable in manufacturing operations is given in various figures and tables throughout this book. As expected, there is a general relationship between tolerances and part

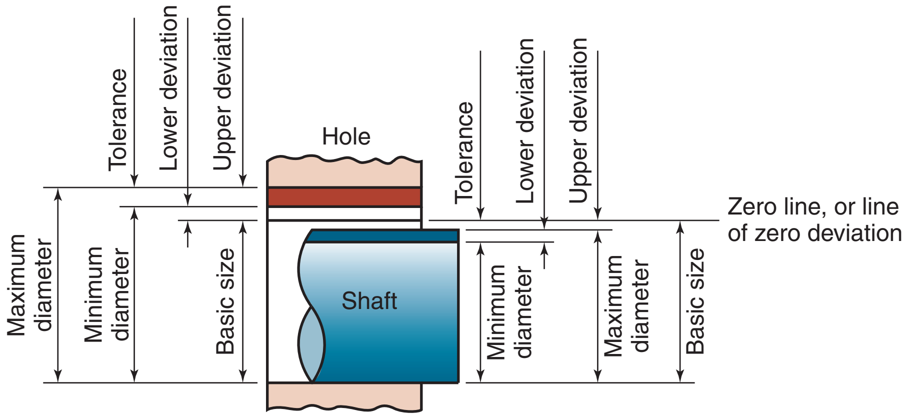


Figure 35.18: Basic size, deviation, and tolerance on a shaft, according to the ISO system.

size (Fig. 35.20) and between tolerances and surface finish of parts manufactured by various processes (Fig. 35.21). Note the wide range of tolerances and surface finish and that the larger the part, the greater is its tolerance range.

Definitions. Several terms are used to describe features of dimensional relationships between mating parts. Details of the definitions are available in the ANSI/ASME B4.2, ANSI/ASME Y14.5, and ISO/TC10/SC5 standards. The commonly used terms for geometric characteristics are defined as follows, in alphabetical order:

- **Allowance,** also called **functional dimension** or **sum dimension**: The specified difference in dimensions between mating parts
- **Basic size**: Dimension from which limits of size are derived with the use of tolerances and allowances
- **Bilateral tolerance**: Deviation (plus or minus) from the basic size
- **Clearance**: The space between mating parts
- **Clearance fit**: Fit that allows for rotation or sliding between mating parts
- **Datum**: A theoretically exact axis, point, line, or plane
- **Feature**: A physically identifiable portion of a part, such as hole, slot, pin, or chamfer

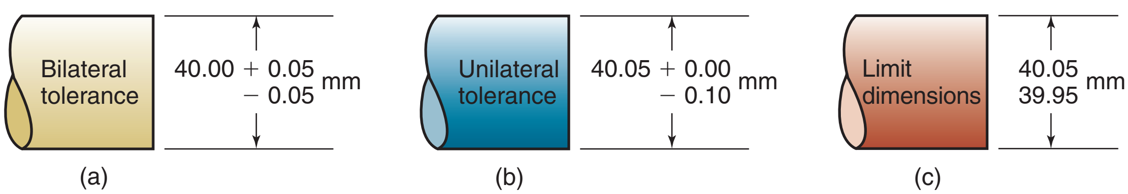


Figure 35.19: Various methods of assigning tolerances on a shaft: (a) bilateral tolerance, (b) unilateral tolerance, and (c) limit dimensions.

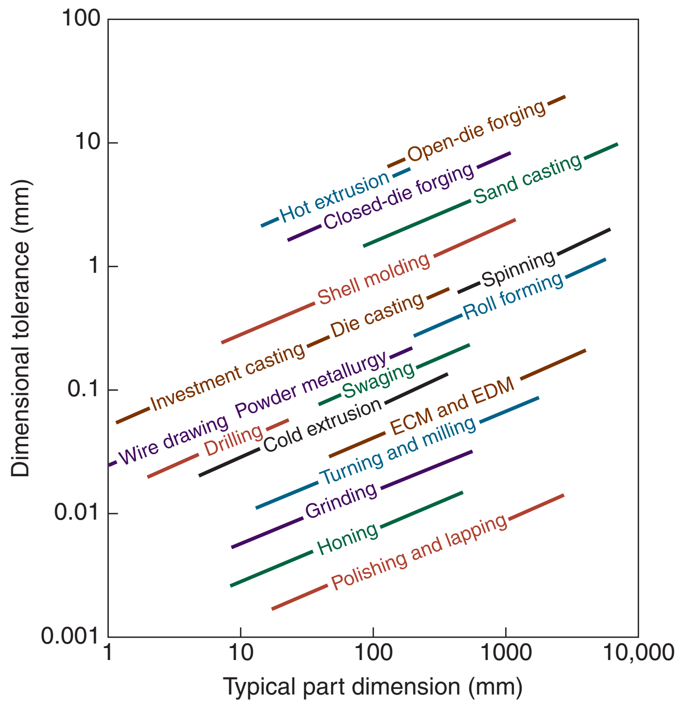


Figure 35.20: Dimensional tolerances as a function of part size for various manufacturing processes; note that because many factors are involved, there is a broad range for tolerances.

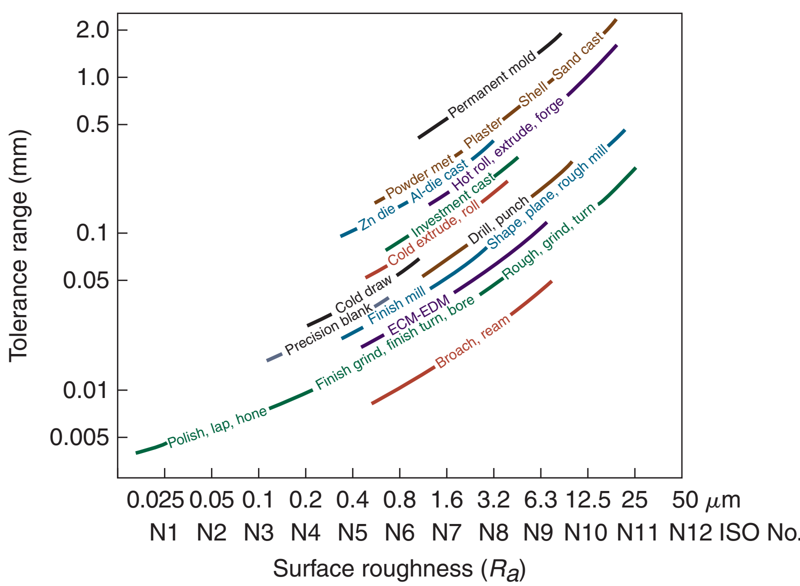


Figure 35.21: Dimensional tolerance range and surface roughness obtained in various manufacturing processes; these tolerances apply to a 25-mm workpiece dimension. *Source:* After J.A. Schey.

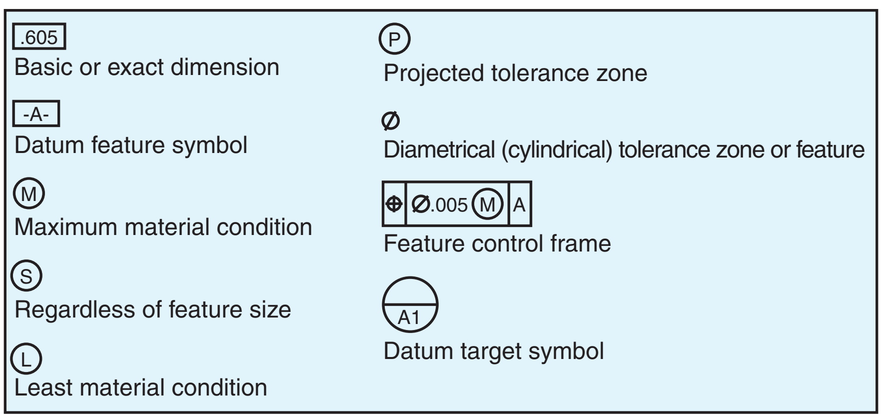
- **Fit**: The range of looseness or tightness that can result from the application of a specific combination of allowance and tolerance in the design of mating-part features
- **Geometric tolerances**: Tolerances that involve shape features of the part
- **Hole-basis system**, also called **standard hole practice** or **basic hole system**: Tolerances based on a zero line on the hole
- **Interference**: Negative clearance
- **Interference fit**: A fit having limits of size so prescribed that an interference always results when mating parts are assembled
- **International tolerance** (IT) **grade**: A group of tolerances that vary with the basic size of the part, but provide the same relative level of accuracy within a grade
- **Limit dimension**, also called **limits**: The maximum and minimum dimensions of a part
- **Maximum material condition** (MMC): The condition whereby a feature of a certain size contains the maximum amount of material within the stated limits of that size
- **Nominal size**: An approximate dimension that is used for the purpose of general identification
- **Positional tolerancing**: A system of specifying the true position, size, and form of the features of a part, including allowable variations
- **Shaft basis system**, also called **standard shaft practice** or **basic shaft system**: Tolerances based on a zero line on the shaft
- **Transition fit**: A fit with small clearance or interference that allows for accurate location of mating parts
- **Unilateral tolerancing**: Deviation from the nominal dimension in one direction only
- **Zero line**: Reference line along the basic size from which a range of tolerances and deviations are specified.

Because the dimensions of holes are more difficult to control than those of shafts, the hole-basis system is commonly used for specifying tolerances in shaft and hole assemblies. The symbols used to indicate geometric characteristics are shown in Fig. 35.22a and b.

Limits and Fits. *Limits and fits* are essential in specifying dimensions for holes and shafts (see Tables 35.1 and 35.2). There are two standards for limits and fits, as described by the American National Standards Institute (see ANSI/ASME B4.1, B4.2, and B4.3). One standard is based on the traditional inch unit; the other is based on the metric unit, and it has been developed in greater detail.

Type of feature	Type of tolerance	Characteristic	Symbol
Individual (no datum reference)	Form	Flatness	⏥
		Straightness	—
		Circularity (roundness)	○
		Cylindricity	⌭
Individual or related	Profile	Profile of a line	⌒
		Profile of a surface	⌓
Related (datum reference required)	Orientation	Perpendicularity	⊥
		Angularity	∠
		Parallelism	//
	Location	Position	⌖
		Concentricity	◎
	Runout	Circular runout	↗
		Total runout	⌰

(a)



(b)

Figure 35.22: Geometric characteristic symbols to be indicated on engineering drawings of parts to be manufactured. *Source:* Courtesy of The American Society of Mechanical Engineers.

Table 35.1: Classes of Fit.

Class	Description	Type	Applications
1	Loose	Clearance	Where accuracy is not essential, such as in building and mining equipment
2	Free	Clearance	In rotating journals with speeds of 600 rpm or greater, such as in engines and some automotive parts
3	Medium	Clearance	In rotating journals with speeds under 600 rpm, such as in precision machine tools and precise automotive parts
4	Snug	Clearance	Where small clearance is permissible and where mating parts are not intended to move freely under load
5	Wringing	Interference	Where light tapping with a hammer is necessary to assemble the parts
6	Tight	Interference	In semipermanent assemblies suitable for drive or shrink fits on light sections
7	Medium	Interference	Where considerable pressure is required for assembly and for shrink fits of medium sections; suitable for press fits on generator and motor armatures and for automotive wheels
8	Heavy force or shrink	Interference	Where considerable bonding between surfaces is required, such as locomotive wheels and heavy crankshaft disks of large engines

Table 35.2: Recommended Tolerances in Millimeter for Classes of Fit.

Class	Allowance	Interference	Hub tolerance	Shaft tolerance
1	$0.0073d^{2/3}$	—	$0.0216d^{1/3}$	$0.0216d^{1/3}$
2	$0.0041d^{2/3}$	—	$0.0112d^{1/3}$	$0.0112d^{1/3}$
3	$0.0026d^{2/3}$	—	$0.0069d^{1/3}$	$0.0069d^{1/3}$
4	0.000	—	$0.0052d^{1/3}$	$0.0035d^{1/3}$
5	—	0.000	$0.0052d^{1/3}$	$0.0035d^{1/3}$
6	—	$0.00025d$	$0.0052d^{1/3}$	$0.0052d^{1/3}$
7	—	$0.0005d$	$0.0052d^{1/3}$	$0.0052d^{1/3}$
8	—	$0.0010d$	$0.0052d^{1/3}$	$0.0052d^{1/3}$

Summary

- Modern manufacturing technology requires measuring instrumentation with several advanced features and characteristics.

- Various devices are now available for measurements, from simple gage blocks to electronic gages with high resolution. Major advances have been made in automated measurement, linking measuring devices to microprocessors and computers for accurate in-process control of manufacturing operations.

- Reliable linking, monitoring, display, distribution, and manipulation of data are important factors, as are the significant costs involved in implementing them.

- Dimensional tolerances and their selection are important factors in manufacturing. The smaller the range of specified tolerances, the higher is the cost of production; thus, tolerances should be as broad as possible but still maintaining the functional and service requirements of the product.

Key Terms

Air gage
Analog instruments
Autocollimator
Bevel protractor
Comparative length–measuring instruments
Coordinate-measuring machine
Dial indicator
Diffraction gratings
Digital instruments
Dimensional tolerance
Electronic gages
Fits
Fixed gage
Gage block
Interferometry
Laser micrometer
Limits
Line-graduated instruments
Measurement standards
Micrometer
Optical contour projector
Optical flat
Plug gage
Pneumatic gage
Precision
Resolution
Ring gage
Sensitivity
Snap gage
Tolerance
Total indicator reading
Vernier caliper

Bibliography

Andersson, A., **Measurement Technology for Process Automation**, CRC Press, 2017.
Bentley, J.P., **Principles of Measurement Systems**, 4th ed., Prentice Hall, 2005.
Bucher, J.L. (ed.), **The Metrology Handbook**, 2nd ed., Quality Press, 2012.
Cogorno, G., **Geometric Dimensioning and Tolerancing for Mechanical Design**, 2nd ed., McGraw-Hill, 2011.
Curtis, M.A., **Handbook of Dimensional Measurement**, 5th ed., Industrial Press, 2013.
Hocken, R.J., and Pereira, P.H., **Coordinate Measuring Machines and Systems**, 2nd ed., CRC Press, 2011.
Krulikowski, A., **Fundamentals of Geometric Dimensioning and Tolerancing**, 3rd ed., Delmar, 2012.
Madsen, D.A., and Madsen, D.P., **Geometric Dimensioning and Tolerancing**, 9th ed., Goodheart-Willcox, 2012.
Morris, A.S., and Langari, R., **Measurement and Instrumentation: Theory and Application**, 2nd ed., Academic Press, 2015.
Rabinowich, S.G., **Measurement Errors and Uncertainties: Theory and Practice**, Springer, 2010.
Whitehouse, D.J., **Handbook of Surface and Nanometrology**, 2nd ed., CRC Press, 2010.

Review Questions

35.1. What is metrology?

35.2. Explain how a meter is defined and measured.

35.3. Explain what is meant by standards for measurement.

35.4. What is the basic difference between direct-reading and indirect-reading linear measurements? Name the instruments used in each category.

35.5. What is meant by comparative length measurement?

35.6. Explain how flatness is measured. What is an optical flat?

35.7. Describe the principle of an optical comparator.

35.8. Why have coordinate measuring machines become important instruments?

35.9. What is the difference between a plug gage and a ring gage?

35.10. What are dimensional tolerances? Why is their control important?

35.11. Why is a sine bar known by that name?

35.12. Explain the difference between tolerance and allowance.

35.13. What is the difference between bilateral and unilateral tolerance?

35.14. How is straightness measured?

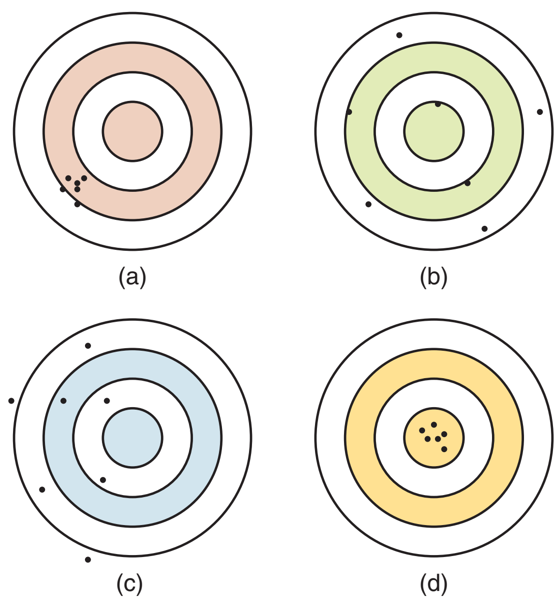
35.15. When is a clearance fit desirable? An interference fit?

35.16. What factors contribute to deviations in the dimensions of the same type of parts made by the same machine?

Qualitative Problems

35.17. Why are the words "accuracy" and "precision" often incorrectly interchanged?

35.18. Review the following results from an archery competition. Indicate which of the targets display an archer with (a) precise placement of the arrows; (b) accurate placement of the arrows.



35.19. Why do manufacturing processes produce parts with a wide range of tolerances? Explain, giving several examples.

35.20. Explain the need for automated inspection.

35.21. Dimensional tolerances for nonmetallic parts usually are wider than for metallic parts. Explain why. Would this also be true for ceramics parts?

35.22. Comment on your observations regarding Fig. 35.21. Why does dimensional tolerance increase with increasing surface roughness?

35.23. Review Fig. 35.20, and comment on the range of tolerances and part dimensions produced by various manufacturing processes.

35.24. In the game of darts, is it better to be accurate or to be precise? Explain.

35.25. What are the advantages and limitations of *GO* and *NOT GO* gages?

35.26. Comment on your observations regarding Fig. 35.19.

35.27. What are gage blocks? Explain three methods that gage blocks can be used in metrology.

35.28. Why is it important to control temperature during the measurement of dimensions? Explain, with examples.

35.29. Describe the characteristics of electronic gages.

35.30. What method would you use to measure the thickness of a foam-rubber part? Explain.

35.31. Review Fig. 35.20 and explain why the dimensional tolerance is related to the part dimensions.

35.32. Review Fig. 35.21 and give reasons that there is a range of tolerance and surface roughness for each manufacturing process.

Quantitative Problems

35.33. Assume that a steel rule expands by 0.07% due to an increase in environmental temperature. What will be the indicated diameter of a shaft with a diameter of 30 mm at room temperature?

35.34. If the same steel rule as in Problem 35.33 is used to measure aluminum extrusions, what will be the indicated diameter at room temperature? What if the part were made of a thermoplastic?

35.35. A shaft must meet a design requirement of being at least 28 mm in diameter, but it can be 0.38 mm oversized. Express the shaft's tolerance as it would appear on an engineering drawing.

35.36. Review Table 35.2 and plot the tolerance of a hub and the shaft it mounts on as a function of diameter.

Synthesis, Design, and Projects

35.37. Describe your thoughts on the merits and limitations of digital measuring equipment over analog instruments. Give specific examples.

35.38. Take an ordinary vernier micrometer (see Fig. 35.2a) and a simple round rod. Ask five of your classmates to measure the diameter of the rod with this micrometer. Comment on your observations.

35.39. Obtain a digital micrometer and a steel ball of, say, 6.25 mm diameter. Measure the diameter of the ball when it (a) has been placed in a freezer, (b) has been put into boiling water, and (c) when it has been held in your hand for different lengths of time. Note the variations, if any, of measured dimensions, and comment on them.

35.40. Repeat Problem 35.39, but with the following parts: (a) the plastic lid of a small jar, (b) a thermoset part such as the knob or handle from the lid of a saucepan, (c) a small juice glass, and (d) an ordinary rubber eraser.

35.41. What is the significance of the tests described in Problems 35.39 and 35.40?

35.42. Explain the relative advantages and limitations of a tactile probe versus a laser probe.

35.43. Make simple sketches of some forming- and cutting-machine tools (as described in Parts III and IV of the book) and integrate them with the various types of measuring equipment described in this chapter. Comment on the possible difficulties involved in doing so.

35.44. Inspect various parts and components in consumer products, and comment on how tight dimensional tolerances have to be in order for these products to function properly.

35.45. As you know, very thin sheet-metal parts can distort differently when held from various locations and edges of the part, just as a thin paper plate or aluminum foil does. How, then, could you use a coordinate-measuring machine for "accurate" measurements? Explain.

35.46. Explain how you would justify the considerable cost of a coordinate-measuring machine such as that shown in Fig. 35.17.

35.47. Explain how you would measure the dimensions of an extruded hexagonal cross-section, and how you would indicate deviation from the hexagon's form.

35.48. Conduct an Internet search, and make a list of the method you would use for length measurement as a function of the length.

35.49. How are dimensions of MEMS devices measured?

Chapter 36

Quality Assurance, Testing, and Inspection

- This chapter outlines the procedures used to ensure the manufacture of high-quality products. It describes the mathematical tools and inspection techniques that have been developed, including total quality management, Taguchi methods, and the Deming and Juran approaches.

- Statistical methods of quality control and control charts are then described, including acceptance sampling in order to ensure that production meets quality standards.
- The chapter concludes with a discussion of the methods used in destructive and nondestructive testing of materials and products.

36.1 Introduction

Manufactured products develop certain external and internal characteristics that result, in part, from the type of production processes employed. External characteristics most commonly involve dimensions, size, and surface finish and integrity considerations. Internal characteristics include defects such as porosity, impurities, inclusions, phase transformations, residual stresses, embrittlement, cracks, and debonding of laminations in composite materials.

Some of these defects may exist in the original material (stock), while others are introduced or induced during the particular manufacturing operation. Before they are marketed, manufactured parts and products are inspected in order to:

- Ensure dimensional accuracy so that parts fit properly into other components during assembly; recall that commercial aircraft have millions of parts to be assembled
- Identify products whose failure or malfunction may have serious implications, including bodily injury or even fatality. Typical examples are switches, brakes, grinding wheels, railroad wheels, welded joints, turbine blades, medical products of all kinds, and pressure vessels.

Product quality always has been one of the most important aspects of manufacturing operations. In view of a global competitive market, *continuous improvement in quality* is a major priority. In Japan, the single term **kaizen** is used to signify *never-ending improvement*. *Quality must be built into a product* and not merely considered *after* the product already has been made. Thus, close cooperation and communication among design and manufacturing engineers and the direct involvement and encouragement of company management are vital.

Major advances in quality engineering and productivity have been made over the years, largely because of the efforts of quality experts such as W.E. Deming, G. Taguchi, and J.M. Juran. The importance of the *quality, reliability*, and *safety* of products in a global economy is now internationally recognized, as evidenced by the establishment of various **ISO and QSO standards** and nationally by the Malcolm Baldrige National Quality Award in the United States.

36.2 Product Quality

What is quality? Unlike most technical terms, quality is difficult to define precisely; generally, it can be defined as a *product's fitness for use*. Thus, quality is a broad-based characteristic or property, and its factors consist not only of well-defined technical considerations, but also of subjective opinions. Several aspects of quality that generally are identified are performance, durability, reliability, robustness, availability, cost, and serviceability, as well as aesthetics and perceived quality.

Consider, for example, the following: (a) The handle on a kitchen utensil is installed improperly or its handle discolors or cracks during its normal use, (b) a weighing scale functions erratically, and (c) a machine tool cannot maintain the specified dimensional tolerances because of lack of stiffness or poor construction. These examples indicate that the product is of low quality. Thus, the general perception is that a high-quality product is one that performs its functions reliably over a long time without breaking down or requiring repairs (see Table I.4 in the General Introduction). The level of quality that a manufacturer chooses for its products depends on the market for which the products are intended. Low-quality, low-cost tools, for example, have their own world market niche.

Contrary to general public perception, high-quality products do not necessarily cost more, especially considering the fact that poor-quality products:

- present difficulties in assembling and maintaining components
- require in-field repairs (see Table I.5)
- have the significant built-in cost of customer dissatisfaction (see Section 36.5).

As described in Section 40.10, the total product cost depends on several variables, including the level of automation in the manufacturing plant. There are many ways for engineers to review and modify overall product design and manufacturing processes in order to minimize a product's cost without affecting its quality. Thus, quality standards are essentially a balance among several considerations; this balance is also called **return on quality** (ROQ) and usually includes some *limit* on the expected life of the product.

36.3 Quality Assurance

Quality assurance is the total effort made by a manufacturer to ensure that its products conform to a detailed set of specifications and standards. It can be defined as all actions necessary to ensure that quality requirements will be satisfied. **Quality control** is the set of operational techniques used to fulfill quality requirements.

The standards cover several types of parameters, such as dimensions, surface finish, tolerances, composition, and color, as well as mechanical, physical, and chemical properties and characteristics. In addition, standards usually are written to ensure proper *assembly*, using *interchangeable* defect-free components and resulting in a product that performs as intended by its designers.

A major aspect of quality assurance is the capability to *analyze* defects as they occur on the production line, and *eliminate* or reduce them promptly to acceptable levels. In an even broader sense, quality assurance involves *evaluating* the product and its customer satisfaction. The sum total of all these activities is referred to by terms such as **total quality control** and **total quality management**.

In order to control quality, it is essential to be able to

- *Measure* the level of quality quantitatively
- *Identify* all of the material and process variables that can be controlled.

The quality level built in during production can then be checked by *continuously* inspecting the product in order to determine whether it meets the relevant specifications for dimensional tolerances, surface finish, defects, and various other characteristics.

36.4 Total Quality Management

Total quality management (TQM) is a *systems approach*, in that both management and employees make a concerted effort to consistently manufacture high-quality products. *Defect prevention* rather than *defect detection* is the major goal.

Leadership and *teamwork* in the organization are essential to ensure that the goal of **continuous improvement** in manufacturing operations is foremost, because they *reduce product variability* and, again, they improve customer satisfaction. The TQM concept also requires *control of the processes*, and not the *control of parts produced*, so that process variability can be reduced and no defective parts are allowed to continue through the production line.

Table 36.1: Deming's 14 Points.

1.	Create constancy of purpose toward improvement of product and service.
2.	Adopt the new philosophy: refuse to accept defects.
3.	Cease dependence on mass inspection to achieve quality.
4.	End the practice of awarding business on the basis of price tag.
5.	Improve the system of production and service constantly and forever, to improve quality and productivity and thus constantly decrease cost.
6.	Institute training for the requirements of a particular task, and document the requirements for future training.
7.	Institute leadership, as opposed to supervision.
8.	Drive out fear so that everyone can work effectively.
9.	Break down barriers between departments.
10.	Eliminate slogans, exhortations, and targets for zero defects and new levels of productivity.
11.	Eliminate quotas and management by numbers, or numerical goals. Substitute leadership.
12.	Remove barriers that rob the hourly worker of pride of workmanship.
13.	Institute a vigorous program of education and self-improvement.
14.	Put everyone in the company to work to accomplish the transformation.

Quality Circle. A *quality circle*, first established in Japan in 1962, consists of groups of employees (workers, supervisors, and managers) who volunteer to meet regularly to discuss how to improve and maintain product quality at *all* stages of the manufacturing operation. Worker involvement, responsibility, and creativity, as well as a team effort, are emphasized. Comprehensive *training* is provided so that the worker can become conscious of quality and also be capable of analyzing statistical data, identifying the causes of poor quality, and taking immediate action to correct the situation. Experience has indicated that quality circles are more effective in *lean-manufacturing* environments, described in Section 39.7.

Quality Engineering as a Philosophy. Experts in quality control have placed many of the quality-control concepts and methods into a larger perspective. Notable among these experts have been Deming, Juran, and Taguchi, whose philosophies of quality and product cost have had, and continue to have, a major impact on modern manufacturing.

36.4.1 Deming Methods

During World War II, W.E. Deming (1900–1993), an American statistician, and several others, developed new methods of *statistical process control* for wartime-industry manufacturing plants. The methods arose from the recognition that there were *variations* (a) in the performance of machines and of people and (b) in the quality and dimensions of raw materials (stock). The efforts of these pioneers involved not only statistical methods of analysis, but also a new way of looking at manufacturing operations from the perspective of *improving quality while lowering costs*.

Deming recognized that manufacturing organizations are *systems* of management, workers, machines, and products. He placed great emphasis on communication, direct worker involvement, and education in statistics and modern manufacturing technology. His basic ideas are summarized in the well-known *14 points*, given in Table 36.1. These points are not to be seen as a checklist or menu of tasks; they are what Deming recognized as *characteristics* of *companies* that produce high-quality goods.

36.4.2 Juran Methods

A contemporary of Deming, J.M. Juran (1904–2008), an electrical engineer and management consultant, emphasized the importance of:

- Recognizing quality at all levels of an organization, including upper management
- Fostering a responsive corporate culture
- Training all personnel in how to plan, control, and improve quality.

The main concern of the top management in an organization is business and management, whereas those in quality control are basically concerned with technology. These different worlds have, in the past, often been at odds, and their conflicts have led to quality problems. Planners determine who the customers are and their needs. An organization’s customers may be external (end users who purchase the product or service), or they may be internal (different parts of an organization that rely on other segments of the organization to supply them with products and services). The planners then develop product and process designs to respond to the customer’s needs. The plans are turned over to those in charge of operations, who then become responsible for implementing both quality control and continued improvement in quality.

36.5 Taguchi Methods

In G. Taguchi’s (1924–2012; an engineer and a statistician) methods, high quality and low costs are achieved by combining engineering and statistical techniques to optimize product design and manufacturing processes. *Taguchi methods* is now a term that refers to the approaches developed to manufacture high-quality products. One fundamental viewpoint put forward is the quality challenge facing manufacturers: provide products that delight your customers, and to do so, manufacturers should offer products with the following product characteristics:

- High reliability
- Perform the desired functions well
- Good appearance
- Inexpensive
- Upgradeable
- Available in the quantities desired when needed
- Robust over their intended life (see Section 36.5.1).

These characteristics clearly are the goals of manufacturers striving to provide high-quality products. Although it is very challenging to provide all of these characteristics, excellence in manufacturing is undeniably a prerequisite.

Taguchi also contributed to the approaches that are used to document quality, recognizing that any deviation from the optimum state of a product represents a financial loss, because of such factors as reduced product life, performance, and economy. *Loss of quality* is defined as the *financial loss* to society after the product is shipped. Loss of quality results in the following problems:

- Poor quality leads to customer dissatisfaction
- Costs are incurred in servicing and repairing defective products, especially when such repairs have to be made in the field
- The manufacturer’s credibility in the marketplace is diminished
- The manufacturer eventually loses its share of the market.

The Taguchi methods of **quality engineering** emphasize the importance of

- **Enhancing cross-functional team interaction**: Design engineers and manufacturing engineers communicate with each other in a common language. They quantify the relationships between design requirements and manufacturing process selection.
- **Implementing experimental design**: The factors involved in a process or operation and their interactions are studied simultaneously.

In *experimental design*, the effects of controllable and uncontrollable variables on the product are identified. This approach minimizes variations in product dimensions and properties and, ultimately, brings the mean to the desired level. The methods used for experimental design are complex, and involve using *factorial design* and *orthogonal arrays*, both of which reduce the number of experiments required. These methods also are capable of identifying the effects of variables that cannot be controlled (called *noise*), such as changes in environmental conditions in a plant.

The use of factorial design and orthogonal arrays results in (a) the rapid *identification* of the controlling variables, referred to as *observing main effects*, and (b) the ability to determine the best method of process control. Control of these variables sometimes requires new equipment or major modifications to existing equipment. Thus, for example, variables affecting dimensional tolerances in machining a particular component can readily be identified and, whenever possible, the correct cutting speed, feed, cutting tool, and cutting fluids can be specified.

36.5.1 Robustness

Another aspect of quality, originally suggested by Taguchi, is *robustness*. A robust design, process, or system is one that continues to function, within acceptable parameters, despite variabilities (often unanticipated) in its environment. In other words, its outputs (such as its performance) have *minimal sensitivity* to its input variations (such as variations in environment, load, and power source).

In a robust design, for example, a part will function sufficiently well even if the loads applied, or their directions, exceed anticipated values. Likewise, a robust machine or a system will undergo minimal deterioration in performance even if it experiences variations in environmental conditions, such as temperature, humidity, air quality, and vibrations. A robust machine will also have no significant reduction in its performance over its life, whereas a less robust design will perform less efficiently as time passes.

As a simple illustration of a robust design, consider a sheet-metal mounting bracket to be attached to a wall with two bolts (Fig. 36.1a). The positioning of the two mounting holes on the bracket will include some error due to the manufacturing process involved; this error will prevent the top edge of the bracket from being perfectly horizontal.

A more robust design is shown in Fig. 36.1b, in which the mounting holes have been moved twice as far apart as in the original design. Even though the precision of hole location remains the same, and the

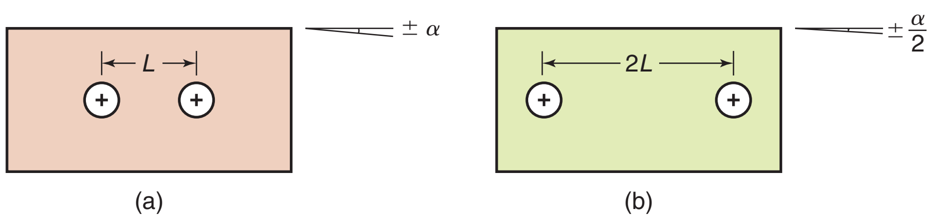


Figure 36.1: A simple example of robust design. (a) Location of two mounting holes on a sheet-metal bracket, where the deviation keeping the top surface of the bracket from being perfectly horizontal is $\pm\alpha$. (b) New locations of holes; the deviation (keeping the top surface of the bracket from being perfectly horizontal) is now reduced to $\pm\alpha/2$.

manufacturing cost also is the same, the variability in the top edge of the bracket (from the horizontal) has now been reduced by one-half. If the bracket is subjected to vibration, however, the bolts may loosen over time. An even more robust design approach would then be to use an adhesive to hold the bolt threads in place or to use a different type of fastener that would not loosen over time (see also Section 32.5).

36.5.2 Taguchi Loss Function

The *Taguchi loss function*, introduced in the early 1980s, is a tool for comparing quality on the basis of minimizing variations. It calculates the increasing loss to the company when the component deviates from the design objective. This function is defined as a parabola where one point is the cost of replacement (including shipping, scrapping, and handling costs) at an extreme of the tolerances, while a second point corresponds to zero loss at the design objective.

Mathematically, the loss cost can be written as

$$\text{Loss cost} = k\left[(Y - T)^2 + \sigma^2\right], \tag{36.1}$$

where Y is the mean value from manufacturing, T is the target value from design, σ is the standard deviation of parts from manufacturing (see Section 36.7), and k is a constant, defined as

$$k = \frac{\text{Replacement cost}}{(\text{LSL} - T)^2}, \tag{36.2}$$

where LSL is the lower specification limit. When the lower (LSL) and upper (USL) specification limits are the same distance from the mean (i.e., the tolerances are balanced), either of the limits can be used in this equation.

Example 36.1 Production of Polymer Tubing

Given: High-quality polymer tubes are being produced for medical applications in which the target wall thickness is 2.6 mm, a USL is 3.2 mm, and an LSL is 2.0 mm (2.6 $\pm$ 0.6 mm). If the units are defective, they are replaced at a shipping-included cost of $10.00. The current process produces parts with a mean of 2.6 mm and a standard deviation of 0.2 mm. The current volume is 10,000 sections of tube per month. An improvement is being considered for the extruder heating system. This improvement will cut the variation in half, but it costs $50,000.

Find: Determine the Taguchi loss function and the payback period for the investment.

Solution: The quantities involved are: USL = 3.2 mm, LSL = 2.0 mm, $T = 2.6$ mm, $\sigma = 0.2$ mm, and $Y = 2.6$ mm. The quantity k is given by Eq. (36.2) as

$$k = \frac{\$10.00}{(3.2 - 2.6)^2} = \$27.28$$

The loss cost before the improvement is, from Eq. (36.1),

$$\text{Loss cost} = (27.78)\left[(2.6 - 2.6)^2 + 0.2^2\right] = \$1.11 \text{ per unit.}$$

After the improvement, the standard deviation is 0.1 mm; thus, the loss cost is

$$\text{Loss cost} = (27.78)\left[(2.6 - 2.6)^2 + 0.1^2\right] = \$0.28 \text{ per unit.}$$

The savings are then ($1.11-$0.28)(10,000)=$8300 per month. Thus, the payback period for the investment is $50,000/($8300/month) = 6.02 months.

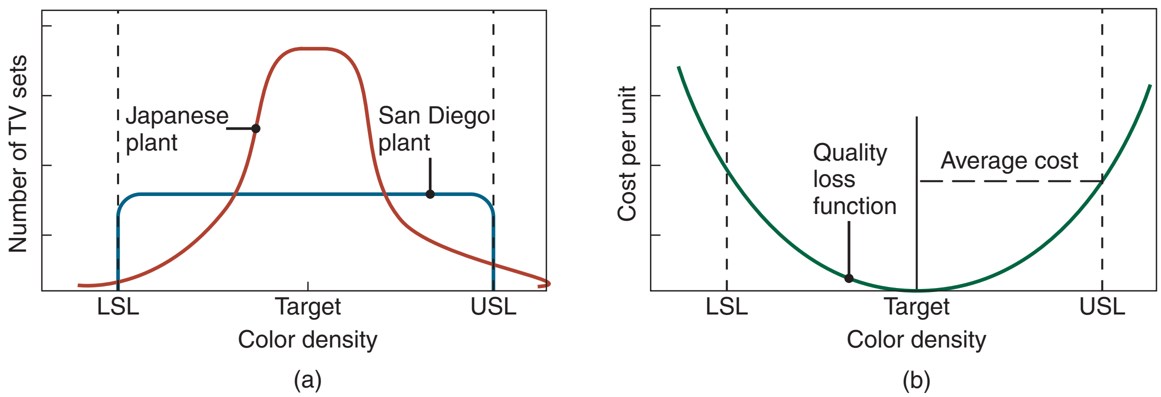


Figure 36.2: (a) Objective-function value distribution of color density for television sets. (b) Taguchi loss function, showing the average replacement cost per unit to correct quality problems.

Case Study 36.1 Manufacture of Television Sets by Sony Corporation

Sony Corporation executives found a confusing situation in the mid-1980s. Television sets manufactured in Japanese production facilities sold faster than those produced in a San Diego, CA facility, even though they were produced from identical designs. There were no identifications to distinguish the sets made in Japan from those made in the United States, and there was no apparent reason for this discrepancy. However, investigations revealed that the sets produced in Japan were superior to the U.S. versions, as color sharpness was better and hues were more brilliant. Since they were on display in stores, consumers could easily detect and purchase the model that had the best picture.

Although the difference in picture quality was obvious, the reasons for the difference were not clear. A further point of confusion was the constant assurance that the San Diego facility had a total quality program in place, and that the plant was maintaining quality-control standards so that no defective parts were produced. Although the Japanese facility did not have a total quality program, there was an emphasis on reducing variations from part to part.

Further investigations found a typical pattern in an integrated circuit that was critical in affecting color density. The distribution of parts meeting the color-design objective is shown in Fig. 36.2a; the Taguchi loss function for these parts is shown in Fig. 36.2b. In the San Diego facility, where the number of defective parts was minimized (to zero in this case), a uniform distribution within the specification limits was achieved.

The Japanese facility actually produced parts outside of the design specification, but the standard deviation about the mean was lower. Using the Taguchi loss-function approach (see Example 36.1) made it clear that the San Diego facility lost about $1.33 per unit while the Japanese facility lost $0.44 per unit.

Traditional quality viewpoints would find a uniform distribution without defects to be superior to a distribution in which a few defects are produced but the majority of parts are closer to the design target values. Consumers, however, can readily detect which product is superior, and the marketplace proves that minimizing deviations is indeed a worthwhile quality goal.

Source: After D.M. Byrne and G. Taguchi.

Example 36.2 Increasing Quality without Increasing the Cost of a Product

A manufacturer of clay tiles noticed that excessive scrap was being produced because of temperature variations in the kiln used to fire the tiles, thus adversely affecting the company's profits. The first solution the manufacturer considered was purchasing new kilns with better temperature controls; however, this solution would require a major capital investment. A study was then undertaken to determine whether modifications could be made in the composition of the clay so that it would be less sensitive to temperature fluctuations during firing.

On the basis of factorial experiment design, in which the factors involved in a process and their interactions are studied simultaneously, it was found that increasing the lime content of the clay made the tiles less sensitive to temperature variations during firing. This modification (which was also the low-cost alternative) was implemented, reducing scrap substantially and improving tile quality.

36.6 The ISO and QS Standards

Customers worldwide are increasingly demanding high-quality products and services at low prices, and are looking for suppliers that can respond to this demand consistently and reliably. This trend in the global marketplace has, in turn, created the need for international conformity and consensus regarding the establishment of methods for quality control, reliability, and safety of products. In addition to these considerations, there are equally important concerns regarding the environment and quality of life that continue to be addressed.

36.6.1 The ISO 9000 Standard

First published in 1987 and revised in 1994, the ISO 9000 standard (**Quality Management and Quality Assurance Standards**) is a deliberately generic series of quality system-management standards. This standard has permanently influenced the manner in which manufacturing companies conduct business in world trade and has become the world standard for quality.

The ISO 9000 series includes the following standards:

- **ISO 9001—***Quality systems: Model for quality assurance in design/development, production, installation, and servicing.*
- **ISO 9002—***Quality systems: Model for quality assurance in production and installation.*
- **ISO 9003—***Quality systems: Model for quality assurance in final inspection and testing.*
- **ISO 9004—***Quality management and quality system elements: Guidelines.*

Companies voluntarily register for these standards and are issued certificates. Registration may be sought generally for ISO 9001 or 9002, and some companies have registration up to ISO 9003. The 9004 standard is simply a guideline and not a model or a basis for registration. For certification, a company's plants are visited and audited by accredited and independent third-party teams to certify that the standard's 20 key elements are in place and are functioning properly.

Depending on the extent to which a company fails to meet the requirements of the standard, registration may or may not be recommended at that time. The audit team does not advise or consult with the company on how to fix discrepancies, but merely describes the nature of the noncompliance. Periodic audits are required to maintain certification. The certification process can take from six months to a year or more and can cost tens of thousands of dollars, depending on the company's size, number of plants, and product line.

The ISO 9000 standard is not a product certification, but a **quality process certification**. Companies establish their own criteria and practices for quality. However, the documented quality system must be in compliance with the ISO 9000 standard; thus, a company cannot write into the system any criterion that opposes the intent of the standard. Registration symbolizes a company's commitment to conform to consistent practices, as specified by the company's own quality system (such as quality in design, development, production, installation, and servicing), including proper documentation of such practices. In this way, customers (including government agencies) are assured that the supplier of the product or service (which may or may not be within the same country) is following specified practices. In fact, manufacturing companies are themselves assured of such practices regarding their own suppliers that have ISO 9000 registration; thus, suppliers also must be registered.

36.6.2 The QS 9000 Standard

Jointly developed by Chrysler, Ford, and General Motors, the QS 9000 standard was first published in 1994. Prior to its development, each of these automotive companies had its own standard for quality system requirements. The ISO/TS 16949 standard superseded QS 9000 in 2002, and is intended to be applied across an entire supply chain. Tier I automotive suppliers have been required to obtain third-party registration to the standard.

36.6.3 The ISO 14000 Standard

ISO 14000 is a family of standards first published in 1996 and pertaining to international **environmental management systems** (EMS). It concerns the way an organization's activities affect the environment throughout the life of its products (see also Section I.6 in the General Introduction). These activities (a) may be internal or external to the organization, (b) range from production to ultimate disposal of the product after its useful life, and (c) include effects on the environment, such as pollution, waste generation and disposal, noise, depletion of natural resources, and energy use.

The ISO 14000 family of standards has several sections: "Guidelines for Environmental Auditing," "Environmental Assessment," "Environmental Labels and Declarations," and "Environmental Management." ISO 14001, *Environmental Management System Requirements,* consists of sections titled "General Requirements," "Environmental Policy, Planning, Implementation and Operation," "Checking and Corrective Action," and "Management Review."

36.7 Statistical Methods of Quality Control

Because of the numerous variables involved in manufacturing processes and operations, the implementation of *statistical methods of quality control* is essential. Some of the more commonly observed variables in manufacturing are as follows:

- Cutting tools, dies, and molds undergo wear, thus part dimensions and surface characteristics vary over time.
- Machinery performs differently depending on its quality, age, condition, and level of maintenance; older machines tend to chatter and vibrate, can be difficult to adjust, and do not maintain tolerances.
- The effectiveness of metalworking fluids declines as they degrade; thus, tool and die life, surface finish and surface integrity of the workpiece, and forces and energy requirements are adversely affected.
- Environmental conditions, such as temperature, humidity, and air quality in the plant, may change from one hour to the next, affecting the performance of machines and workers.

- Different shipments, at different times, of raw materials to a plant may have significantly different dimensions, properties, surface characteristics, and overall quality.
- Operator attention may vary during the day or from operator to operator.

Those events that occur *randomly*—that is, without any particular trend or pattern—are called **chance variations** or **special causes**; those that can be traced to *specific causes* are called **assignable variations** or **common causes**.

Although the existence of **variability** in production operations had been recognized for centuries, it was E. Whitney (1765–1825), an American inventor and arms manufacturer, who first understood its full significance when he observed that *interchangeable parts* were indispensable to the mass production of firearms. Modern statistical concepts relevant to manufacturing engineering were first developed in the early 1900s, notably through the work of W.A. Shewhart (1891–1967), a physicist, engineer, and statistician.

36.7.1 Statistical Quality Control

Statistical quality control (SQC) involves the use of probability theory along with the testing of random subsets of parts produced to obtain an understanding of quality. To understand statistical quality control, the following commonly used terms must first be defined:

- **Sample size:** The number of parts to be inspected in a sample. The properties of the parts in the sample are studied to gain information about the whole population.
- **Random sampling:** Taking a sample from a population or lot in which each item has an equal chance of being included in the sample. Thus, when taking samples from a large bin, the inspector must not take only those that happen to be within reach.
- **Population:** The total number of individual parts of the same design from which samples are taken; also called the **universe**.
- **Lot size:** The size of a subset of the population. One or more lots can be considered subsets of the population and may also be considered as representative of the population.

The sample is inspected for several characteristics and features, such as tolerances, surface finish, and defects, using the instruments and techniques described in Chapter 35 and in Sections 36.10 and 36.11. These characteristics fall into two categories: (a) those that are measured quantitatively (*method of variables*) and (b) those that are measured qualitatively (*method of attributes*).

1. The **method of variables** is the *quantitative measurement* of the part's characteristics, such as dimensions, tolerances, surface finish, and physical or mechanical properties. The measurements are made for each of the units in the group under consideration, and the results are then compared against specifications.
2. The **method of attributes** involves observing the presence or absence of *qualitative characteristics* (such as external or internal defects in machined, formed, or welded parts, and dents in sheet-metal parts) in each of the units in the group under consideration. The sample size for attributes-type data generally is larger than that for variables-type data.

Consider measuring the diameters of machined shafts produced on a lathe (Fig. 35.2). For a variety of reasons described in this chapter, the diameters will vary. When the measured diameters of the turned shafts in a given population are listed, one or more parts will have the smallest diameter and one or more will have the largest diameter. The rest of the turned shafts will have diameters that lie between these two extremes.

All the diameter measurements can be grouped and plotted in a bar graph, called a *histogram*, representing the number of parts in each diameter group (Fig. 36.3a). The bars show a **distribution**, also called a

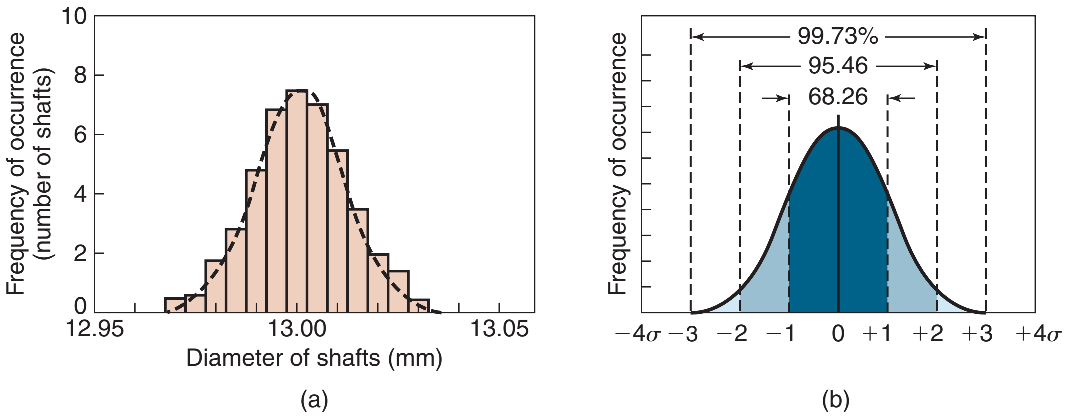


Figure 36.3: (a) A histogram of the number of shafts measured and their respective diameters. This type of curve is called a frequency distribution. (b) A normal distribution curve indicating areas within each range of standard deviation. Note that the greater the range, the higher is the percentage of parts that fall within it.

spread or **dispersion** of the diameter measurements. The *bell-shaped curve* in Fig. 36.3a is called a **frequency distribution**, and shows the frequencies with which parts of each diameter are being produced.

Data from manufacturing processes often fit curves represented by a mathematically derived **normal-distribution curve** (Fig. 36.3b), called *Gaussian*, after K.F. Gauss (1777–1855), a German mathematician and physical scientist, who developed it on the basis of *probability*. The bell-shaped **normal distribution curve** fitted to the data shown in Fig. 36.3a has two features. First, it shows that most part diameters tend to cluster around an *average value* (**arithmetic mean**), and is designated as $\bar{x}$ and calculated from the expression

$$\bar{x} = \frac{x_1 + x_2 + x_3 + \cdots + x_n}{n}, \tag{36.3}$$

where the numerator is the sum of all of the measured values (shaft diameters) and n is the number of measurements (number of shafts).

The second feature of this curve is its width, indicating the **dispersion** of the diameters measured; the wider the curve, the greater is the dispersion. The difference between the largest value and the smallest value is called the **range**, R:

$$R = x_{\max} - x_{\min}. \tag{36.4}$$

The dispersion is estimated by the **standard deviation**, given by the expression

$$\sigma = \frac{\sqrt{(x_1 - \bar{x})^2 + (x_2 - \bar{x})^2 + \cdots + (x_n - \bar{x})^2}}{n - 1}, \tag{36.5}$$

where x_i is the measured value for each part.

Note from the numerator in Eq. (36.5) that as the curve widens, the standard deviation becomes greater, and that σ has the same units as x_i. Since the number of turned parts that fall within each group is known, the percentage of the total population represented by each group can be calculated. Thus, Fig. 36.3b shows that, in the measurement of shaft diameters,

- 99.73% of the population falls within the range $\pm 3\sigma$,
- 95.46% within $\pm 2\sigma$, and
- 68.26% within $\pm 1\sigma$.

These quantities are only valid for distributions that are normal, as shown in Fig. 36.3, and are not skewed. It will be noted that only 0.27% fall outside the $\pm 3\sigma$ range, which means that there will be 2700 defective parts per one million parts produced. In modern manufacturing, that is *not* an acceptable rate, because at this level of defects no modern computer would function reliably.

36.7.2 Six Sigma

Six sigma is a set of statistical tools based on total quality management principles of *continually* measuring the quality of products and services. Although six sigma indicates 3.4 defective parts per million, it includes considerations such as understanding *process capabilities* (described Section 36.8.2), delivering defect-free products, and thus ensuring customer satisfaction. This approach consists of a clear focus on (a) defining quality problems, (b) measuring relevant quantities, and (c) analyzing, controlling, and improving processes and operations.

As stated in Section 36.7.1, three sigma would result in 0.27% defective parts, an unacceptable rate in modern manufacturing. Also, in the service industries, at this rate 270 million incorrect credit-card transactions would be recorded each year in the United States alone. It has further been estimated that companies operating at three- to four-sigma levels lose about 10 to 15% of their total revenue due to defects. Extensive efforts continue to be made to eliminate virtually all defects in products, processes, and services, resulting in savings estimated to be in billions of dollars. Because of its major impact on business, six sigma is now widely recognized as a good management philosophy.

36.8 Statistical Process Control

If the number of parts that do not meet set standards begin to increase during a production run, it is essential to determine the cause (such as incoming materials, machine controls, degradation of metalworking fluids, operator boredom, or various other factors) and take appropriate action. Although this statement at first appears to be self-evident, it was only in the early 1950s that a systematic statistical approach was developed in order to guide operators in manufacturing plants.

The statistical approach advises the operator to take certain measures and actions and tells the operator when to take them to avoid producing further defective parts. Known as *statistical process control* (SPC), this technique consists of:

- Control charts and control limits
- Capabilities of the particular manufacturing process
- Characteristics of the machinery involved.

36.8.1 $\bar{x}$ and *R* Charts (Shewhart Control Charts)

The frequency distribution curve shown in Fig. 36.3b indicates a range of shaft diameters being produced that may fall beyond the design tolerance range. The same bell-shaped curve is shown in Fig. 36.4 but now includes the *specified tolerances* for the diameter of the turned shafts.

Control charts graphically represent the variations of a process over time; they consist of data taken and plotted *during* production. Typically, there are two plots. The quantity (Fig. 36.5a) is the average for each subset of samples taken and inspected—say, each subset consists of five parts. A sample size of between 2 and 10 parts is sufficiently accurate (although more parts are better), provided that the sample size is held constant throughout the inspection.

The frequency of sampling depends on the nature of the process; some processes may require continual sampling, whereas others may require only one sample per day. Quality-control analysts are best qualified to determine this frequency for a particular operation. Since the measurements in Fig. 36.5a are made consecutively, the abscissa of the control charts also represents time.

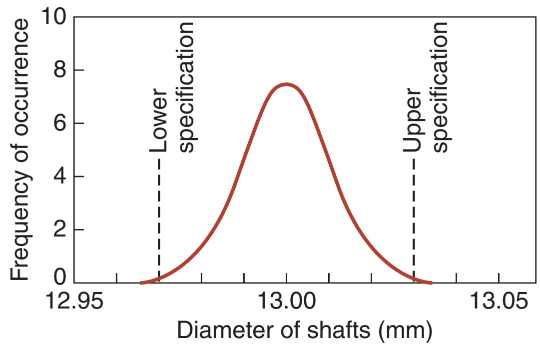


Figure 36.4: Frequency distribution curve showing lower and upper specification limits.

The solid horizontal line in this figure is the **average of averages (grand average)**, denoted as $\bar{\bar{x}}$ and represents the population mean. The upper and lower horizontal broken lines indicate the **control limits** for the process. The control limits are set on these charts according to statistical-control formulas designed to keep actual production within acceptable levels of variation. One common approach is to ensure that all parts are within three standard deviations of the mean ($\pm 3\sigma$).

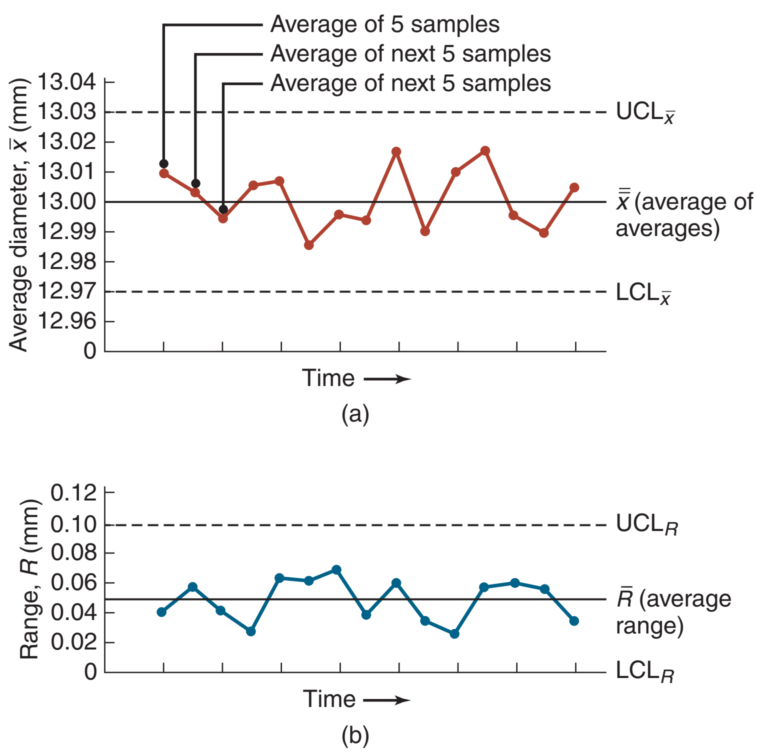


Figure 36.5: Control charts used in statistical quality control. The process shown is in good statistical control because all points fall within the lower and upper control limits. In this illustration, the sample size is 5 and the number of samples is 15.

The standard deviation also can be expressed as a function of range. Thus, for $\bar{x}$,

$$\text{Upper control limit } (\text{UCL}_{\bar{x}}) = \bar{x} + 3\sigma = \bar{\bar{x}} + A_2\bar{R} \tag{36.6}$$

and

$$\text{Lower control limit } (\text{LCL}_{\bar{x}}) = \bar{x} - 3\sigma = \bar{\bar{x}} - A_2\bar{R} \tag{36.7}$$

where A_2 is obtained from Table 36.2 and $\bar{R}$ is the average of R values. The quantities $\bar{\bar{x}}$ and $\bar{R}$ are estimated from the measurements taken.

The control limits are calculated on the basis of the past production capability of the equipment itself, and are not associated with either design tolerance specifications or dimensions. They indicate the limits within which a certain percentage of measured values normally are expected to fall, because of the inherent variations of the process itself and upon which the limits are based. The major goal of statistical process control is to improve the manufacturing process with the aid of control charts so as to eliminate assignable causes. The control chart continually indicates progress in this area.

The second control chart, shown in Fig. 36.5b, indicates the range, R, in each subset of samples. The solid horizontal line represents the average of R values in the lot, denoted as $\bar{R}$, and is a measure of the variability of the samples. The upper and **lower control limits** for R are obtained from the equations

$$\text{UCL}_R = D_4\bar{R} \tag{36.8}$$

and

$$\text{LCL}_R = D_3\bar{R}, \tag{36.9}$$

where the constants D_4 and D_3 take on the values given in Table 36.2. The table also includes the constant d_2, which is used to estimate the standard deviation of the process distribution shown in Fig. 36.4 from the equation

$$\sigma = \frac{\bar{R}}{d_2}. \tag{36.10}$$

When the curve of a control chart is like the one shown in Fig. 36.5a, it is said that the process is in *good statistical control*, meaning that

- There is no discernible trend in the pattern of the curve
- The points (measured values) are random with time
- The points do not exceed the control limits.

Table 36.2: Constants for Control Charts.

Sample size	A_2	D_4	D_3	d_2
2	1.880	3.267	0	1.128
3	1.023	2.575	0	1.693
4	0.729	2.282	0	2.059
5	0.577	2.115	0	2.326
6	0.483	2.004	0	2.534
7	0.419	1.924	0.078	2.704
8	0.373	1.864	0.136	2.847
9	0.337	1.816	0.184	2.970
10	0.308	1.777	0.223	3.078
12	0.266	1.716	0.284	3.258
15	0.223	1.652	0.348	3.472
20	0.180	1.586	0.414	3.735

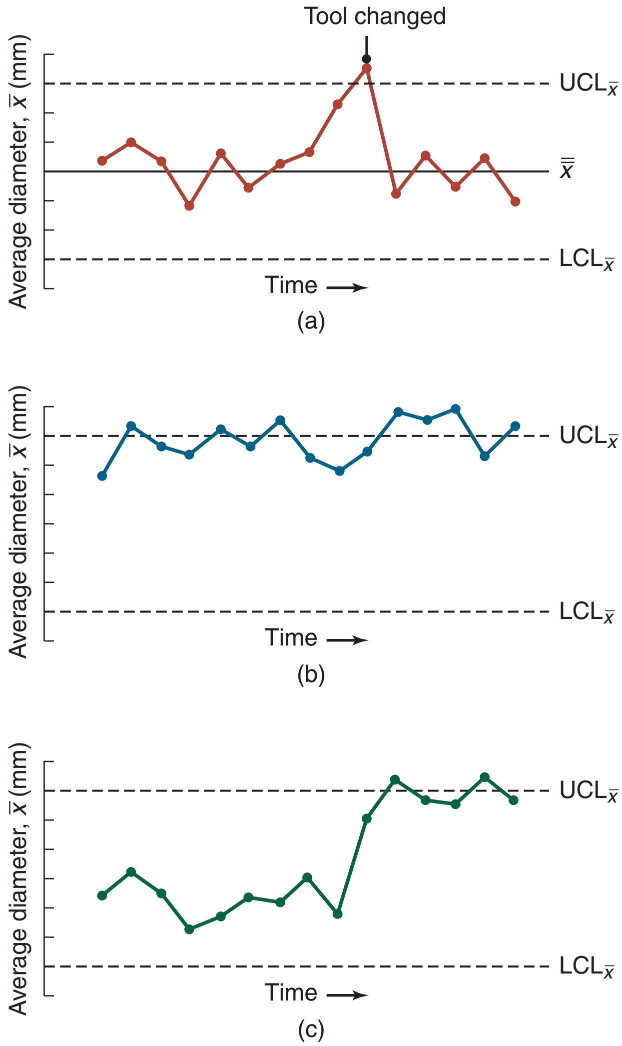


Figure 36.6: Control charts. (a) Process begins to become out of control because of such factors as tool wear (*drift*); the tool is changed and the process is then in statistical control. (b) Process parameters are not set properly; thus, all parts are around the upper control limit (*shift in mean*). (c) Process becomes out of control, because of factors such as a change in the properties of the incoming material (*shift in mean*).

It can be seen that in curves such as those in Fig. 36.6a, b, and c, there are certain *trends*. Note, for example, that in the middle of the curve in Fig. 36.6a, the diameter of the shafts is increasing with time, a reason for which may be a change in one of the process variables, such as wear of the cutting tool.

If the trend is consistently towards large diameters, as in the curve in Fig. 36.6b, with diameters hovering around the **upper control limit,** it could mean that the tool settings on the lathe may be incorrect and, as a result, the parts being turned are consistently too large. The curve in Fig. 36.6c shows two distinct trends that may be due to such factors as a change in the properties of the incoming material or a change in the performance of the cutting fluid (e.g., its degradation). These situations place the process *out of control*. Warning limits to this effect are sometimes set at $\pm 2\sigma$.

Analyzing patterns and trends in control charts requires considerable experience so that one may identify the specific cause(s) of an out-of-control situation. Among such causes may be one or more of those variables listed at the beginning of Section 36.7. Overcontrol of the manufacturing process (such as setting upper and lower control limits too close to each other, resulting in a smaller standard-deviation range) is

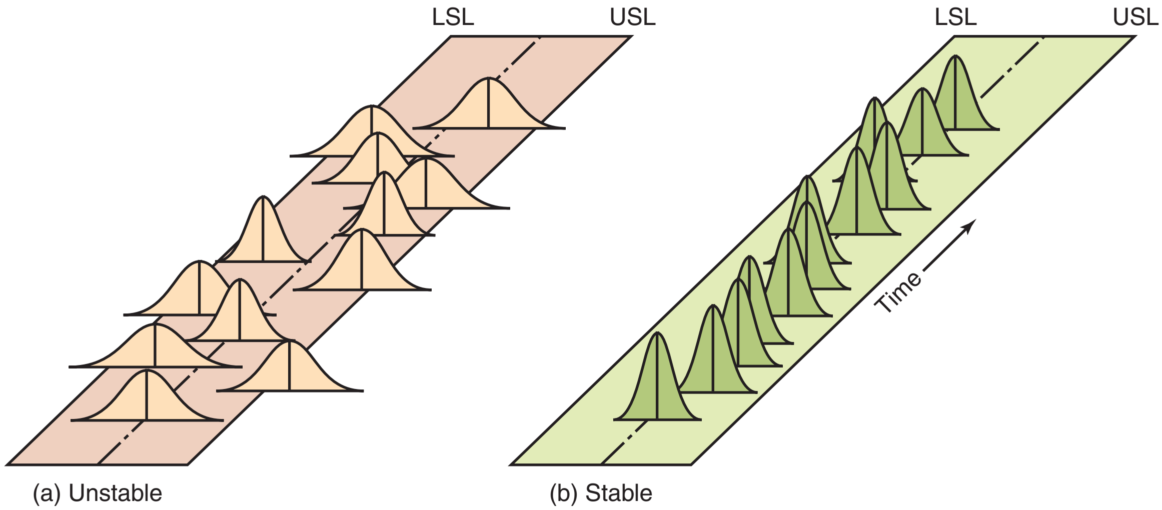


Figure 36.7: Illustration of processes that are (a) unstable or out of control and (b) stable or in control. Note in part (b) that all distributions standard deviations that are lower than those of the distributions in part (a) and have means closer to the desired value. *Source:* After K. Crow.

another cause of out-of-control situations, and is the reason why control limits are calculated on the basis of *process variability* rather than on potentially inapplicable criteria.

It is evident that operator training is critical for the successful implementation of SPC on the shop floor. Once process control guidelines are established, operators should also have some responsibility for making adjustments in processes that are beginning to become out of control. The capabilities of individual operators also should be taken into account so that they are not overloaded with data input and thus fail to interpret the data properly.

This task is now greatly simplified through dedicated software. For example, digital readouts on electronic measuring devices are now directly integrated into a computer system for real-time SPC. Figure 35.2 shows such a multifunctional computer system in which the output from a digital caliper or micrometer is analyzed by a microprocessor, in real time, and is displayed in several ways, such as frequency distribution curves and control charts.

36.8.2 Process Capability

Process capability is defined as the ability of a process to manufacture defect-free parts in controlled production. It indicates an ability to produce parts consistently and repeatedly within specific limits of precision (Fig. 36.7). Various indices are used to determine process capability, describing the relationship between the variability of a process and the spread of lower and upper specification limits. Since a manufacturing process typically involves materials, machinery, and operators, each factor can be analyzed individually to identify a problem when process capabilities do not meet **specification limits.**

Example 36.3 Calculation of Control Limits and Standard Deviation

Given: The data given in Table 36.3 show length measurements (in mm) taken on a machined workpiece. The sample size is 5, and the number of samples is 10; thus, the total number of parts measured is 50. The quantity $\bar{x}$ is the average of five measurements in each sample.

Find: Determine the upper and lower control limits and the standard deviation for the population of machined parts.

Solution: The average of averages, $\bar{\bar{x}}$, is

$$\bar{\bar{x}} = \frac{1125.4}{10} = 112.51 \text{ mm}$$

The average of the R values is

$$\bar{R} = \frac{26}{10} = 2.6 \text{ mm}$$

Since the sample size is 5, it can be determined from Table 36.2 that $A_2 = 0.577$, $D_4 = 2.115$, and $D_3 = 0$. The control limits can now be calculated from Eqs. (36.4) through (36.7). Thus, for averages,

$$\text{UCL}_{\bar{x}} = 112.51 + (0.577)(2.6) = 114.01 \text{ mm}$$

and

$$\text{LCL}_{\bar{x}} = 112.51 - (0.577)(2.6) = 111.01 \text{ mm}$$

For ranges,

$$\text{UCL}_R = (2.115)(2.6) = 5.5 \text{ mm}$$

and

$$\text{LCL}_R = (0)(2.6) = 0 \text{ mm}$$

From Eq. (36.10), the standard deviation, σ, for the population, for a value of $d_2 = 2.326$, can be estimated as

$$\sigma = \frac{2.6}{2.326} = 1.18 \text{ mm}$$

Table 36.3: Data for Example 36.3.

Sample number	x_1	x_2	x_3	x_4	x_5	$\bar{x}$	R
1	113.3	111.8	112.8	113.3	112.5	112.74	1.5
2	113.0	112.5	113.5	111.5	111.8	112.46	2.0
3	111.3	113.8	112.3	112.3	110.5	112.09	3.3
4	112.3	112.8	115.2	114.0	110.5	112.94	4.6
5	112.3	113.0	112.5	112.8	112.0	112.52	1.0
6	112.8	113.0	112.8	111.5	111.8	112.38	1.5
7	111.5	112.0	112.3	113.3	113.5	112.52	2.0
8	113.0	112.0	112.5	112.0	114.3	112.76	2.3
9	112.8	113.3	109.2	111.3	114.0	112.12	4.8
10	112.3	112.5	111.0	113.5	114.0	112.60	3.0

36.8.3 Acceptance Sampling and Control

Acceptance sampling consists of taking a few random samples from a lot and inspecting them to judge whether the entire lot is acceptable or whether it should be rejected or reworked. Developed in the 1920s and used extensively during World War II for military hardware, this statistical technique is used widely. Acceptance sampling is particularly useful for inspecting high-production-rate parts when 100% inspection would be too costly. There are certain critical devices, such as pacemakers, and prosthetic devices, for example, that must be subjected to 100% inspection.

A number of acceptance sampling plans have been prepared for both military and national standards on the basis of an acceptable, predetermined, and limiting percentage of nonconforming parts in the sample.

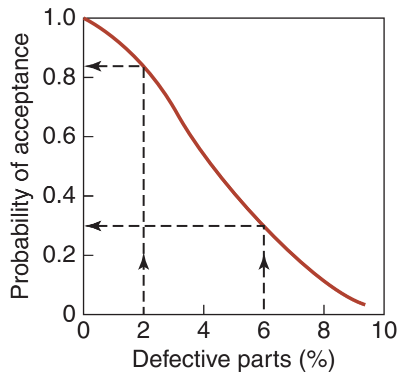


Figure 36.8: A typical operating-characteristics curve used in acceptance sampling. The higher the percentage of defective parts, the lower is the probability of acceptance by the consumer.

If this percentage is exceeded, the entire lot is rejected or it is reworked, if economically feasible. Note that the actual number of samples, but not the percentages of the lot that are in the sample, can be significant in acceptance sampling.

The greater the number of samples taken from a lot, the greater is the chance that the sample will contain nonconforming parts. However, this does not change the probability of detecting a nonconforming part. **Probability** is defined as the relative occurrence of an event. The probability of acceptance is obtained from various operating characteristics curves, one example of which is shown in Fig. 36.8.

The **acceptance quality level** (AQL) is commonly defined as the level at which there is a 95% acceptance probability for the lot. This percentage indicates to the manufacturer that 5% of the parts in the lot may be rejected by the consumer (**producer's risk**); likewise, the consumer knows that 95% of the parts are acceptable (**consumer's risk**).

The manufacturer can salvage those lots that do not meet the desired quality standards through a secondary rectifying inspection. In this method, a 100% inspection is made of the rejected lot, and the defective parts are removed. The process is time consuming and costly, and is an important incentive for the manufacturer to better control its production processes.

Acceptance sampling requires less time and fewer inspections than do other sampling methods. Consequently, inspection of the parts can be more detailed. Automated inspection techniques (Section 36.12) have been developed so that 100% inspection of all parts is indeed possible and inspection can also be economical.

36.9 Reliability of Products and Processes

All products eventually fail in some manner or other: Automobile tires become worn, electric motors burn out, water heaters begin to leak, dies and cutting tools wear out or break, and machinery stops functioning properly. **Product reliability** is generally defined as the probability that a product will perform its intended function, and without failure, in a given environment for a specified period of time while in normal use by the customer.

The more critical the application of a particular product, the higher its reliability must be. Thus, the reliability of an aircraft jet engine, a medical instrument, or an elevator cable, for example, must be much higher than that of a kitchen faucet or a mechanical pencil. From the topics described in this chapter, it can be noted that as the quality of each component of a product increases, so, too, does the reliability of the whole product.

Predicting reliability involves complex mathematical relationships and calculations. The importance of predicting the reliability of the critical components of civilian or military aircraft is obvious. The reliability of an automated and computer-controlled high-speed production line, with all of its complex mechanical and electronic components, is also important, as its failure can result in major economic losses to the manufacturer.

Process reliability can be defined as the capability of a particular manufacturing process to operate predictably and smoothly over time. It is implicit that there must be no significant deterioration in performance, which otherwise would require downtime on machines, interrupt production, and result in major economic loss.

36.10 Nondestructive Testing

Nondestructive testing (NDT) is carried out in such a manner that product integrity and surface texture remain unchanged. The techniques employed generally require considerable operator skill and interpreting test results accurately, which may be a difficult task. The extensive use of computer graphics and other enhancement techniques, however, have significantly reduced the likelihood of human error. Current systems have various capabilities for data acquisition and for qualitative and quantitative inspection and analysis.

Liquid Penetrants. In this technique, fluids are applied to the surfaces of the part and allowed to penetrate into cracks, seams, and pores (Fig. 36.9). By capillary action, the penetrant can seep into cracks as small as 0.1 μm in width. Two common types of liquids used for this test are (a) *fluorescent penetrants*, with various sensitivities and which fluoresce under ultraviolet light and (b) *visible penetrants*, using dyes (usually red) that appear as bright outlines on the workpiece surface.

The liquid penetrants method can be used to detect a variety of surface defects. The equipment is simple and easy to use, can be portable, and is less costly to operate than those of other methods. However, the method can detect only defects that are open to the surface or are external.

Magnetic-particle Inspection. This technique consists of placing fine ferromagnetic particles on the surface of the part. The particles can be applied either dry or in a liquid carrier, such as water or oil. When the part is magnetized with a magnetic field, a discontinuity (defect) on the surface causes the particles to gather visibly around the defect (Fig. 36.10). The ferromagnetic particles may be colored with pigments for better visibility on metal surfaces.

The defect then becomes a magnet, due to flux leakages where the magnetic-field lines are interrupted by the defect; this, in turn, creates a small-scale N-S pole at either side of the defect as field lines exit the surface. The particles generally take the shape and size of the defect. Subsurface defects also can be detected

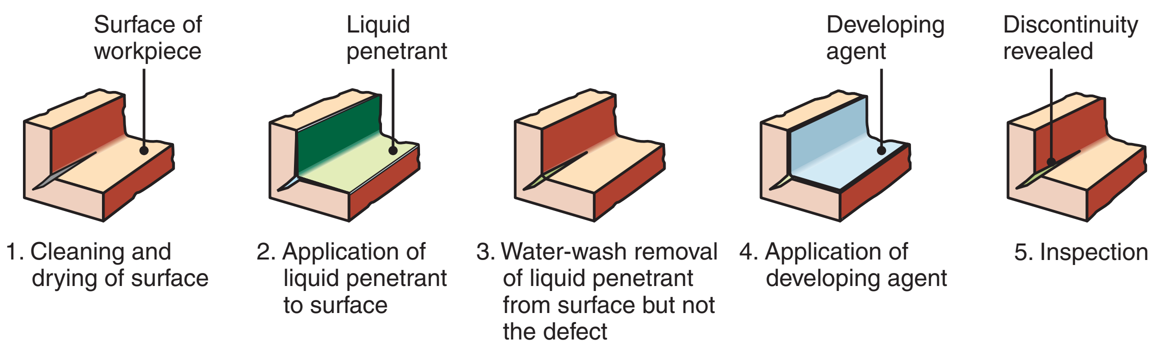


Figure 36.9: Sequence of operations for liquid-penetrant inspection to detect the presence of cracks and other flaws in a workpiece. *Source:* ASM International.

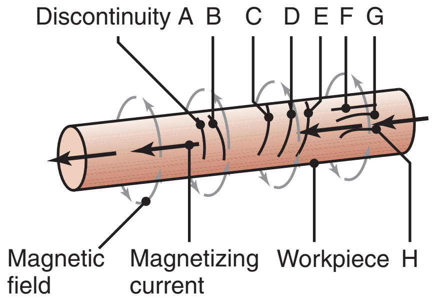


Figure 36.10: Schematic illustration of magnetic-particle inspection of a part with a defect in it. Cracks that are in a direction parallel to the magnetic field (such as discontinuity A) would not be detected, whereas the others shown would. Discontinuities F, G, and H are the easiest to detect. *Source:* ASM International.

by this method, provided that they are not too deep. The magnetic fields can be generated with either DC or AC, and with yokes, bars, and coils. The magnetic-particle method can also be used on pure ferromagnetic materials, but the parts have to be demagnetized and cleaned after inspection. The equipment may be portable or stationary.

Ultrasonic Inspection. In this technique, an ultrasonic beam travels through the part; an internal defect (such as a crack) interrupts the beam and reflects back a portion of the ultrasonic energy. The amplitude of the energy reflected and the time required for its return indicate the presence and location of any flaws in the workpiece.

The ultrasonic waves are generated by transducers, called *search units* or *probes*, available in various types and shapes. Transducers operate on the principle of *piezoelectricity* (Section 3.7), using materials such as quartz, lithium sulfate, or various ceramics. Most inspections are carried out at a frequency 1–25 MHz. Couplants, such as water, oil, glycerin, and grease, are used to transmit the ultrasonic waves from the transducer to the test piece. The ultrasonic-inspection method has high penetrating power and sensitivity. It can also be used from various directions to inspect flaws in large parts, such as railroad wheels, pressure vessels, and die blocks. The method requires experienced personnel to properly conduct the inspection and to correctly interpret the results.

Acoustic Methods. The **acoustic-emission technique** (see also Section 21.5.4) detects signals (high-frequency stress waves) generated by the workpiece itself during plastic deformation, crack initiation and propagation, phase transformation, and abrupt reorientation of grain boundaries. Bubble formation during the boiling of a liquid and friction and wear of sliding interfaces are other sources of acoustic signals.

Acoustic-emission inspection is usually performed by elastically stressing the part or structure, such as bending a beam, applying torque to a shaft, or internally pressurizing a vessel. **Sensors,** typically consisting of piezoelectric ceramic elements, detect acoustic emissions. This method is particularly effective for continuous surveillance of load-bearing structures.

The **acoustic-impact technique** consists of tapping the surface of an object, listening to the signals produced, and analyzing them to detect discontinuities and flaws. The principle is basically the same as that employed when tapping walls, desktops, or countertops in various locations, with a finger or a hammer, and listening to the sound emitted. Vitrified grinding wheels (Section 26.2) are tested in a similar manner (called *ring test*) to detect cracks in the wheel that may not be visible to the naked eye. The acoustic-impact technique is easy to perform and can be instrumented and automated.

Radiography. *Radiography* involves X-ray inspection to detect such internal flaws as cracks and porosity. The technique detects differences in density within a part. Thus, for example, on an X-ray film, the metal surrounding a defect is typically denser, and thus shows up as lighter than the flaws. The source of

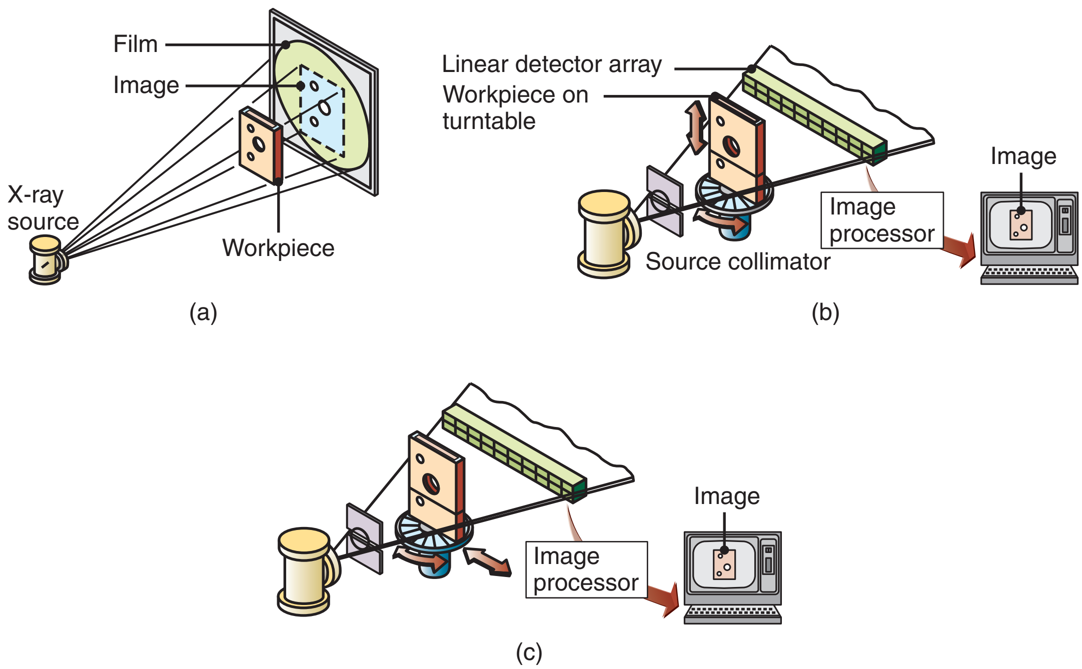


Figure 36.11: Three methods of radiographic inspection: (a) conventional radiography, (b) digital radiography, and (c) computed tomography. *Source:* ASM International.

radiation is typically an X-ray tube, and a visible, permanent image is made on a film or radiographic paper (Fig. 36.11a). **Fluoroscopes** also are used to produce X-ray images very quickly, and it is a real-time radiography technique that shows events as they are occurring. Radiography requires expensive equipment and proper interpretation of results, and can be a radiation hazard.

Three radiographic technics are:

- **Digital radiography.** The film is replaced by a linear array of detectors (Fig. 36.11b). The X-ray beam is collimated into a fan beam (compare Fig. 36.11a and b) and the workpiece is moved vertically. The detectors digitally sample the radiation, and the data are stored in computer memory; the monitor then displays the data as a two-dimensional image of the workpiece.

- **Computed tomography.** Also called *computer-assisted tomography* or **CAT scans**, this technique is based on the same system as described for digital radiography. The main differences are that the workpiece is rotated along a vertical axis as it is being moved vertically (Fig. 36.11c) and the monitor produces X-ray images of thin cross-sections of the workpiece. The translation and rotation of the workpiece provide several angles from which to precisely view the object.

Eddy-current Inspection. This method is based on the principle of *electromagnetic induction*. The part is placed in or adjacent to an electric coil through which alternating current (exciting current) flows, at frequencies of 60 Hz to 6 MHz. The current causes eddy currents to flow in the part. Defects in the part impede and change the direction of the eddy currents (Fig. 36.12) and cause changes in the electromagnetic field. These changes then affect the exciting coil (inspection coil), the voltage of which is monitored to determine the presence of flaws. Inspection coils can be made in various sizes and shapes to suit the shape of the part being inspected. Parts must be conductive electrically, and flaw depths detected usually are limited to 13 mm. The technique requires the use of a standard reference sample to set the sensitivity of the tester.

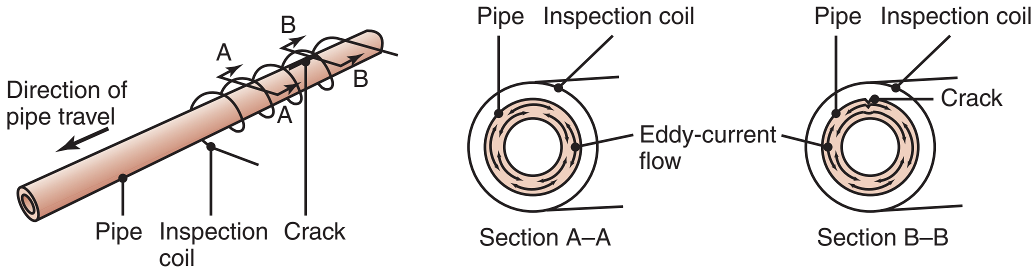


Figure 36.12: Changes in eddy-current flow caused by a defect in a part. *Source:* ASM International.

Thermal Inspection. *Thermal inspection* involves using contact- or noncontact-type heat-sensing devices that detect temperature changes. Defects in the workpiece (such as cracks, and poor joints, debonded regions in laminated structures) cause a change in temperature distribution. In **thermographic inspection**, materials such as heat-sensitive paints and papers, liquid crystals, and other coatings are applied to the workpiece surface; any changes in their color or appearance indicate defects. The most common method of noncontact thermographic inspection uses infrared detectors (usually infrared scanning microscopes and cameras), which have a high response time and sensitivities of as low as 1°C. **Thermometric inspection** utilizes devices such as thermocouples, radiometers, and pyrometers, as well as materials, such as waxlike crayons.

Holography. The *holography technique* creates a three-dimensional image of the part by utilizing an optical system (Fig. 36.13). Generally used on simple shapes and highly polished surfaces, this technique records the image on a photographic film.

The use of holography has been extended to **holographic interferometry** for the inspection of parts with various shapes and surface features. Using double- and multiple-exposure techniques, while the part is being subjected to external forces or time-dependent variations, any changes in the images reveal defects in the part.

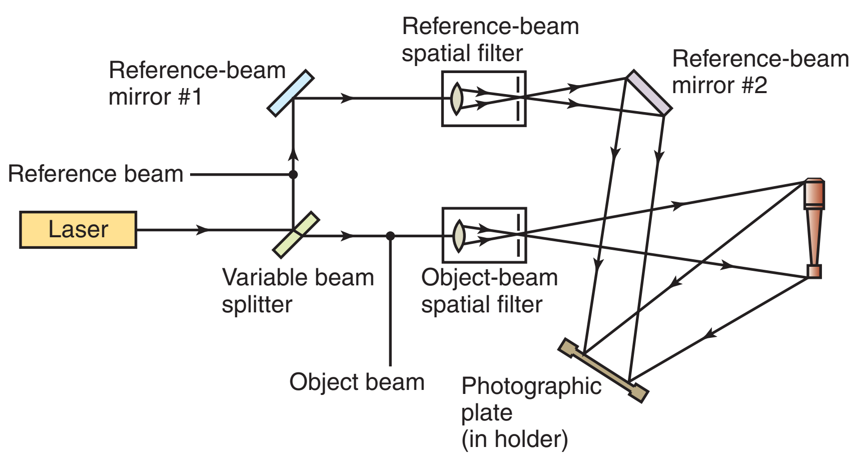


Figure 36.13: Schematic illustration of the basic optical system used in holography elements in radiography for detecting flaws. *Source:* ASM International.

- In **acoustic holography**, information on internal defects is obtained directly from the image of the interior of the part. In *liquid-surface acoustical holography*, the part and two ultrasonic transducers (one for the object beam and the other for the reference beam) are immersed in a water-filled tank. A holographic image is then obtained from the ripples in the tank.
- In **scanning acoustical holography**, only one transducer is used and the hologram is produced by electronic-phase detection. In addition to being more sensitive, the equipment is usually portable and can accommodate very large workpieces by using a water column instead of a tank.

36.11 Destructive Testing

As the name suggests, the part tested by *destructive-testing* methods no longer retains its integrity, original shape, or surface characteristics. Mechanical test methods, described in Chapter 2, are all destructive, in that a sample or specimen has to be removed from the product in order to test it. Examples of other destructive tests include the speed testing of grinding wheels to determine their bursting speed (Section 26.4) and the high-pressure testing of pressure vessels to determine their bursting pressure.

Hardness tests that leave relatively large indentations (Figs. 2.13 and 2.14) also may be regarded as destructive testing. Microhardness tests, however, may be regarded as nondestructive because of the very small permanent indentations produced; this distinction is based on the assumption that the material is not notch sensitive (see Section 2.9), as typically is the case with brittle materials. Generally, most glasses, highly heat treated metals, and ceramics are notch sensitive; consequently, a small indentation produced by the indenter can greatly reduce their strength and toughness.

36.12 Automated Inspection

Traditionally, individual parts and sub-assemblies have been manufactured in batches, sent to inspection in quality-control rooms (**postprocess inspection**) and, if approved, placed into inventory. If the parts do not pass the quality inspection, they are either scrapped or are kept and used on the basis of having a certain acceptable deviation from the standard. Parts also may be inspected immediately after they are produced (**in-process inspection**).

In contrast, *automated inspection* uses a variety of sensor systems that monitor the relevant parameters during the manufacturing operation (**online** or **in-situ inspection**). Using the measurements obtained, the process automatically corrects itself to produce acceptable parts. Thus, further inspection of the part at another location in the plant becomes unnecessary. The development of accurate sensors and advanced computer-control systems has greatly enabled automated inspection to be integrated into manufacturing operations (Chapters 37 and 38). Such integration ensures that no part is moved from one process to another (such as a turning operation followed by cylindrical grinding), unless the part is made correctly and meets the standards in the first operation.

Automated inspection is flexible and responsive to product design changes. Furthermore, because of automated equipment, less operator skill is required, productivity is increased, and parts have higher quality, dimensional accuracy, and reliability.

Sensors for Automated Inspection. Continuing advances in *sensor technology*, described in Section 37.7, have made real-time monitoring of manufacturing processes feasible. Directly or indirectly, and with the use of various *probes*, sensors can detect dimensions, surface finish, temperature, force, power, vibration, tool wear, and the presence of defects.

Sensors operate on the principles of strain gages, inductance, capacitance, ultrasonics, acoustics, pneumatics, infrared radiation, optics, lasers, or various electronic gages. They may be *tactile* (touching) or *nontactile*. They are linked to microprocessors and computers for graphic data display (see also *programmable logic controllers*, Section 37.2.6). This capability allows rapid online adjustment of any processing parameter, thus resulting in the production of parts that consistently are within specified standards. For example, such systems already are standard equipment on machine tools, described in Part IV of this book.

Summary

- Quality must be built into products. Quality assurance concerns various aspects of production, such as design, manufacturing, assembly, and especially inspection, at each step of production for conformance to specifications.
- The traditional approach of inspecting a part after it is made has largely been replaced by online and 100% inspection of all parts being manufactured.
- Statistical quality control and process control are indispensable in modern manufacturing; they are particularly important in the production of interchangeable parts and in reduction of manufacturing costs.
- Although all quality-control approaches have their limits of applicability, the implementation of total quality management, the ISO and QSO 9000 standards, and the ISO 14000 standard are among the most significant developments in quality management in manufacturing.
- Several nondestructive and destructive testing techniques, each of which has its own applications, advantages, and limitations, are available for inspection of parts produced.

Key Terms

Acceptance quality level
Acceptance sampling
Assignable variations
Automated inspection
Chance variations
Common cause
Consumer's risk
Continuous improvement
Control charts
Control limits
Defect prevention
Deming methods
Destructive testing
Dispersion
Distribution
Environmental management systems
Experimental design
Factorial design
Frequency distribution
Grand average
ISO standards
Juran methods
Kaizen
Lot size
Lower control limit
Method of attributes
Method of variables
Nondestructive testing
Normal distribution curve
Population
Probability
Process capability
Process reliability
Product reliability
Producer's risk
Quality
Quality assurance
Quality circle
QS standards
Random sampling
Range
Reliability
Return on quality
Robustness

Sample size
Sensors
Shewhart control charts
Six sigma
Special cause
Specification limits
Standard deviation
Statistical process control
Statistical quality control
Statistics
Taguchi loss function
Taguchi methods
Total quality control
Total quality management
Upper control limit
Variability

Bibliography

Allen, T.T., **Introduction to Engineering Statistics and Lean Sigma: Statistical Quality Control and Design of Experiments and Systems**, 2nd ed,. Springer, 2010.
Besterfield, D.H., **Quality Improvement**, 9th ed., Prentice Hall, 2012.
De Feo, J., and Juran, J.M., **Juran's Quality Handbook**, 7th ed., McGraw-Hill, 2016.
Fowlkes, W.Y., **Engineering Methods for Robust Product Design: Using Taguchi Methods in Technology and Product Development**, Pearson/Education, 2012.
Heiller, C., **Handbook of Nondestructive Evaluation**, 2nd ed., McGraw-Hill, 2013.
Hull, B., **Manufacturing Best Practices: Optimizing Productivity and Product Quality,** Wiley, 2011.
Imai, M., **Gemba Kaizen: A Commonsense Approach to Continuous Improvement**, 2nd ed., McGraw-Hill, 2012.
Montgomery, D.C., **Introduction to Statistical Quality Control**, 7th ed., Wiley, 2012.
Pyzdek, T., and Keller, P., **The Six Sigma Handbook,** 4th ed., McGraw-Hill, 2014.
Roy, R.K., **A Primer on the Taguchi Method**, 2nd ed., SME, 2010.
Ryan, T.P., **Modern Experimental Design**, Wiley, 2007.
Ryan, T.P., **Statistical Methods for Quality Engineering**, 3rd ed., Wiley, 2011.
Schilling, E.G., and Neubauer, D.V., **Acceptance Sampling in Quality Control**, 2nd ed., Chapman and Hall/CRC, 2009.
Taguchi, G., Chowdhury, S., and Wu, Y., **Taguchi's Quality Engineering Handbook**, Wiley-Interscience, 2004.
Wang, J.X., **Engineering Robust Designs with Six Sigma**, Prentice Hall, 2005.
Yang, G., **Life Cycle Reliability Engineering**, Wiley, 2007.

Review Questions

36.1. Define the terms *sample size*, *random sampling*, *population*, and *lot size*.

36.2. What are chance variations?

36.3. Explain the difference between method of variables and method of attributes.

36.4. Define standard deviation. Why is it important in manufacturing?

36.5. Describe what is meant by statistical process control.

36.6. When is a process out of control? Explain.

36.7. Explain why control charts are developed. How are they used?

36.8. What is a loss function? How is it used?

36.9. What do control limits indicate?

36.10. Define process capability. How is it used?

36.11. What is acceptance sampling? Why was it developed?

36.12. Explain the difference between series and parallel reliability?

36.13. What is meant by six sigma quality?

36.14. Explain the difference between (a) probability and reliability and (b) robustness and reliability?

36.15. How are liquid penetrants used, and what can they detect?

36.16. Give three methods of non-destructive testing that measure material properties.

Qualitative Problems

36.17. Explain why major efforts are continually being made to build quality into products.

36.18. Give examples of products for which 100% sampling is not possible or feasible.

36.19. What is the consequence of setting lower and upper specifications closer to the peak of the curve in Fig. 36.4?

36.20. Identify several factors that can cause a process to become out of control.

36.21. Describe situations in which the need for destructive testing techniques is unavoidable.

36.22. Which of the nondestructive inspection techniques are suitable for nonmetallic materials? Why?

36.23. What are the advantages of automated inspection? Why has it become an important part of manufacturing engineering?

36.24. Why is reliability important in manufacturing engineering? Give several examples.

36.25. Give examples of the acoustic-impact inspection technique other than those given in the chapter.

36.26. Explain why *GO* and *NOT GO* gages (see Section 35.4.4) are incompatible with the Taguchi philosophy.

36.27. Review Case Study 36.1 and list the important lessons that are illustrated.

36.28. List advantages and disadvantages of incorporating fluoroscopy as an in-line inspection tool.

36.29. Search the technical literature and give examples of robust design in addition to that shown in Fig. 36.1.

36.30. What is a Taguchi loss function? What is its significance?

Quantitative Problems

36.31. Beverage-can manufacturers try to achieve failure rates of less than one can in ten thousand. If this corresponds to n-sigma quality, find n.

36.32. Assume that in Example 36.3 the number of samples was 8 instead of 10. Using the top half of the data in Table 36.3, recalculate the control limits and the standard deviation. Compare your observations with the results obtained by using 10 samples.

36.33. Calculate the control limits for averages and ranges for the following: number of samples = 9; $\bar{\bar{x}} = 55$; $\bar{R} = 6$.

36.34. Calculate the control limits for the following: number of samples = 5; $\bar{\bar{x}} = 42$; $\mathrm{UCL}_R = 3.500$.

36.35. In an inspection with a sample size of 8, it was found that the average range was 12 and the average of averages was 64. Calculate the control limits for averages and ranges.

36.36. Determine the control limits for the data shown in the following table:

x_1	x_2	x_3	x_4
0.57	0.61	0.50	0.55
0.59	0.55	0.60	0.58
0.55	0.50	0.55	0.51
0.54	0.57	0.50	0.50
0.58	0.58	0.60	0.56
0.60	0.61	0.55	0.61
0.58	0.55	0.61	0.53

36.37. The average of averages of a number of samples of size 7 was determined to be 125. The average range was 17.82 and the standard deviation was 5.85. The following measurements were taken in a sample: 120, 143, 124, 130, 105, 132, and 121. Is the process in control?

36.38. A manufacturer is ring rolling ball-bearing races (see Fig. 13.16a). The inner surface has a surface roughness specification of 0.10 ± 0.006 μm. Measurements taken from rolled rings indicate a mean roughness of 0.112 μm with a standard deviation of 0.02 μm. Fifty thousand rings per month are manufactured, and the cost of rejecting a defective ring is $10.00. It is known that by changing lubricants to a special emulsion, the mean roughness could be made essentially equal to the design specification. What additional cost per month can be justified for the lubricant?

36.39. For the data in Problem 36.38, assume that the lubricant change can cause the manufacturing process to achieve a roughness of 0.10 ± 0.01 μm. What additional cost per month for the lubricant can be justified? If the lubricant did not add any new cost, would you recommend its use?

Synthesis, Design, and Projects

36.40. Which aspects of the quality-control concepts of Deming, Taguchi, and Juran would, in your opinion, be difficult to implement in a typical manufacturing facility? Why?

36.41. Describe your thought on whether products should be designed and built for a certain expected life. Would your answer depend on whether the products were consumer or industrial products? Explain.

36.42. Survey the available technical literature, contact various associations, and prepare a comprehensive table concerning the life expectancy of various consumer products.

36.43. Deming's 14 points have been formulated as the characteristics of a company that produces high-quality products, so that very few defects are seen in products. Could the same rules be used to minimize a company's carbon footprint? Explain.

36.44. Would it be desirable to incorporate nondestructive inspection techniques into metalworking machinery? Give a specific example, make a sketch of such a machine, and explain its features.

36.45. Name several material and process variables that can influence product quality in metal (a) casting, (b) forming, and (c) machining.

36.46. Identify the nondestructive techniques that are capable of detecting internal flaws and those which detect external flaws only.

36.47. Explain the difference between in-process and postprocess inspection of manufactured parts. What trends are there in such inspections? Explain.

36.48. Review Table 36.1 and outline changes that would occur if your instructor incorporated all of Deming's 14 points.

36.49. Many components of products have a minimal effect on part robustness and quality. For example, the hinges in the glove compartment of an automobile do not have an impact on the owner's satisfaction, and the glove compartment is opened so infrequently that a robust design is easy to achieve. Would you advocate using Taguchi methods (such as loss functions) on this type of component? Explain.

36.50. You are instructed to design an automated inspection system for an orthogonal cutting operation. What property or properties would you attempt to measure, and with what instruments/probes?

PART IX
Manufacturing in a Competitive Environment

In a highly competitive global marketplace for consumer and industrial goods, advances in manufacturing processes, machinery, tooling, and operations are being driven by goals that may be summarized as follows:

- Products must fully meet **design** and **service requirements, specifications,** and **standards**
- Manufacturing activities must continually strive for higher levels of **quality** and **productivity**; quality must be built into the product at each stage of design and manufacture
- Manufacturing processes and operations must have sufficient **flexibility** to respond rapidly to constantly changing market demands
- The most **economical methods** of manufacturing must be explored and implemented.

Although numerical control of machine tools, beginning in the early 1950s, was a key factor in setting the stage for modern manufacturing, much of the progress in manufacturing activities stems from our ability to view these activities and operations as a large *system*, with often complex interactions among all of its components. In implementing a *systems approach* to manufacturing, various functions and activities, that for a long time had been separate and distinct entities, can now be *integrated* and *optimized*.

As the first of the four chapters in the final part of this book, Chapter 37 introduces the concept of *automation* and its implementation in terms of key developments in numerical control and, later, in computer numerical control. This introduction is followed by a description of the advances made in automation and controls, involving major topics such as adaptive control, industrial robots, sensor technology, material handling and movement, and assembly systems, and how they are implemented in modern production.

Manufacturing systems and how their individual components and operations are integrated are described in Chapter 38, along with the critical role of computers, communications, and various *enabling technologies* as an aid to such activities as product design, engineering, manufacturing, and process planning. The enabling technologies include industrial robots, sensor technology, adaptive control, flexible fixturing, and assembly systems.

Computer-integrated manufacturing, with its various features, such as cellular manufacturing, flexible manufacturing systems, just-in-time production, lean manufacturing, and artificial intelligence, are then described in Chapter 39.

The aim of Chapter 40 is to highlight the importance of the numerous, and often complex, factors and their interactions that have a major impact on competitive manufacturing in a global marketplace. Among the factors involved are product design, quality, and product life cycle; selection of materials and processes and their substitution in production; process capabilities, and costs involved, including machinery, tooling, and labor.

Chapter 37

Automation of Manufacturing Processes and Operations

- This chapter describes automation in all aspects of manufacturing processes and operations, by which parts are produced reliably, accurately, and economically at high production rates. It begins with a description of the types of automation and their various applications.
- Flexibility in manufacturing through numerical control of machines is then discussed, with detailed descriptions of their important features.

- The chapter also investigates the different control strategies that can be used, including open-loop, closed-loop, and adaptive control.
- Industrial robots are then reviewed, including their capabilities and guidelines for applications. A discussion of sensor technology and its important applications follows.
- The chapter concludes with a comprehensive description of flexible fixturing and assembly systems and their design considerations.

37.1 Introduction

Until the early 1950s, most operations in a typical manufacturing plant were carried out on traditional machinery, such as lathes, milling machines, drill presses, and various equipment for forming, shaping, and joining materials. Such equipment generally lacked flexibility, and it required considerable skilled labor to produce parts with acceptable dimensional accuracy and surface characteristics. Moreover, each time a different product was to be manufactured, the machinery had to be retooled, fixtures had to be prepared or modified, and the movement of materials among various machines had to be rearranged. The development of new products and of parts with complex shapes required numerous trial-and-error attempts by the operator to set the proper processing parameters. Also, because of human involvement, making parts that were exactly alike was often difficult, time consuming, and costly.

These circumstances meant that processing methods generally were inefficient and that labor costs were a significant portion of the overall production cost. The necessity for reducing the labor share of product cost became increasingly apparent, as did the need to improve the efficiency and flexibility of manufacturing operations.

Productivity also became a major concern; generally defined as output per employee per hour, it basically measures operating efficiency. An efficient operation makes optimum use of all resources, such as materials, energy, capital, labor, machinery, and available technologies. With rapid advances in the science and technology of manufacturing, the efficiency of manufacturing operations began to improve and the percentage of total cost represented by labor began to decline.

In the past decade, *free trade zones* have proliferated across the entire globe. One effect of this development is that labor-intensive manufacturing, such as furniture, shoes, assembly, and clothing, often has been relocated to regions where labor costs are low (see Table I.7 in the General Introduction). Still, countries with higher labor costs continue to try to remain competitive, mainly by achieving significant increases in productivity.

The important elements in improving productivity have been *mechanization, automation, sensing*, and *control* of manufacturing equipment and systems, as well as widespread adoption of communications and software. **Mechanization** controls a machine or process with the use of various mechanical, hydraulic, pneumatic, or electrical devices; it reached its peak by the 1940s. In spite of the obvious benefits of mechanized operations, however, the worker would still be directly involved in a particular operation and would continually check a machine's performance.

Consider, for example, the following situations:

- A cutting tool wears or fractures during a machining operation
- A part is overheated during its heat treatment
- The surface finish of a part begins to deteriorate during grinding
- Dimensional tolerances and springback become too large in sheet-metal forming.

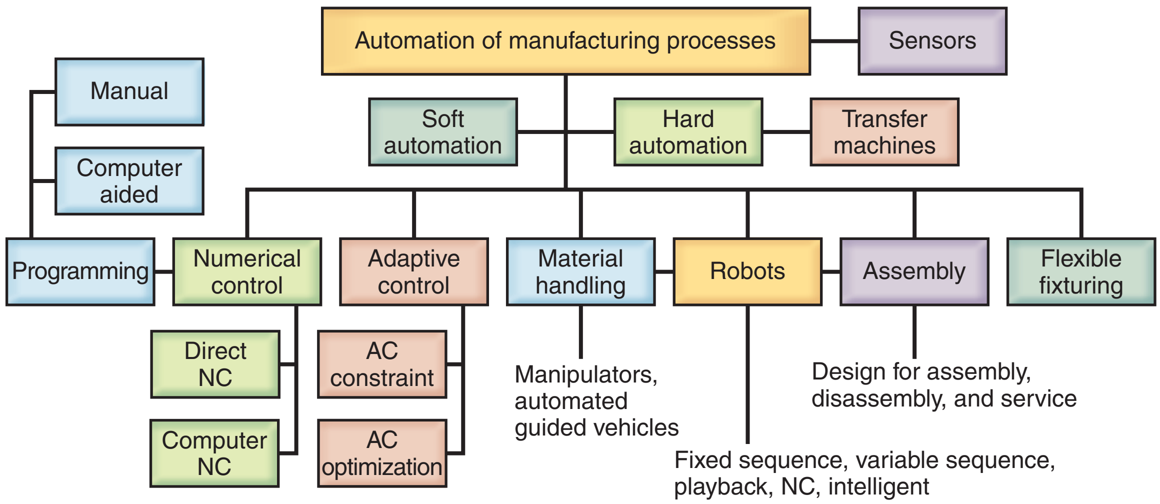


Figure 37.1: Outline of topics covered in Chapter 37.

In each of these situations, the operator had to intervene and change one or more of the relevant process parameters and machine settings, a task that requires considerable training and experience.

The next step in improving the efficiency of manufacturing operations was **automation**, a word coined in the mid-1940s by the U.S. automobile industry to indicate the **automatic handling and processing of parts** in and among production machines. Efficiency became further increased by rapid advances in automation and the development of several enabling technologies, largely through advances in **control systems**, with the help of increasingly powerful computers and software.

This chapter follows the outline shown in Fig. 37.1. It first reviews the history and principles of automation and how it has helped to *integrate* various key operations and activities in a manufacturing plant. It then introduces the concept of the control of machines and systems through **numerical control** and **adaptive control** techniques. The chapter also describes how the important activity of material handling and movement has been developed into various systems, particularly those including the use of **industrial robots** to greatly improve handling efficiency.

The subject of **sensor technology** is then described; this is a topic that is an essential element in the control and optimization of machinery, processes, and systems. Significant developments in **flexible fixturing** and **assembly operations** also are covered; these methods are essential in advanced manufacturing technologies, particularly flexible manufacturing systems. Also included is a discussion of **design for assembly, disassembly, and service,** with specific recommendations to help improve the efficiency of each of these operations. The final topic of the chapter describes the **economics** of the equipment, processes, and manufacturing operations.

37.2 Automation

Automation, from the Greek word *automatos*, meaning "self acting," is generally defined as the process of enabling machines to follow a predetermined sequence of operations with little or no human intervention, and using specialized equipment and devices that perform and control processes and operations. Table 37.1 shows the development of automation throughout history. Full automation is achieved through various devices, sensors, actuators, techniques, and equipment that are capable of (a) monitoring all aspects of, (b) making decisions concerning changes that should be made in, and (c) controlling all aspects of the operation. Modern systems also have the capability to store data regarding these activities and communicate the data over a network to computers or storage devices for analysis.

Table 37.1: History of the Automation of Manufacturing Processes.

Date	Development
1500–1600	Water power for metalworking; rolling mills for coinage strips
1600–1700	Hand lathe for wood; mechanical calculator
1700–1800	Boring, turning, and screw-cutting lathe; drill press
1800–1900	Copying lathe, turret lathe, universal milling machine; advanced mechanical calculators
1808	Sheet-metal cards with punched holes for automatic control of weaving patterns in looms
1863	Automatic piano player (Pianola)
1900–1920	Geared lathe; automatic screw machine; automatic bottle-making machine
1920	First use of the word robot
1920–1940	Transfer machines; mass production
1940	First electronic computing machine
1943	First digital electronic computer
1945	First use of the word automation
1947	Invention of the transistor
1952	First prototype numerical-control machine tool
1954	Development of the symbolic language APT (Automatically Programmed Tool); adaptive control
1957	Commercially available NC machine tools
1959	Integrated circuits; first use of the term group technology
1960s	Industrial robots
1965	Large-scale integrated circuits
1968	Programmable logic controllers
1970	First integrated manufacturing system; spot welding of automobile bodies with robots
1970s	Microprocessors; minicomputer-controlled robot; flexible manufacturing systems; group technology
1980s	Artificial intelligence; intelligent robots; smart sensors; untended manufacturing cells
1990s–2000s	Integrated manufacturing systems; intelligent and sensor-based machines; telecommunications and global manufacturing networks; fuzzy-logic devices; artificial neural networks; Internet tools; virtual environments; high-speed information systems
2010	Cloud-based storage and computing; MTConnect for information retrieval; three-dimensional geometry files; STEP-NC and autogenerated G-code.

Automation is an *evolutionary* rather than a revolutionary concept, and has been implemented especially in the following basic areas of manufacturing activities:

- **Production processes:** Machining (including hybrid machines), forging, cold extrusion, casting, powder metallurgy, and grinding operations
- **Material handling and movement:** Materials and parts in various stages of completion (called *work in progress* or WIP) are moved throughout a plant by computer-controlled equipment, with little or no human guidance
- **Inspection:** Parts are inspected automatically for dimensional accuracy, surface finish, quality, and various specific characteristics during their production (*in-process inspection*)
- **Assembly:** Individually manufactured parts and components are assembled automatically into subassemblies and then into assemblies to complete a product
- **Packaging:** Products are packaged automatically for shipment.

37.2.1 Evolution of Automation

As shown in Table I.2, some metalworking processes were used as early as 4000 B.C. However, it was not until the beginning of the Industrial Revolution in the 1750s (also referred to as the *First Industrial Revolution*) that automation began to be introduced in the production of goods. The *Second Industrial Revolution* began in the mid 1950s, with advances in several areas.

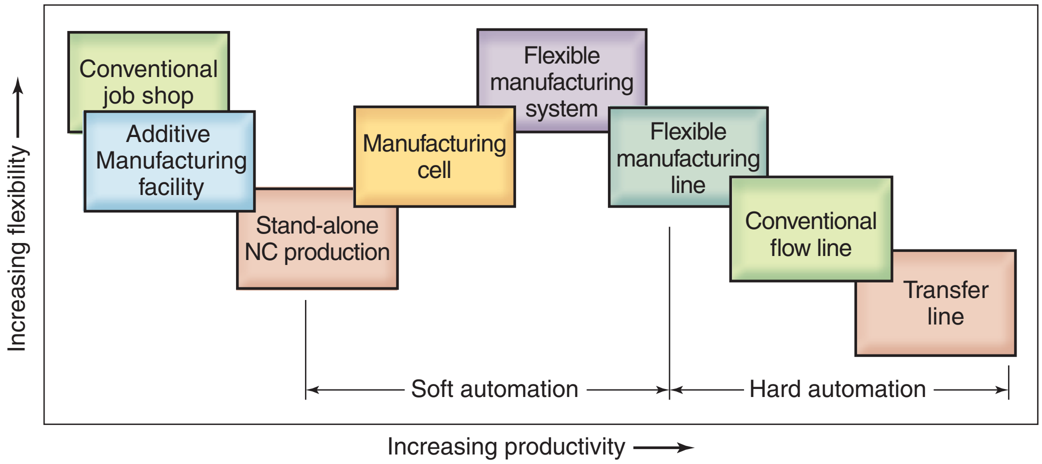


Figure 37.2: Flexibility and productivity of various manufacturing systems. Note the overlap between the systems, due to the various levels of automation and computer control that are possible in each group. (See Chapter 39 for details.)

Machine tools, such as turret lathes, automatic screw machines, and automatic glass bottle–making equipment, began to be developed in the late 1890s. Mass-production techniques and transfer machines were developed in the 1920s. These machines had *fixed* automatic mechanisms and were designed to produce *specific* parts, best represented by the automobile industry, which produced passenger cars at high rates and low cost. For example, by the end of the 1920s, Ford Motor Co. was producing more than one car per minute and it could be easily afforded by the typical factory employee.

The major breakthrough in automation began with numerical control (NC) of machine tools in the early 1950s (see Section 37.3). Since this historic development, rapid progress has been made in automating almost all aspects of manufacturing, from the introduction of computers into automation, to computerized numerical control (CNC) and adaptive control (AC), to industrial robots, to computer-aided design, engineering, and manufacturing (CAD/CAE/CAM) and computer-integrated manufacturing (CIM) systems.

As described throughout various chapters, manufacturing involves different levels of automation, depending on the processes used, the products to be made, and production volumes required. Manufacturing systems include the following classifications, in order of increasing automation (see also Fig. 37.2):

- **Job shops:** These facilities use general-purpose machines and machining centers, with high levels of labor involvement.
- **Stand-alone NC production:** This method uses numerically controlled machines, but with significant operator–machine interaction.
- **Manufacturing cells:** These cells use a cluster of machines with integrated computer control and flexible material handling, often with industrial robots.
- **Flexible manufacturing systems:** These systems use computer control of all aspects of manufacturing, the simultaneous incorporation of several manufacturing cells and automated material-handling systems.
- **Flexible manufacturing lines:** These lines involve computer-controlled machinery organized in production lines instead of cells. Part transfer is through hard automation and product flow is more limited than in flexible manufacturing systems, but the throughput is larger for higher production quantities.

- **Flow lines** and **transfer lines:** These lines consist of organized groupings of machinery with automated material handling between machines. Because the goal is to produce only one type of par, the manufacturing line is designed with limited or no flexibility.

37.2.2 Implementation of Automation

Automation generally has the following primary goals:

- *Integrate* various aspects of manufacturing operations so as to improve product quality and uniformity, minimize cycle times and effort, and reduce labor costs.
- *Improve productivity* by reducing manufacturing costs through better control of production; parts are loaded, fed, and unloaded on machines more efficiently, machines are used more effectively, and production is organized more efficiently.
- *Improve quality* by using more repeatable processes.
- *Reduce human involvement*, boredom, and thus the possibility of human error.
- *Reduce workpiece damage* caused by the manual handling of parts.
- *Raise the level of safety* for personnel, especially under hazardous working conditions.
- *Economize on floor space* in the plant by arranging machines, material handling and movement, and auxiliary equipment more efficiently.

Automation and Production Quantity. The production quantity is crucial in determining the type of machinery and the level of automation required to produce parts economically. *Total production quantity* is defined as the total number of parts to be made, whereas *production rate* is defined as the number of parts produced per unit time. The production quantity is produced in batches of various *lot sizes*. The approximate and generally accepted ranges of production volume are shown in Table 37.2 for some typical applications. Note that, as expected, experimental or prototype products represent the lowest volume (see Chapter 20).

Job shops typically produce small quantities per year (Fig. 37.2), using various standard general-purpose machine tools (called *stand-alone machines*) or *machining centers* (Chapter 25). The operations performed typically have high part variety, meaning that different parts can be produced in a short time without extensive changes in tooling or in operations. Machinery in job shops generally requires skilled labor to operate, and production quantities and rates are typically low; as a result, production cost per part is high (Fig. 37.3). When parts involve a large labor component, the production is called *labor intensive*.

Digital manufacturing and *additive manufacturing* have significantly transformed low-volume production. Along with the development of computer software with three-dimensional geometric modeling ability, piece parts can now be designed and manufactured with less effort and cost than previously possible. Operator/programmer skill is still quite high, but additive manufacturing facilities can achieve almost the same

Table 37.2: Approximate Annual Production Quantities.

Type of production	Number produced	Typical products
Experimental or prototype	1–10	All products
Piece or small-batch	10–5000	Aircraft, industrial robots,special machinery, dies, jewelry, and orthopedic implants
Batch or high-volume	5000–100,000	Trucks, agricultural machinery, jet engines, diesel engines, and sporting goods
Mass production	100,000 and over	Automobiles, appliances, fasteners, and food and beverage containers

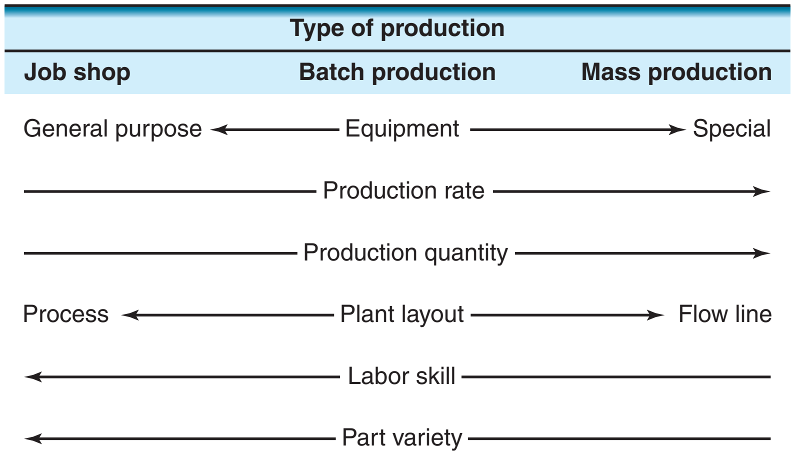


Figure 37.3: General characteristics of three types of production methods: job shop, batch, and mass production.

flexibility as job shops, being limited only in the number of materials that can be effectively processed. Modern computer software allows machining centers to produce complex parts quickly and efficiently.

- **Piece-part production** generally involves very small quantities and is suitable for job shops; the majority of piece-part production is done in lot sizes of 50 or less.
- **Small-batch production** quantities typically are in the range of 10 to 100; the equipment consists of general-purpose machines and machining centers.
- **Batch production** usually involves lot sizes of 100 to 5000; the machinery is similar to that used for small-batch production but with specially-designed fixtures for higher productivity.
- **Mass production** involves quantities typically over 100,000. It requires special-purpose machinery, called **dedicated machines**, and automated equipment for transferring materials and parts in progress. These production systems are organized for a specific type of product; consequently, they lack flexibility. Although the machinery, equipment, and specialized tooling are expensive, both the labor skills required and the labor costs are relatively low.

37.2.3 Applications of Automation

Automation can be applied to the manufacturing of all types of goods, from raw materials to finished products, and in all types of production, from job shops to large manufacturing facilities. The decision to automate a new or existing production facility includes the following considerations:

- Type of product manufactured
- Production quantity and rate of production required
- Particular phase of the manufacturing operation to be automated if not all phases are to be automated
- Level of skill in the available workforce
- Reliability and maintenance problems that may be associated with automated machinery and systems
- Economics of the whole operation.

Because automation generally involves high initial equipment cost and requires a knowledge of the operation and maintenance principles, a decision about the implementation of even low levels of automation must involve a careful study of the actual needs of an organization. In some situations, **selective automation**, rather than total automation, of a facility is desirable. As described in the rest of this final part of this book, there are several important and complex issues involved in making decisions about the appropriate level of automation. Generally, if a manufacturing facility is already automated, the operator skill level required is lower.

On the other hand, the application of computers and computer-controlled machinery in the manufacturing environment requires higher sophistication in the workforce than for hard automation. Also, automation requires greater skill in the maintenance and setup workers employed. Thus, it should be recognized that the skills and attributes of a plant's workforce are now different and are tied to the level of automation to be achieved.

37.2.4 Hard Automation

In *hard automation*, also called **fixed-position automation**, the machines are designed to produce a standard product, such as a gear, a shaft, or an engine block. Although product size and processing parameters, such as machining speed, feed, and depth of cut, can be modified, these machines are specialized, hence lack flexibility.

Because the machines are expensive to design and build, their economical use requires production of parts in very large quantities—for example, automotive engines. The machines, generally called *transfer machines* and consisting of *power-head production units* and *transfer mechanisms*, usually are built on the **modular (building-block) principle** (see also Section 25.2.4).

Power-head Production Units. Consisting of a frame or bed, electric drive motors, gearboxes, and tool spindles, these units are self-contained; their components are available commercially in various standard sizes and capacities. Because of this inherent modularity, and thus their adaptability and flexibility, they can easily be regrouped to produce a different part.

Transfer Machines. Typically consisting of two or more powerhead units, these machines can be arranged on the shop floor in linear, circular, or U-shaped patterns. Transfer machines also are used extensively in automated assembly, as described in Section 37.9. *Transfer mechanisms* are used to move the workpiece from one station to another in the machine or from one machine to another, in order to enable various operations to be performed on the part. Workpieces are transferred by such methods as (a) rails, along which the parts (usually placed on pallets) are pushed or pulled by various mechanisms (Fig. 37.4a), (b) rotary indexing tables (Fig. 37.4b), and (c) overhead conveyors.

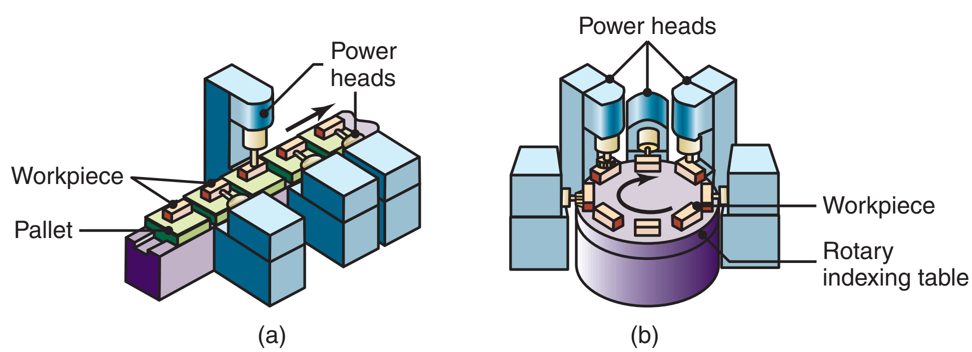


Figure 37.4: Two types of transfer mechanisms: (a) straight rails and (b) circular or rotary.

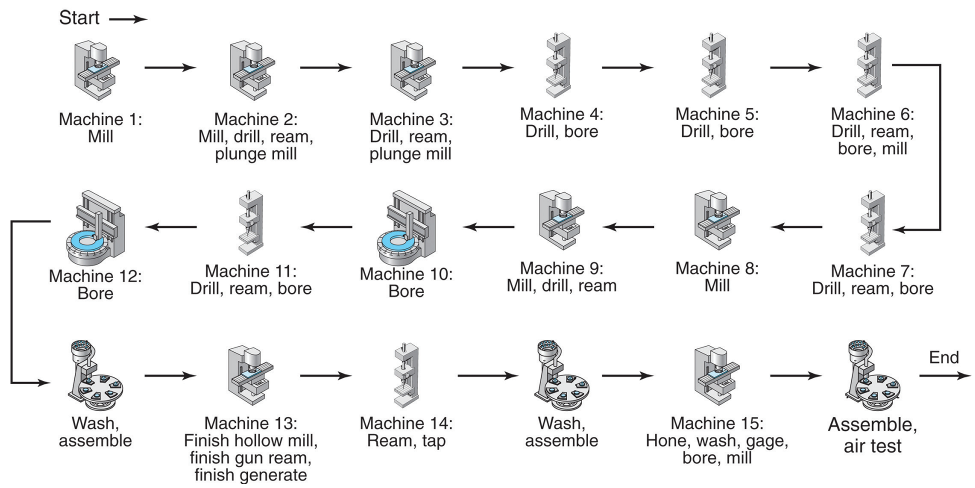


Figure 37.5: A large transfer line for producing engine blocks and cylinder heads.

Transfer Lines. A *transfer line*, or **flow line**, is a very large system for typically producing cylinder heads for engine blocks consisting of a number of transfer machines, as shown in Fig. 37.5. This system is capable of producing 100 cylinder heads per hour; note the variety of operations performed: milling, drilling, reaming, boring, tapping, honing, washing, and gaging. The weight and shape of the workpieces influence the arrangement of the individual machines, important for continuity of operation in the event of tool failure or machine breakdown in one or more of the units. *Buffer storage* features are incorporated in these machines to permit continued operation in such an event.

The transfer of parts from station to station is usually controlled by sensors and other devices. Tools on the machines can easily be changed using toolholders with quick-change features. The machines can be equipped with various automatic gaging and inspection systems, ensuring that the dimensions of a part produced in one station are within acceptable tolerances before that part is transferred to the next station along the line.

37.2.5 Soft Automation

Recall that hard automation generally involves mass-production machines that lack flexibility. Greater flexibility is achieved in *soft automation*, also called **flexible** or **programmable automation**, through the use of computer control of the machine and of its functions; thus, soft automation can produce parts with complex shapes. The machines can easily be reprogrammed to produce a part that has a shape or dimensions different from the one produced just prior to it. Advances in soft automation include the extensive use of modern computers, leading to the development of **flexible manufacturing systems**, with high levels of efficiency and productivity (Section 39.3).

37.2.6 Programmable Logic Controllers

The control of a production process in the proper sequence of operations, especially one involving groups of machines and material-handling equipment, has traditionally been performed by switches, relays, timers, counters, and similar hardwired devices that are based on mechanical, electromechanical, and pneumatic principles.

Beginning in 1968, *programmable logic controllers* (PLCs) were introduced to replace these hardwired devices. Because PLCs eliminate the need for relay control panels and they can be reprogrammed and take less space, they have been adopted widely in manufacturing systems and operations. Their basic functions are (a) on–off, (b) motion, (c) sequential operations, and (d) feedback control. These controllers perform reliably in industrial environments and improve the overall efficiency of an operation. Although they have become less common in new installations (because of advances in numerical-control machines), PLCs still represent a very large installation base. Their sustained popularity is due to their low cost and the proliferation of powerful software to allow programming of PLCs from personal computers, which upload control programs via Ethernet or wireless communications. Modern PLCs are often programmed in specialized versions of BASIC or C **programming languages.** *Microcomputers* are now used more often because they are less expensive and are easier to program and to network. PLCs are also used in system control with high-speed digital-processing and communication capabilities.

Programmable automation controllers is a term often used interchangeably with programmable logic controller, but generally refers to a higher performance product with more advanced communications tools, greater memory, and ability for modular architecture to be integrated with other devices and networks.

Microcontrollers. A number of low-cost but capable systems have been developed, notably the Arduino and Raspberry Pi platforms that are widely used in custom systems and machines. While not as robust as PLCs, these systems use a programming environment that allows rapid incorporation of control logic. These devices use proprietary, relatively simple programming languages, and are being increasingly applied in manufacturing.

37.2.7 Total Productive Maintenance and Total Productive Equipment Management

The management and maintenance of a wide variety of machines, equipment, and systems are among the important aspects affecting productivity in a manufacturing organization. The concepts of *total productive maintenance* (TPM) and *total productive equipment management* (TPEM) include the continued analysis of such factors as:

- Equipment breakdown and equipment problems
- Monitoring and improvement of equipment productivity
- Implementation of preventive and predictive maintenance
- Reduction in setup time, idle time, and cycle time
- Full utilization of machinery and equipment and the improvement of their effectiveness
- Reduction in product defects.

As expected, teamwork is an important component of this activity and involves the full cooperation of machine operators, maintenance personnel, engineers, and the management of the organization (see also **kaizen**, Section 36.1).

37.3 Numerical Control

Numerical control (NC) is a method of controlling the movements of machine components by directly inserting coded instructions into the system. The system then automatically interprets these data and converts them to output signals that, in turn, control various machine components—for example, turning spindles on and off, changing tools, moving the workpiece or the tools along specific paths, and turning cutting fluids on and off.

The importance of numerical control in manufacturing can be illustrated by the following example: Assume that several holes are to be drilled on a part in the positions shown in Fig. 37.6. In the traditional manual method of machining this part, the operator positions the drill bit with respect to the workpiece (see Fig. 23.25), using reference points given by any of the three methods shown in the figure; the operator then

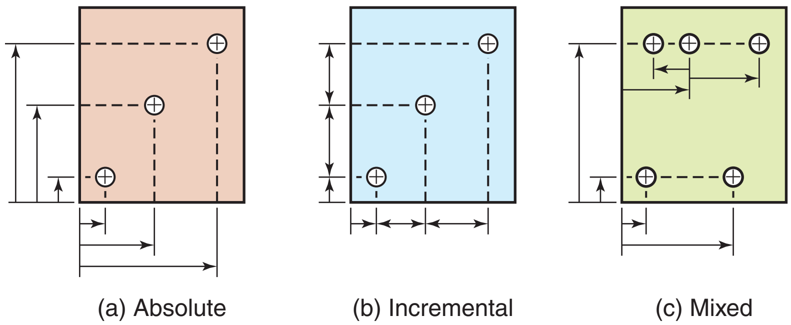


Figure 37.6: Positions of drilled holes in a workpiece. Three methods of measurements are shown: (a) Absolute dimensioning, referenced from one point at the lower left of the part; (b) Incremental dimensioning, made sequentially from one hole to another; and (c) Mixed dimensioning, a combination of both methods.

proceeds to drill the holes. Assume now that 100 parts, all having exactly the same shape and dimensional accuracy, are to be drilled. Obviously, this operation is going to be tedious and time consuming, because the operator has to go through the same motions repeatedly. Moreover, the probability is high that, for a variety of reasons, some of the parts drilled will be different from others.

Assume further that during this production run, the order of processing these parts is changed and that 10 of the parts now require holes in different positions. The machinist now has to reposition the worktable, an operation that is time consuming and subject to error. Such operations can be performed easily by numerical-control machines (see Fig. 25.7), which are capable of producing parts repeatedly and accurately and of handling different parts simply by loading different part programs.

In numerical-control operations, data concerning all aspects of the machining operation, such as tool locations, speeds, feeds, and cutting fluids, are stored on hard disks. On the basis of input information, relays and other devices, known as *hardwired controls*, can be actuated to obtain a desired machine setup. Complex operations, such as turning a part having various contours or die sinking in a milling machine, are now carried out easily. NC machines are used extensively in small- and medium-quantity production (typically 500 or fewer parts) of a wide variety of parts, both in small shops and in large manufacturing facilities.

Example 37.1 Historical Origin of Numerical Control

The basic concept behind numerical control appears to have been implemented in the early 1800s, when punched holes in sheet-metal cards were used to automatically control the movements of weaving machines. Needles were activated by sensing the presence or absence of a hole in the card. This invention was followed by automatic piano players (*Pianolas*™), in which the keys were activated by air flowing through holes punched in a perforated roll of paper.

The principle of numerically controlling the movements of machine tools was first conceived in the 1940s by J.T. Parsons (1913–2007), an American inventor, in his attempt to machine complex helicopter blades. The first prototype NC machine, built in 1952 at the Massachusetts Institute of Technology, was a vertical-spindle, two-axis copy-milling machine retrofitted with servomotors; the machining operations consisted of end milling and face milling (Chapter 24) on a thick aluminum plate.

The numerical data to be punched into the paper tapes were generated by a digital computer (another invention that was being developed at the same time at MIT). In the first experiments, parts were machined successfully, accurately, and repeatedly without any operator intervention. On the basis of this success, the machine-tool industry began designing, building, and marketing NC machine tools. Later, these machines were equipped with computer-numerical controls (CNCs), with greater flexibility, accuracy, versatility, and ease of operation.

37.3.1 Computer Numerical Control

In the next step in the development of numerical control, the control hardware (mounted on the NC machine) was converted to local computer control by software.

- In **direct numerical control** (DNC), several machines were controlled directly (step by step) by a central mainframe computer. In this system, the operator had access to the central computer through a remote terminal. With DNC, the status of *all* machines in a manufacturing facility can be monitored and assessed from a central computer. However, DNC has a crucial disadvantage in that if the computer shuts down for some reason, all of the machines become inoperative.

- In **distributed numerical control**, a central computer serves as the control system over a number of individual CNC machines having onboard microcomputers. This approach was valuable when computers lacked the memory necessary to store all of the data for machining operations necessary to produce a part. DNC became less popular in the 1990s, although using a central computer is still common for collecting, storing, and analyzing data. A recent development is to use **cloud storage**, where files are stored on the servers of a third company and accessed, as needed, through Internet tools.

- **Computer numerical control** (CNC) is a system in which a control computer is an integral part of a machine (*onboard computer*). The machine operator can program onboard computers, modify the programs directly, prepare programs for different parts, and store the programs. CNC systems are widely used today because of the availability of (a) small computers with large memory, (b) low-cost programmable controllers and microprocessors, and (c) program-editing capabilities.

- A current activity is to apply *cloud computing* for CNC. In cloud computing, sufficient data (from incorporated sensors) is communicated from the machine tool to another computer that has more processing power and has the ability to solve complicated physics models, and then provide instructions to the machine in real time. Cloud computing in this context is controversial, because of the concerns of cybersecurity.

37.3.2 Principles of NC Machines

The basic elements and operation of a typical NC machine are shown in Fig. 37.7. The functional elements in numerical control and the components involved are the following:

- **Data input:** The numerical information is read and stored in computer memory

- **Data processing:** The programs are read into the machine control unit for processing

- **Data output:** This information is translated into commands (typically, pulsed commands) to the servomotor (Fig. 37.8). The servomotor then moves the worktable to specific positions, through linear or rotary movements by means of stepping motors, leadscrews, or other similar devices.

Types of Control Circuits. An NC machine can be controlled through two types of circuits. In the **open-loop** system (Fig. 37.8a), the signals are sent to the servomotor by the controller, but the movements and final positions of the worktable are not checked for accuracy. In contrast, the **closed-loop** system (Fig. 37.8b) is equipped with various transducers, sensors, and counters that accurately measure the position of the worktable. Through *feedback control*, the position of the worktable is compared against the signal; the table movements terminate when the proper coordinates are reached.

The feedback signal can be simply computed, such as comparing a measured distance and comparing it to a desired position; the difference can be multiplied by a constant (the *gain*) to obtain the correcting signal. On the other hand, complex physics-based models of the manufacturing process can be used. This

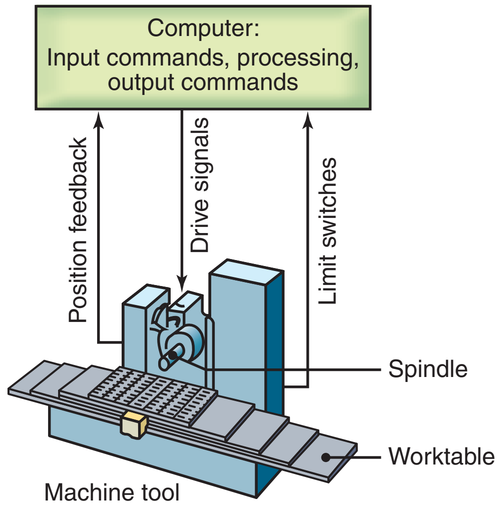


Figure 37.7: Schematic illustration of the major components for position control on a numerical-control machine tool.

would allow, for example, the ability to decide on a number of corrective actions, compilation of data for the workpiece that can be used for performance models (see *digital twin*, Section 39.8.1), or determination of preferred machine states (as with stability lobes, see Section 25.4).

Position measurement in NC machines is carried out through two methods (Fig. 37.9). In *indirect measuring systems, rotary encoders*, or *resolvers*, rotary movement is converted to translation. Backlash (the play between two adjacent mating gear teeth), however, can significantly affect measurement accuracy. Position feedback mechanisms utilize various sensors that are based mainly on magnetic and photoelectric principles. In *direct measuring systems*, a sensing device reads a graduated scale on the machine table, or slide, for linear movement (Fig. 37.9c). This system is more accurate because the scale is built into the machine and thus backlash in the mechanisms is not significant.

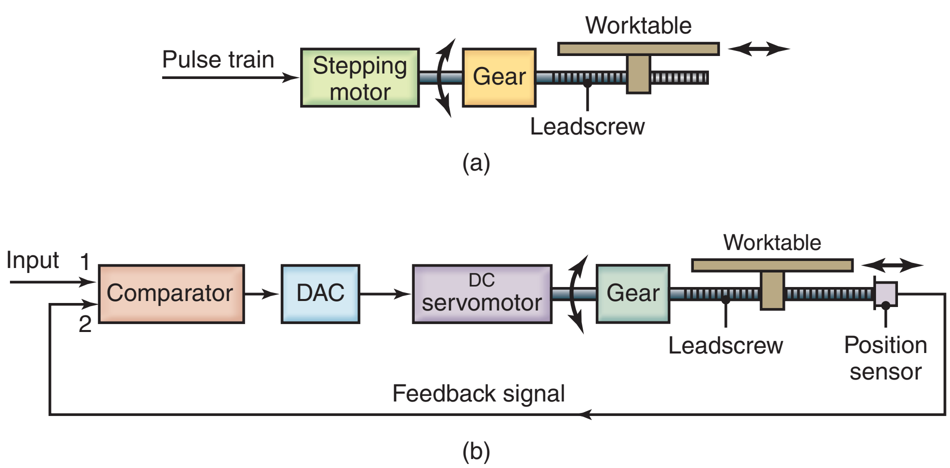


Figure 37.8: Schematic illustration of the components of (a) an open-loop and (b) a closed-loop control system for a numerical-control machine. DAC = digital-to-analog converter.

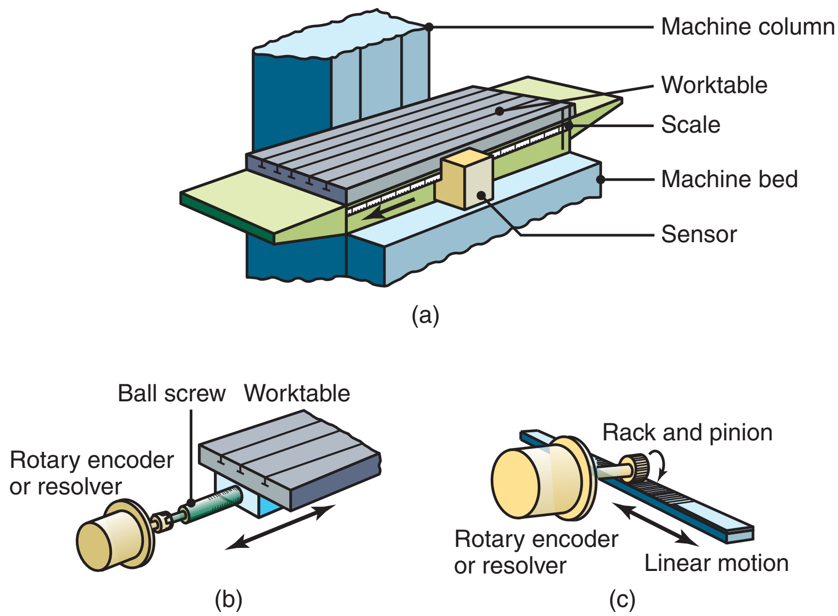


Figure 37.9: (a) Direct measurement of the linear displacement of a machine-tool worktable. (b) and (c) Indirect measurement methods.

37.3.3 Types of Control Systems

There are two basic types of control systems in numerical control:

1. In a **point-to-point system**, also called a *positioning system*, each axis of the machine is driven separately by leadscrews and at different velocities, depending on the type of operation. The tool moves initially at maximum velocity, in order to reduce nonproductive time, but then it decelerates as the tool approaches its numerically defined position. Thus, in an operation such as drilling a hole, the positioning and drilling take place *sequentially* (Fig. 37.10a).

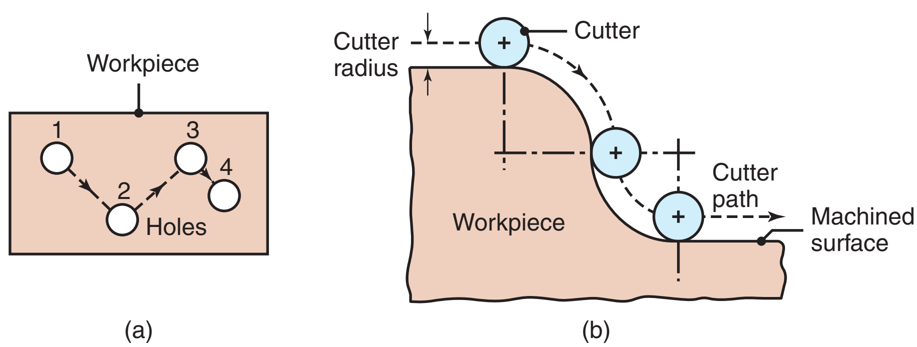


Figure 37.10: Movement of tools in numerical-control machining. (a) Point-to-point, in which the drill bit drills a hole at position 1, is retracted and moved to position 2, and so on. (b) Continuous path by a milling cutter. Note that the cutter path is compensated for by the cutter radius; the path also can be compensated for cutter wear.

After the hole is drilled, the tool retracts upward and moves rapidly to another specified position, and the operation is repeated. The tool path followed from one position to another is important in only one respect: it must be chosen to minimize the time of travel for better efficiency. Point-to-point systems are used mainly in drilling, punching, and straight milling operations.

2. In a **contouring system**, also known as a *continuous-path system*, the positioning and the operations are both performed along controlled paths but at different velocities. Because the tool cuts as it travels along a prescribed path (Fig. 37.10b), the control and synchronization of velocities and movements are important for accuracy. The contouring system is typically used on lathes, milling machines, grinders, welding machinery, and machining centers.

Interpolation. Movement of the tool along a path (*interpolation*) occurs incrementally by one of several basic methods (Fig. 37.11). Examples of actual paths in drilling, boring, and milling operations are shown in Fig. 37.12. In all interpolations, the path controlled is that of the *center of rotation* of the tool. However, compensation for different types of tools, different diameters of tools, or for tool wear during machining can be made in the NC program.

- In **linear interpolation**, the tool moves in a straight line from start to end (Fig. 37.11a) along two or three axes. Although, theoretically, all types of profiles can be produced by this method by making the increments between the points small (Fig. 37.11b), a large amount of data has to be processed in order to do so.
- In **circular interpolation** (Fig. 37.11c), the inputs required for the path are the coordinates of the end points, the coordinates of the center of the circle and its radius, and the direction of the tool travels along the arc.
- In **parabolic interpolation** and **cubic interpolation**, the tool path is approximated by curves based on higher-order mathematical equations (see *B Splines*, Section 38.4.2). This method is effective in 5-axis machines and is useful in die-sinking operations, such as for dies for sheet forming of automotive bodies. The interpolations also are used for the movements of industrial robots (Section 37.6).

37.3.4 Positioning Accuracy in Numerical Control

Positioning accuracy in numerical-control machines is defined by how accurately the machine can be positioned with respect to a certain coordinate system. *Repeat accuracy* is defined as the closeness of the agreement of the repeated position movements under the same operating conditions of the machine. *Resolution*, also called *sensitivity*, is the smallest increment of motion of the machine components (see also Section 35.7).

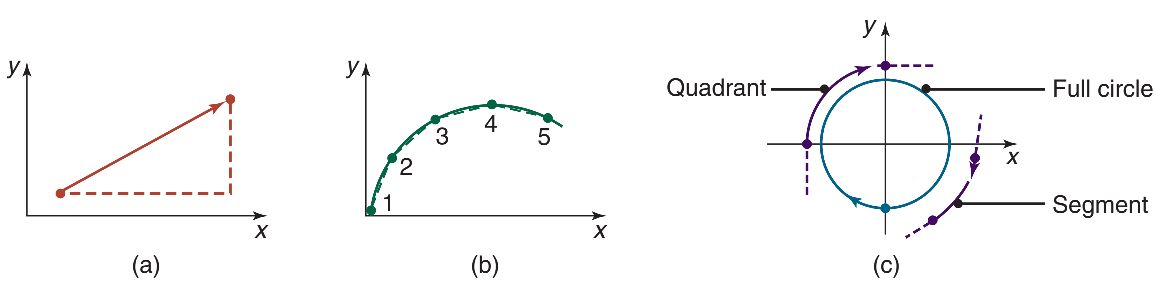


Figure 37.11: Types of interpolation in numerical control: (a) linear, (b) continuous path, approximated by incremental straight lines, and (c) circular.

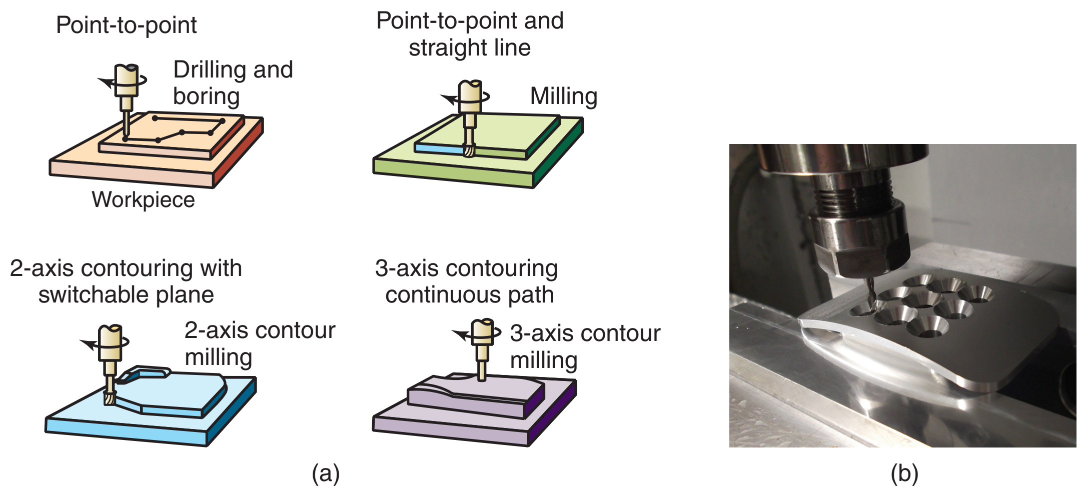


Figure 37.12: (a) Schematic illustration of drilling, boring, and milling along various paths. (b) Machining a sculptured surface on a 5-axis numerical-control machine. *Source:* Courtesy of Scott Smith and Geoff Heacock UNC –Charlotte.

The *stiffness* of the machine tool and the backlash in gear drives and leadscrews are important factors in achieving dimensional accuracy. Backlash in modern machines is eliminated by using preloaded ball screws. Also, a rapid response to command signals requires that friction in machine slideways and inertia be minimized. The latter can be achieved by reducing the mass of moving components of the machine, such as, for example, by using lightweight materials, including ceramics (see also Section 25.3).

37.3.5 Advantages and Limitations of Numerical Control

Numerical control has the following advantages over conventional methods of machine control:

- Greater flexibility of operation, as well as the ability to produce complex shapes with good dimensional accuracy and repeatability; high production rates, productivity, and product quality; and lower scrap loss
- Machine adjustments are easy to make
- More operations can be performed with each setup, and the lead time required for setup and machining is less than the lead time required in conventional methods
- Programs can be prepared rapidly and can be recalled at any time
- Operator skill required is less than that for a qualified machinist and the operator has more time to attend to other tasks in the work area.

The major limitations of NC are (a) the relatively high initial cost of the equipment, (b) the need and cost for programming, and (c) the special maintenance required.

37.3.6 Programming for Numerical Control

A *program* for numerical control consists of a sequence of directions that cause an NC machine to carry out a certain operation, machining operations of all forms being the most commonly applied. *Programming for NC* may be

- Performed by computer software, such as *G-Code generators*, that produce numerical code from geometry data files; such software is now very common, with numerous open-source codes available.

- Done on the shop floor (rare but minor changes in G-code are common)
- Purchased from an outside source.

The program contains instructions and commands:

1. *Geometric instructions* pertain to relative movements between the tool and the workpiece
2. *Processing instructions* concern spindle speeds, feeds, cutting tools, cutting fluids, and so on
3. *Travel instructions* pertain to the type of interpolation and to the speed of movement of the tool or the worktable
4. *Switching instructions* concern the on–off position for coolant supplies, direction or lack of spindle rotation, tool changes, workpiece feeding, clamping, and so on.

Complex parts can be machined with the use of graphics-based computer-aided machining programs. Standardized programming languages such as STEP-NC and the older but still common G-Code are used for communicating machining instructions to the CNC hardware.

The STEP-NC software is a standardized language that has been extended beyond machine tools, and incorporates models for milling, turning, and EDM. Plasma cutting and laser welding and cutting systems also have been developed for use with STEP-NC.

In *shop-floor programming*, CNC programming software is used directly on the machine tool controller. This allows higher-level geometry and processing information to be sent to the CNC controller instead of G-code. G-code is then developed by the dedicated computer under the control of the machine operator. The advantage of this method is that any changes that are made to the machining program is sent back to the programming group and is stored as a shop-proven design iteration; it can be re-used or standardized across a family of parts.

37.4 Adaptive Control

Adaptive control (AC) is basically a *dynamic-feedback* system in which the operating parameters automatically adapt themselves to conform to new circumstances; it is thus a logical extension of computer numerical control systems. Human reactions to occurrences in everyday life already contain dynamic-feedback control. For example, driving a car on a smooth road is relatively easy; on a rough road, however, it has to be steered in order to avoid potholes by *visually* and *continuously* observing the condition of the road directly ahead of the car. Moreover, the driver's body feels the car's rough movements and vibrations, and then reacts by changing the direction or the speed of the car to minimize these effects. An **adaptive controller** continuously checks road conditions, calculates an appropriate desired braking profile (e.g., an antilock brake system and traction control), and then uses feedback to implement it.

As described in Section 37.3, in manufacturing operations the part programmer sets the processing parameters on the basis of the existing knowledge of the workpiece material and relevant data on the particular operation. In CNC machines, these parameters are held constant during a particular process cycle; in AC, however, the system is capable of automatic adjustments *during* the operation, through closed-loop feedback control (Fig. 37.13). Several adaptive-control systems are commercially available for a variety of applications.

Adaptive Control in Manufacturing. The main purposes of adaptive control in manufacturing are to

- Optimize production rate
- Optimize product quality
- Minimize production cost.

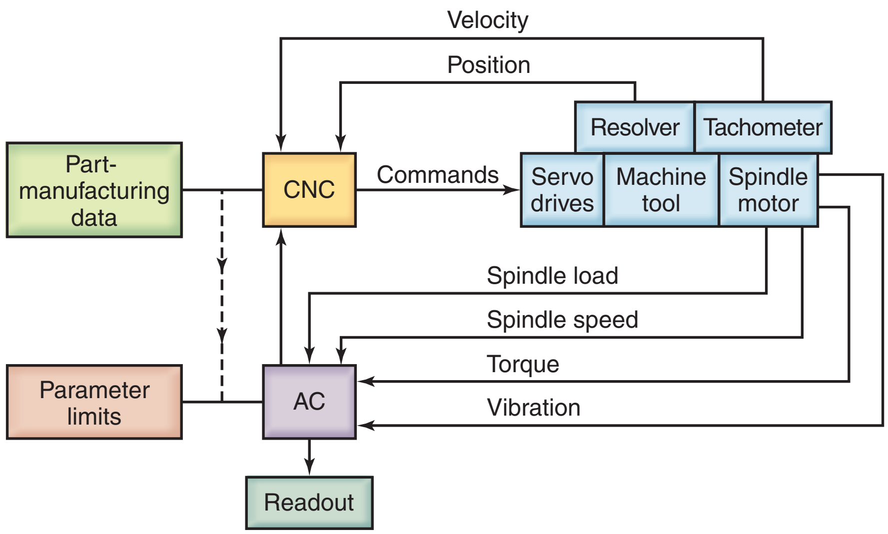


Figure 37.13: Schematic illustration of the application of adaptive control (AC) for a turning operation. The system monitors such parameters as cutting force, torque, and vibrations; if these parameters are excessive, it modifies process variables (such as feed and depth of cut) to bring the parameters back to acceptable levels.

Consider a machining operation, such as turning on a lathe (Section 23.2). The adaptive control system senses, *in real time*, parameters such as cutting forces, spindle torque, temperature rise, tool wear, tool condition, and surface finish produced. The AC system converts this information into commands that modify the process parameters to hold them within certain limits and, thus, optimize the machining operation.

Those systems that place a constraint on a process variable, such as force, torque, or temperature, are called **adaptive-control constraint** (ACC) systems. Thus, for example, if the cutting force (and hence the torque) increases excessively (due, say, to the presence of a hard region in a casting), the system modifies the cutting speed or the feed as necessary, in order to lower the cutting force to an acceptable level (Fig. 37.14).

Without AC or the direct intervention of the operator (as has been the case in traditional machining operations), high cutting forces may cause tool failure or the workpiece to deflect or distort excessively. As a result, workpiece dimensional accuracy and surface finish begin to deteriorate. Those systems that

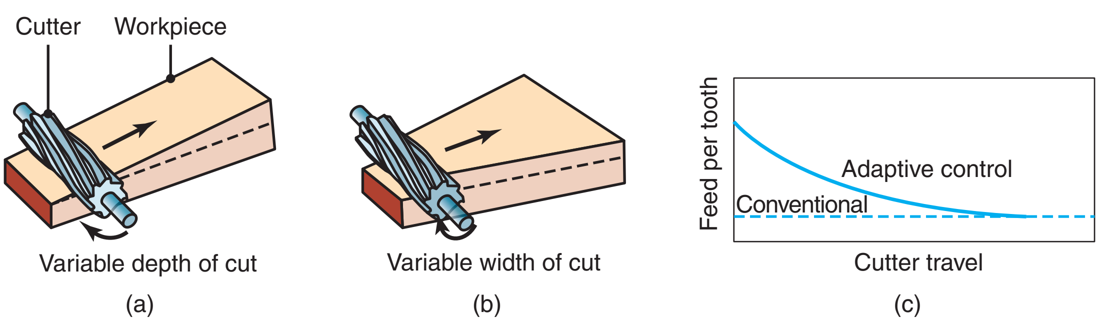


Figure 37.14: An example of adaptive control in milling. As the depth of cut (a) or the width of cut (b) increases, the cutting forces and the torque increase. The system senses this increase and automatically reduces the feed (c) to avoid excessive forces or tool breakage in order to maintain cutting efficiency. *Source:* After Y. Koren.

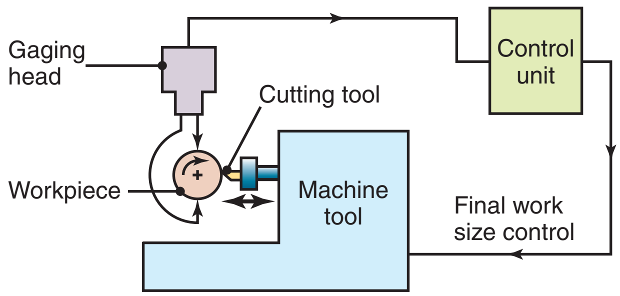


Figure 37.15: In-process inspection of a workpiece diameter in a turning operation. The system automatically adjusts the radial position of the cutting tool in order to produce the correct diameter.

optimize an operation are called **adaptive-control optimization** (ACO) systems. Optimization may involve maximizing the material-removal rate between tool changes or improving the surface finish of the part.

Gain scheduling is the simplest form of adaptive control, where a different gain is selected depending on the measured operating conditions and a different gain is assigned to each region of the system's operating space. With advanced adaptive controllers, the gain may vary continuously with changes in operating conditions. As described in Section 37.3, the part programmer sets the processing parameters, based on the existing knowledge of the workpiece material and various data on the particular manufacturing process. Whereas in CNC machines, these parameters are held constant during a particular processing cycle, in adaptive control the system is capable of automatic adjustments *during* processing through closed-loop feedback control (Fig. 14.13).

Response time must be short for AC to be effective, particularly in high-speed machining operations (Section 25.5). Assume, for example, that a turning operation is being performed on a lathe at a spindle speed of 1000 rpm, and the tool suddenly breaks, adversely affecting the surface finish and dimensional accuracy of the part. Obviously, in order for the AC system to be effective, the sensing system must respond within a very short time; otherwise the damage to the workpiece will become extensive.

In adaptive control, *quantitative* relationships must therefore be established and coded in the computer software as mathematical models. If, for instance, the tool-wear rate in a machining operation is excessive (Section 21.5), the computer must be able to (a) calculate how much of a change in speed or in feed is necessary and (b) decide whether to increase or decrease the speed or the feed, in order to reduce the tool-wear rate to an acceptable level. The system also must be able to compensate for dimensional changes in the workpiece due to such factors as tool wear and temperature rise (Fig. 37.15).

It is apparent that, coupled with CNC, adaptive control is a powerful tool in optimizing manufacturing operations. While it has significant demonstrable benefits with respect to quality and productivity, it has not been extensively applied to date. Most machine-tool control systems are geometry-based and do not incorporate adaptive control. The latest versions of STEP-NC programming languages, however, allows for deflection and wear management.

37.5 Material Handling and Movement

During a typical manufacturing operation, raw materials and parts in progress are moved from storage to machines, from machine to machine, from assembly to inventory, and, finally, to shipment. For example, (a) workpieces are loaded on machines, as when a forging is mounted on a milling-machine bed for machining; (b) sheet metal is fed into a press for stamping; (c) parts are removed from one machine and loaded onto another, as when a machined forging is to be subsequently ground for better surface finish and dimensional accuracy; and (d) finished parts are assembled into a final product.

Material handling is defined as the functions and systems associated with the transportation, storage, and control of materials and parts in the total manufacturing cycle of a product. The total time required for actual manufacturing operations depends on the part size and shape and on the set of operations to be performed. Note that idle time and the time required for transporting materials can constitute the majority of the time consumed, thus reducing productivity. Material handling must be an integral part of the planning, implementing, and control of manufacturing operations; moreover, it must be repeatable and predictable.

Plant layout is an important aspect of the orderly flow of materials and components throughout the manufacturing cycle. The time and distances required for moving raw materials and parts should be minimized (see *lean manufacturing*, Section 39.7). For parts requiring multiple operations, which typically is the case, equipment should be grouped around the operator or an industrial robot or robots (see also *cellular manufacturing*, Section 39.2).

Material-handling Methods. Several factors must be considered in selecting a suitable material handling method for a particular manufacturing operation:

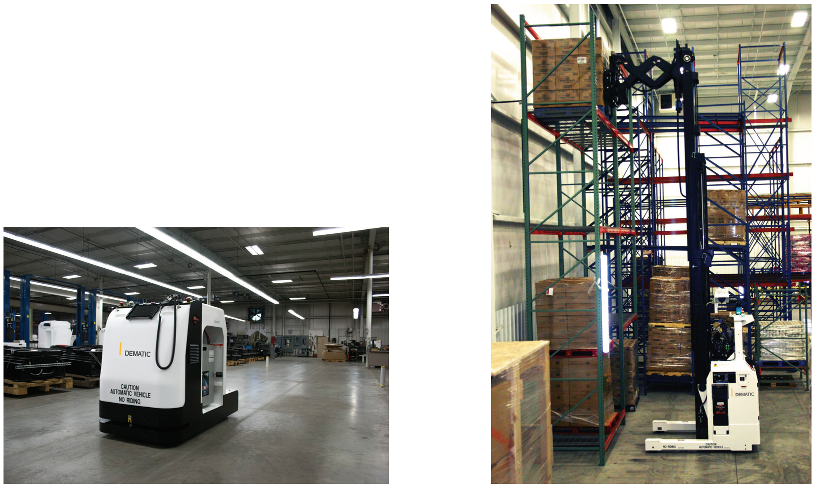
1. Shape, weight, and characteristics of the parts
2. Distances involved and the position and orientation of the parts during their movement and at final destination
3. Conditions of the path along which the parts are to be transported
4. Level of automation, the controls needed, and any integration with other equipment and systems
5. Operator skill required
6. Economic considerations.

For small-batch manufacturing operations, raw materials and parts can be handled and transported manually, but this method is generally costly; also, because it involves humans, the practice can be unpredictable, unreliable, and unsafe. In contrast, in automated manufacturing plants, computer-controlled material and parts flow are implemented, resulting in improved productivity at lower labor costs.

Equipment. The types of equipment that can be used for moving materials and parts in progress may consist of conveyors, rollers, carts, forklift trucks, self-powered monorails, and various mechanical, electrical, magnetic, pneumatic, and hydraulic devices and manipulators. **Manipulators** are designed to be controlled directly by the operator, or they can be automated for repeated operations, such as the loading and unloading of parts from machine tools, presses, and furnaces. Manipulators are capable of gripping and moving heavy parts and of orienting them, as necessary, between the manufacturing and assembly operations. Workpieces are transferred directly from machine to machine. Machinery combinations having the capability of conveying parts without the use of additional material-handling apparatus are called **integral transfer devices**.

Flexible material handling and movement with real-time control has become an integral part of modern manufacturing. Industrial robots, specially-designed pallets, and **automated guided vehicles** (AGV) are used extensively in *flexible manufacturing systems* to move parts and to orient them as required (Fig. 37.16). AGVs operate automatically along pathways in a plant, with in-floor wiring or tapes for optical scanning for guidance without operator intervention. This transport system has high flexibility and is capable of random delivery to different workstations. It optimizes the movement of materials and parts in cases of congestion around workstations, machine breakdown (*downtime*), or the failure of one section of the production system.

The movements of AGVs are planned such that they interface with **automated storage/retrieval systems** (AS/RS), in order to utilize warehouse space efficiently and to reduce labor costs. However, these systems are now considered undesirable because of the current focus on *minimal inventory* and on *just-in-time* production methods (Section 39.6).



(a) (b)

Figure 37.16: (a) An automated guided vehicle (Tugger type). This vehicle can be arranged in a variety of configurations to pull caster-mounted cars; it has a laser sensor to ensure that the vehicle operates safely around people and various obstructions. (b) An automated guided vehicle configured with forks for use in a warehouse. *Source:* Courtesy of Dematic Corp.

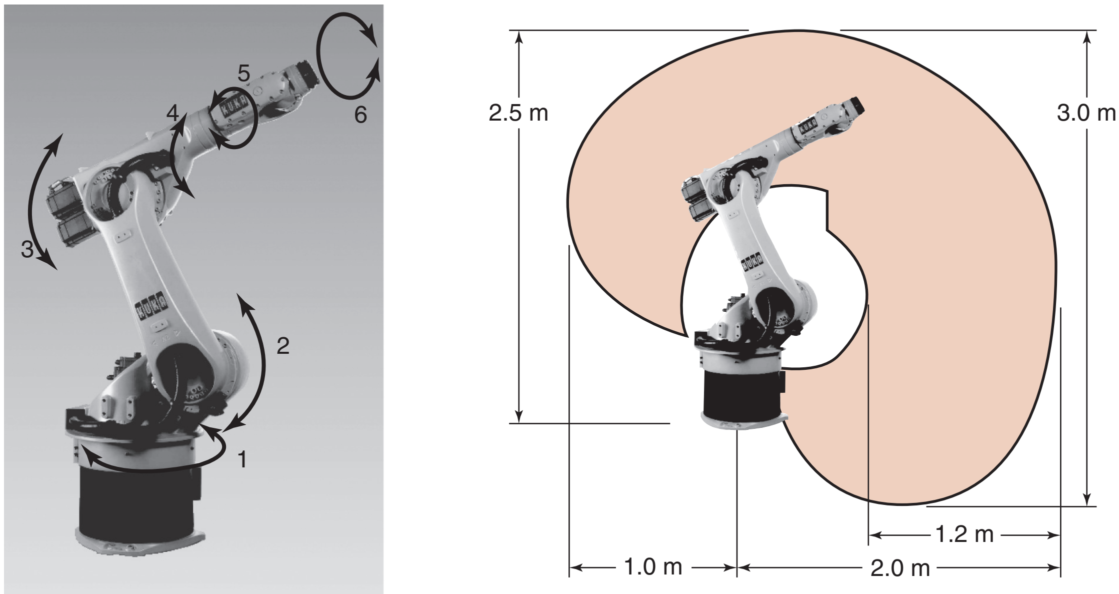
Coding systems that have been developed to locate and identify parts in progress throughout the manufacturing system and to transfer them to their appropriate stations are outlined as:

- **Bar coding**, including **QR barcodes**, is the most widely used coding system and the least costly.
- **Magnetic strips** constitute the second most common system.
- **RF** (radio frequency) **tags** are sometimes used; although expensive, they do not require the clear line-of-sight essential in the previous two systems, have a long range, and are rewritable.
- **Acoustic waves, optical character recognition**, and **machine vision** are the principles of other identification systems.

37.6 Industrial Robots

The word **robot** was coined in 1920 by the Czech author K. Čapek in his play *R.U.R.* (Rossum's Universal Robots); it is derived from the word *robota*, meaning "worker." An **industrial robot** has been described by the International Organization for Standardization (ISO) as a machine formed by a mechanism including several degrees of freedom, often having the appearance of one or several arms ending in a wrist capable of holding a tool, a workpiece, or an inspection device. In particular, an industrial robot's control unit must use a memorizing method and may also use sensing or adapting features to take into account environment and special circumstances.

Industrial robots were first introduced in the early 1960s. Computer-controlled robots were commercialized in the early 1970s, and the first robot controlled by a microcomputer appeared in 1974. Industrial robots were first employed in hazardous operations, such as the handling of toxic and radioactive materials and the loading and unloading of hot workpieces from furnaces in foundries.



(a) (b)

Figure 37.17: (a) Schematic illustration of a 6-axis KR-30 KUKA robot; the payload at the wrist is 30 kg and repeatability is $\pm$ 0.15 mm. The robot has mechanical brakes on all of its axes, which are directly coupled to one another. (b) The work envelope of the robot, as viewed from the side. *Source:* Courtesy of KUKA Robotics Corp.

Simple rule-of-thumb applications for robots are described as the three D's, for *dull, dirty,* and *dangerous* (a fourth D would be for *demeaning*) and the three H's, for *hot, heavy,* and *hazardous*. Industrial robots are now indispensable components in almost all manufacturing operations and have greatly improved productivity at reduced labor costs.

37.6.1 Robot Components

An industrial robot (Fig. 37.17) has a number of basic components.

Manipulator. Also called an **arm and wrist**, the *manipulator* is a mechanical device that provides motions (trajectories) similar to those of a human arm and hand. The end of the wrist can reach a point in space having a specific set of coordinates and in a specific orientation. Most robots have six rotational joints; seven degrees of freedom (*redundant* robots) also are available for special applications.

End Effector. The end of the wrist in a robot is equipped with an *end effector*, also called *end-of-arm tooling*. Depending on the type of operation, conventional end effectors may be equipped with any of the following devices (Fig. 37.18):

- Grippers, hooks, scoops, electromagnets, vacuum cups, and adhesive fingers for material handling (Fig. 37.19a)
- Spray guns for painting
- Attachments for spot and arc or laser welding and for arc cutting
- Power tools, such as drills, nut drivers, and burrs
- Measuring instruments.

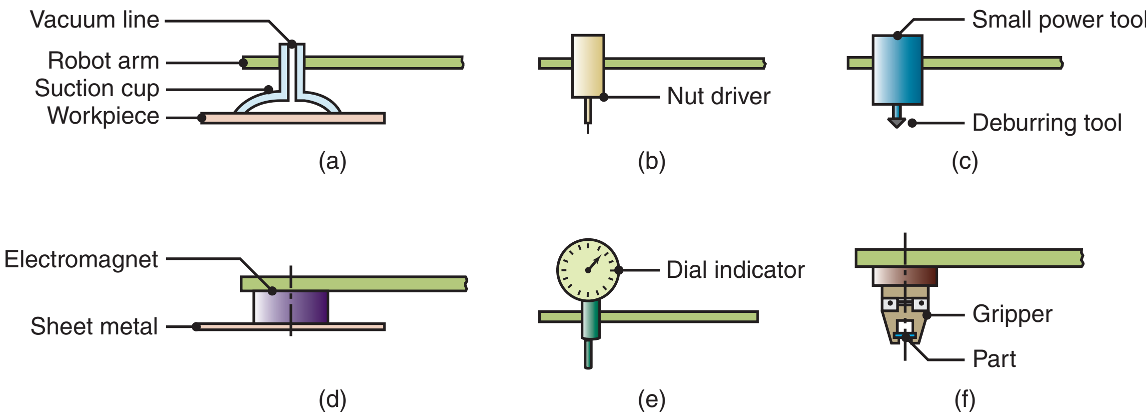


Figure 37.18: Types of tools and devices attached to end effectors to perform a variety of operations.

Equipped with two or more fingers, mechanical grippers are used most commonly. **Compliant end effectors** are typically used to handle fragile materials or to facilitate assembly; they can use mechanisms that limit the force applied to the workpiece or they can be designed with a specific stiffness. The selection of an appropriate end effector for a specific application depends on such factors as the payload, environment, reliability, and cost; consequently, end effectors generally are custom made to meet specific handling requirements.

Figure 37.19: Example of end effectors Gripper with three fingers. *Source:* Courtesy of Robotiq.

Power Supply. Each motion of the manipulator (linear or rotational) is controlled and regulated by independent actuators that use an electrical, pneumatic, or hydraulic power supply. Each source of energy and each type of motor has its own characteristics, applications, advantages, and limitations.

Controller. Also known as the *control system*, the controller is the communications and information-processing system that gives commands for the movements of the robot; it is the brain of the robot and stores data to initiate and terminate movements of the manipulator. The controller is also the nervous system of the robot; it interfaces with computers and other equipment, such as manufacturing cells or assembly systems.

Feedback devices, such as transducers, are an important part of the control system. Robots with a fixed set of motions have *open-loop control*, in which commands are given and the robot arm goes through its motions. Unlike feedback in *closed-loop systems*, in open-loop systems the accuracy of the movements is not monitored; consequently, an open-loop system does not have a self-correcting capability.

Depending on the particular task, the *positioning repeatability* required may be as small as 0.050 mm, as would be required in assembly operations for electronic printed circuit boards (Section 28.13). Accuracy and repeatability vary greatly with payload and with position within the work envelope.

37.6.2 Classification of Robots

Robots may be classified by basic type (Fig. 37.20):

1. **Cartesian**, or **rectilinear**
2. **Cylindrical**
3. **Spherical**, or **polar**
4. **Articulated**, *revolute, jointed*, or *anthropomorphic*.

Robots may be attached permanently to the floor of a plant; they may also move along overhead rails (*gantry robots*) or be equipped with wheels to move along the factory floor (*mobile robots*).

Fixed-sequence and Variable-sequence Robots. The *fixed-sequence robot*, also called a **pick-and-place robot**, is programmed for a specific sequence of operations; its movements are from point to point and the cycle is repeated continuously. These robots are simple and relatively inexpensive. The *variable-sequence* robot also is programmed for a specific sequence of operations but it can be *reprogrammed* to perform a different sequence of operations.

Playback Robot. An operator leads or walks the *playback robot* and its end effector through the desired path; in other words, the operator *teaches* the robot by showing it what to do. The robot records the path

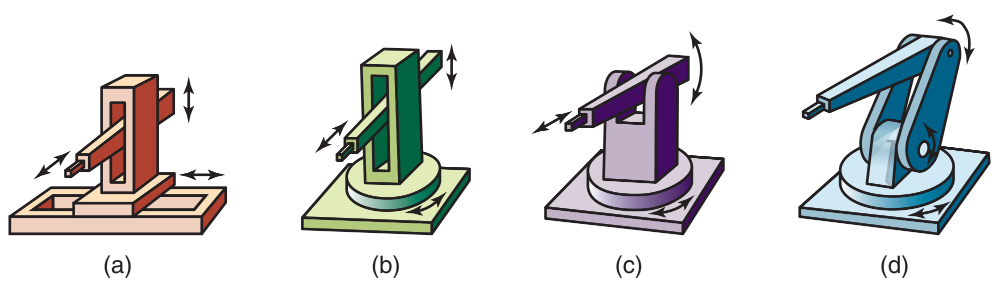


Figure 37.20: Four types of industrial robots: (a) Cartesian (rectilinear), (b) cylindrical, (c) spherical (polar), and (d) articulated (revolute, jointed, or anthropomorphic).

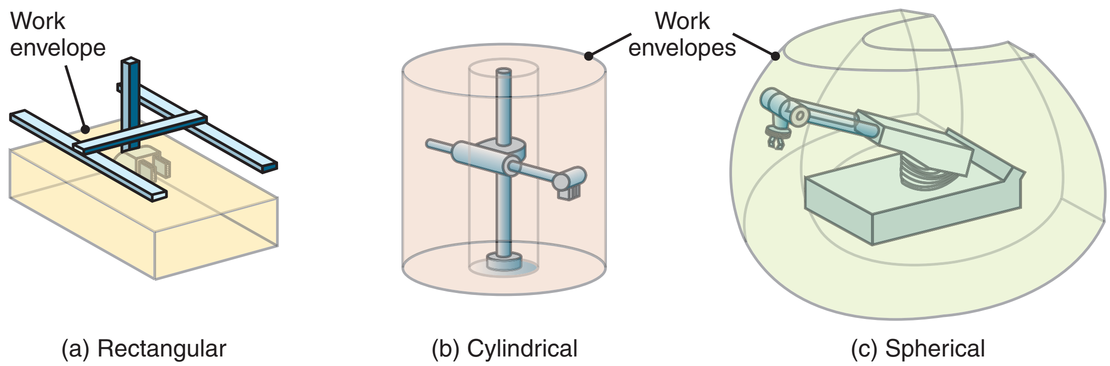


Figure 37.21: Work envelopes for three types of robots; their selection depends on the particular application. (See also Fig. 37.17b.)

and sequence of the motions and can repeat them continually without further action or guidance by the operator. The **teach pendant** robot utilizes handheld button boxes connected to the control panel; the pendants are used to control and guide the robot and its tooling through the work to be performed. These movements are then registered in the memory of the controller and are automatically reenacted by the robot whenever needed.

Numerically Controlled Robot. This type of robot is programmed and operated much like a numerically controlled machine; it is servocontrolled by digital data and its sequence of movements can be modified with relative ease. As in NC machines, there are two basic types of control. (1) *Point-to-point* robots are easy to program and have a higher load-carrying capacity and a larger **work envelope**, the maximum extent or reach of the robot hand or working tool in all directions, as shown in Fig. 37.21. (2) *Continuous-path* robots have greater accuracy than point-to-point robots but they have lower load-carrying capacity.

Intelligent Robot. The *intelligent robot*, also called a *sensory robot*, is capable of performing some of the functions and tasks carried out by humans. The robot is equipped with a variety of sensors with *visual* (*computer vision*) and *tactile* capabilities (Section 37.7). Much like humans, the robot observes and evaluates its immediate environment and its own proximity to other objects (especially machinery) by perception and *pattern recognition*. It then makes appropriate decisions for the next movement and proceeds accordingly.

Developments in intelligent robots include the following:

- Behaving more like humans and performing tasks such as moving among a variety of machines and equipment on the shop floor and avoiding collisions
- Recognizing, selecting, and properly gripping the correct raw material or workpiece for further processing
- Transporting a part from machine to machine
- Assembling components into subassemblies or into a final product.

37.6.3 Applications and Selection of Robots

Major applications of industrial robots include the following:

- Material-handling operations can be performed reliably and repeatedly, thereby improving quality and reducing scrap losses. Some examples are (a) casting and molding operations in which molten

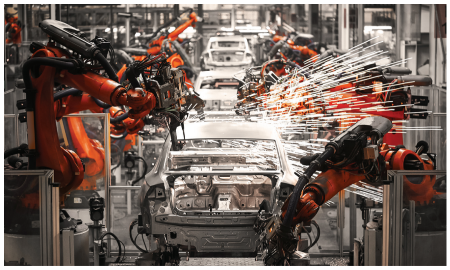


Figure 37.22: Example of industrial robot applications. Spot welding automobile bodies with industrial robots. *Source:* Shutterstock/Jensen.

metal, raw materials, and parts in various stages of completion are handled without operator interference; (b) heat-treating operations in which parts are loaded and unloaded from furnaces and quench baths; and (c) operations in which parts are loaded and unloaded from presses and various other types of metalworking machinery.

- Spot welding of automobile and truck bodies, producing welds of good quality (Fig. 37.22). Robots also perform other, similar operations, such as arc welding, arc cutting, and riveting.
- Operations such as deburring, grinding, and polishing can be done by using appropriate tools attached to end effectors.
- Applying adhesives and sealants.
- Spray painting (particularly of complex shapes) and cleaning operations are frequent applications because the motions required for treating one piece are repeated accurately for the next piece.
- Automated assembly (Fig. 37.23).
- Inspection and gaging at speeds much higher than those that can be achieved by humans.

Robot Selection. Factors that influence the selection of robots in manufacturing are:

- Load-carrying capacity
- Work envelope (see Figs. 37.17b and 37.21)
- Speed of movement
- Reliability
- Repeatability
- Arm configuration
- Degrees of freedom
- The control system
- Program memory

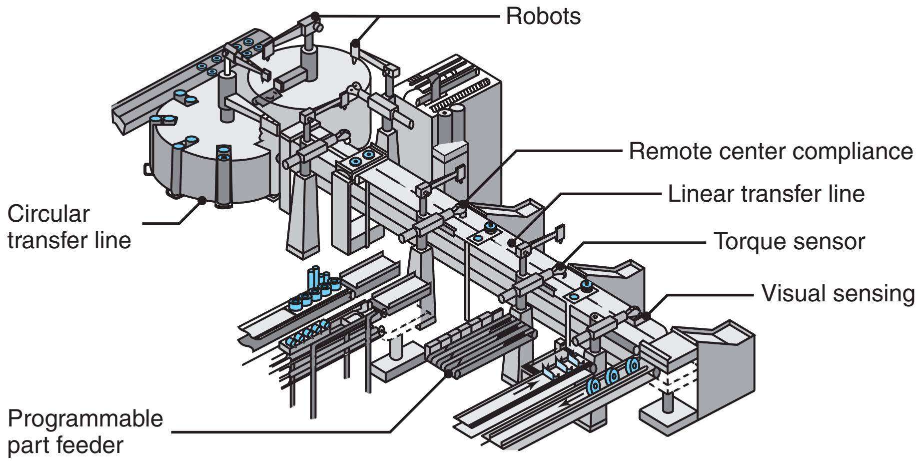


Figure 37.23: Automated assembly operations using industrial robots and circular and linear transfer lines.

Economics. In addition to the technical factors, cost and benefit considerations are major aspects of robot selection and their use. The increasing availability, reliability, and their reduced costs, intelligent robots are having a major impact on manufacturing operations.

Robot Safety. Depending on the size of the robot's work envelope, speed, and proximity to humans, safety considerations in a robot environment are important, particularly for programmers and maintenance personnel who are in direct physical interaction with robots. In addition, the movement of the robot with respect to other machinery requires a high level of reliability in order to avoid collisions and damage to

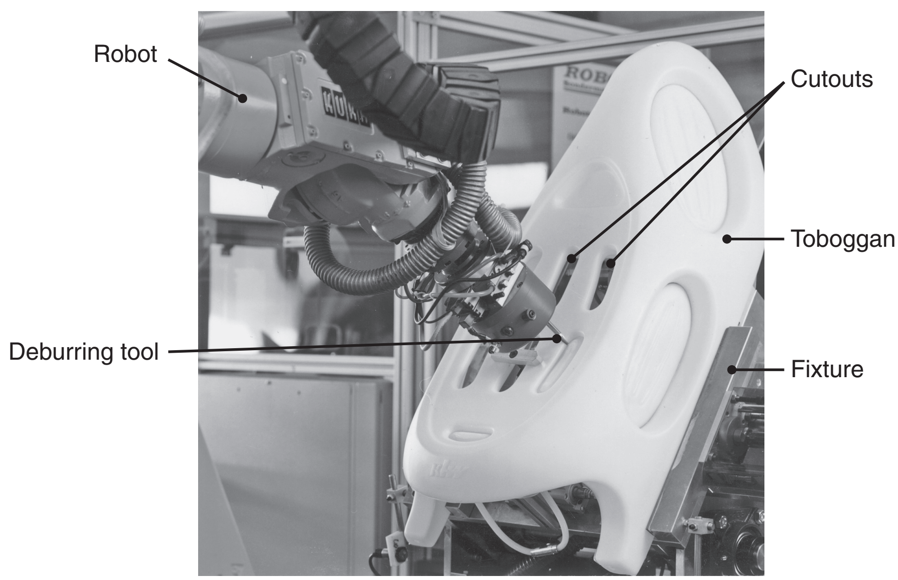


Figure 37.24: Robotic Deburring of a Blow-Molded Toboggan. *Source:* Courtesy of Kuka Robotics, Inc., and Roboter Technologie, GmbH. (See also Fig. 26.35.)

nearby equipment. The robot's material-handling activities require the proper securing of raw materials and parts in its gripper at various stages in the production line.

Cobots. Robots are typically programmed to perform a certain task without human guidance and are then isolated behind a fence or guard for the protection of operating personnel. *Cobots*, or *collaborative robots*, are designed to interact with plant personnel and have sensors to aide interaction with humans. Most applications to date have been in assembly or in materials handling; depending on the definition used, modern automobiles may indeed be considered to be cobots.

Case Study 37.1 Robotic Deburring of a Blow-molded Toboggan

Roboter Technologie, of Basel, Switzerland, produces high-quality toboggans and car seats made of plastic by injection molding (Section 19.3) or by blow molding (Section 19.4). After molding and while the part is cooling, holes have to be cut and deburred (Section 26.8). The deburring operation is ideal for a robot to perform but it is very difficult to automate. If a rotating burr tool is used, fumes and particles are generated that pose a health risk, and a nonrotating cutter requires that the robot allow deviations in its programmed path in order to accommodate shrinkage variations in the molded parts.

Roboter Technologie found a solution: a float-mounted tool (see *compliant end effectors*, Section 37.6.1) that can accommodate various cutter blades. To remove burrs, the blade must maintain the correct cutting angle and a constant cutting force. This task is achieved using a KUKA KR-15 robot, which performs the cutting and deburring operations in a single step while also compensating for the plastic shrinkage (Fig. 37.24).

As soon as the blow-molded part leaves the molding machine, an operator removes the flash and places the part on a rotary indexing table (Fig. 37.4b). When a part is rotated into the robot's workspace, the robot cuts and deburrs the holes; once the side parts have been cut and deburred, the fixture tilts the toboggan to a vertical position so that its top can be accessed. An automatic tool changer (Section 25.2) is used during the processing of each toboggan to switch from a convex cutter (used for the cutouts) to a straightedge cutter in order to produce a smooth exterior contour.

The complex shape of a toboggan is a good example of the flexibility achievable by robots. The robot completes the machining operation in 40 to 50 seconds, as opposed to the cycle time of the blow molding process of 120 seconds; thus, the implementation of the robot is consistent with the goals of a pull system (see Section 39.6). The robot was successfully used in a hazardous and dirty environment and it eliminated tasks that were associated with work-related injuries, incurred by fume exposure, and stresses on the wrist in manual deburring. Moreover, because of the higher quality of deburring and the elimination of rejected parts, the robot cell paid for itself within only three months.

Source: Courtesy of Kuka Robotics, Inc. and Roboter Technologie, GmbH.

37.7 Sensor Technology

A *sensor* is a device that produces a signal in response to its detecting or measuring a specific property, such as position, force, torque, pressure, chemistry, temperature, humidity, speed, acceleration, or vibration. Traditionally, actuators and switches have been used to set limits on the performance of machines, such as (a) stops on machine tools to restrict worktable movements, (b) pressure and temperature gages with automatic shutoff features, and (c) governors on engines to prevent excessive speed of operation.

Sensor technology is an important aspect of modern manufacturing, and is essential for data acquisition, monitoring, communication, and computer control. Low-cost sensors, often based on *flexible electronics* (Section 28.15), are being increasingly applied to process monitoring and control (Fig. 37.25). The

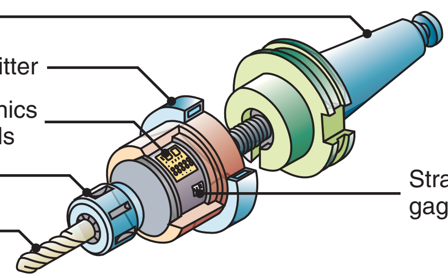


Figure 37.25: A toolholder equipped with thrust-force and torque sensors (smart toolholder), capable of continuously monitoring the cutting operation; such toolholders are essential for the adaptive control of manufacturing operations.

widespread use of sensors and the large amount of data they generate is the foundation for the application of **Big Data** and the **Internet of Things** approaches in manufacturing (see Section 39.8.1).

37.7.1 Sensor Classification

Sensors that are of greatest interest in manufacturing operations are generally classified as:

- **Mechanical** sensors measure such quantities as strain, mass, position, shape, velocity, force, torque, pressure, and vibration.
- **Chemical** sensors measure concentrations of target chemicals. A common example is the O_2 sensor in an automobile, which measures the amount of oxygen gas in the exhaust; this allows inference of the proper amounts of gas and fuel intake in order to minimize harmful emissions.
- **Electrical** sensors measure voltage, current, charge, and electrical conductivity.
- **Magnetic** sensors measure magnetic field, flux, and permeability.
- **Thermal** sensors measure temperature, flux, thermal conductivity, and specific heat.
- Other types of sensors are **acoustic, ultrasonic, optical, radiation, laser, and fiber optic**.

Depending on its application, a sensor may consist of metallic, nonmetallic, organic, or inorganic materials, as well as fluids, gases, plasmas, or semiconductors. Using the special characteristics of these materials, sensors convert the quantity or property measured to analog or digital output. The operation of an ordinary mercury thermometer, for example, is based on the difference between the thermal expansion of mercury and that of glass. Likewise, a machine part, a physical obstruction, or a barrier in a space can be detected by breaking the beam of light sensed by a photoelectric cell. A *proximity sensor*, which senses and measures the distance between it and an object or a moving member of a machine, can be based on acoustics, magnetism, capacitance, or optics. Other types of sensors contact the object and then take appropriate action, usually by electromechanical means.

Tactile Sensing. *Tactile* sensing involves the continuous sensing of variable contact forces, commonly by an array of sensors. Such a system is capable of operating within an arbitrary three-dimensional space. Fragile parts, such as eggs, thin glass objects, and electronic devices, can be handled by robots with *compliant (smart) end effectors*. These effectors can sense the force applied to the object being handled, using, for example, strain gages, piezoelectric devices, magnetic induction, ultrasonics, and optical systems of fiber optics and light-emitting diodes (LEDs).

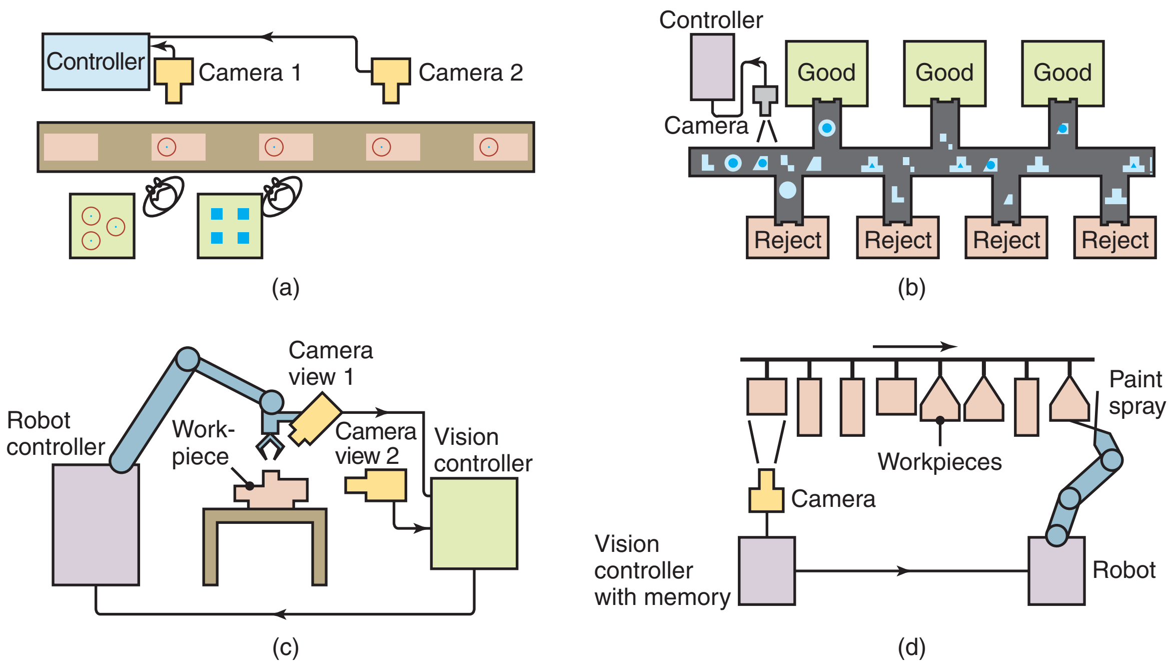


Figure 37.26: Examples of machine-vision applications. (a) In-line inspection of parts. (b) Identification of parts with various shapes, and inspection and rejection of defective parts. (c) Use of cameras to provide positional input to a robot relative to the workpiece. (d) Painting parts having different shapes by means of input from a camera. The system's memory allows the robot to identify the particular shape to be painted and to proceed with the correct movements of a paint spray attached to the end effector.

The force that is sensed is monitored and controlled through closed-loop feedback devices. Compliant grippers with force-feedback capabilities and sensory perception are complex and require powerful computers. *Anthropomorphic end effectors* are designed to simulate the human hand and fingers and to have the capability of sensing touch, force, and movement. The ideal tactile sensor must also sense *slip*, a normal capability of human fingers and hands; note, for example, how, even with closed eyes, one can sense if an object is beginning to slip from one's hand.

Visual Sensing. In visual sensing, cameras scan an image, and the software processes the data. Its most important applications in manufacturing are *pattern recognition* (Fig. 37.26), edge detection, and the transfer of information, such as with simple barcodes.

Machine vision commonly uses digital cameras that communicate with a computer through wireless, USB, or Ethernet connections. Scanning can take place in (a) one (line scan, such as with bar codes), (b) two (2D scanning, as with QR barcodes), or (c) three dimensions (3D scanning, as with CT scanning or confocal cameras, as shown in Fig. 37.27). Three-dimensional scanning has become more common, with powerful software available to allow simple digital cameras on smart phones to take three-dimensional data. Most manufacturing applications require 2D scanning, although 3D scanners are useful for geometry capture.

Machine vision is particularly suitable for parts with inaccessible features, in hostile manufacturing environments, measuring a large number of small features, and in situations where physical contact with the part may cause damage to the part. Sensors for machine tools can now sense tool breakage, verify part placement and fixturing, and monitor surface finish.

Figure 37.27: The use of a 3D scanner to digitize the geometry of a stamped part to create a 3D data file describing the geometry. The data file can then be used for quality control, or it can be additively manufactured or machined to produce a part with the same geometry or a mold for stamping. *Source:* Courtesy of Creaform.

Applications. Several applications of machine vision in manufacturing are shown in Fig. 37.26. With visual sensing capabilities, end effectors are capable of picking up parts and gripping them in the proper orientation and location. Machine vision is capable of in-line identification and inspection of parts and of rejecting defective ones. *Robust sensors* have been developed to withstand extremes of temperature, shock and vibration, humidity, corrosion, dust and various contaminants, fluids, electromagnetic radiation, and other interferences.

The *selection* of a sensor for a particular application depends on such factors as

- The particular quantity to be measured or sensed
- The sensor's interaction with other components in the system
- The sensor's expected service life
- Level of sophistication
- Difficulties associated with its use
- The power source
- Cost.

Smart Sensors. These sensors have the capability to perform a logic function, conduct two-way communication, make decisions, and take appropriate actions. The knowledge required to make a decision can be built into a smart sensor; in machining, for example, a computer chip with sensors can be programmed to turn a machine tool off when a cutting tool breaks. Likewise, a smart sensor can stop a mobile cobot, or a robot arm, from accidentally coming into contact with an object or people by sensing quantities such as distance, heat, and noise.

37.7.2 Sensor Fusion

Sensor fusion basically involves the integration of several sensors in such a manner that individual data from each of the sensors (such as force, vibration, temperature, and dimension data) are combined to provide a higher level of information and reliability. A simple application of sensor fusion can be demonstrated when someone drinks from a cup of hot coffee. Although such a common event is often taken for granted, it can readily be seen that drinking from a cup involves data input from eyes, lips, tongue, and hands. Through the five basic senses of sight, hearing, smell, taste, and touch, there is now real-time monitoring of relative movements, positions, and temperatures. If the coffee is too hot, for example, the hand movement of the cup toward the lip is slowed and adjusted accordingly. Note that the fingers and the hand also already sense the temperature, thus becoming an input into the control system.

The earliest engineering applications of sensor fusion were in robot movement control, missile flight tracking, and similar military applications, primarily because these activities involve movements that mimic human behavior. An important aspect in sensor fusion is *sensor validation*: the failure of any one sensor is detected, so that the control system maintains high reliability; thus it is essential to receive redundant data from different sensors.

Sensor fusion has become practical and available at relatively low cost, largely because of the advances made in sensor size, quality, and technology, as well as continued developments in computer-control systems, artificial intelligence, expert systems, and artificial neural networks, all described in Chapter 39.

37.8 Flexible Fixturing

In describing *work-holding devices* for the manufacturing operations throughout this book, the words *clamp, jig,* and *fixture* often were used interchangeably and sometimes in pairs, such as in jigs and fixtures. Briefly,

- **Clamps** are simple multifunctional work-holding devices
- **Jigs** have various reference surfaces and points for accurate alignment of parts or tools
- **Fixtures** generally are designed for specific applications.

Other common work-holding devices are chucks, collets, and mandrels. Some work-holding devices, such as power chucks, are designed and operated at various levels of mechanization and automation, and are driven by mechanical, hydraulic, or electrical means. Work-holding devices generally have specific ranges of capacity. For example, (a) a particular collet can accommodate bars only within a certain range of diameters; (b) four-jaw chucks can accommodate square or prismatic workpieces having a certain range of dimensions, (c) *dedicated fixtures* are designed and made for specific workpiece shapes and dimensions and for specific tasks; and (d) if a workpiece has curved surfaces, the contacting surfaces of the jaws are shaped by machining them to conform to the workpiece surfaces, known as *machinable jaws*.

The emergence of flexible manufacturing systems has necessitated the design and use of work-holding devices and fixtures that have *built-in flexibility*. There are several methods of **flexible fixturing**, based on different principles and called **intelligent fixturing systems**. These devices are capable of quickly accommodating a range of part shapes and dimensions, without requiring extensive changes, adjustments, or operator intervention.

Modular Fixturing. Modular fixturing is often used for small or moderate lot sizes, especially when the cost of dedicated fixtures and the time required to produce them would be difficult to justify. Complex workpieces can be located within machines using fixtures produced quickly from standard components, and can then be disassembled when a production run is completed. Modular fixtures usually are based on tooling plates or blocks configured with grid holes or T-slots upon which a fixture is constructed.

To quickly produce a fixture, several standard components, such as locating pins, workpiece supports, V-blocks, clamps, springs, and adjustable stops can be mounted onto a *base plate* or *block*. Using computer-aided fixture-planning techniques, such fixtures can be designed, assembled, and modified using industrial robots. Compared with dedicated fixturing, modular fixturing has been shown to be low in cost, have a shorter lead time, have more easily repaired components, and possess more intrinsic flexibility of application.

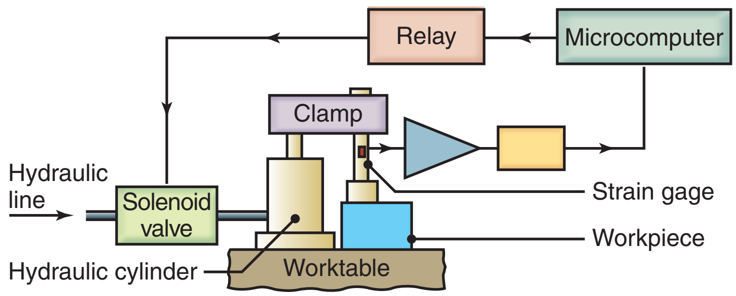


Figure 37.28: Schematic illustration of an adjustable-force clamping system; the clamping force is sensed by the strain gage, and the system automatically adjusts this force. *Source:* After P.K. Wright.

Tombstone Fixtures. Also referred to as *pedestal-type fixtures*, *tombstone fixtures* have between two and six vertical faces (hence resembling tombstones) onto which workpieces can be mounted. These fixtures are typically used in automated or robot-assisted manufacturing. The machine tool performs the desired operations on one face of the workpiece, then flips or rotates the fixture to work on other surfaces. Tombstone fixtures are commonly used for higher volume production, typically in the automotive industry (see also Case Study 24.1).

Bed-of-nails Device. This type of fixture consists of a series of *air-actuated pins* that conform to the shape of the external surfaces of the workpiece. Each pin moves as necessary to conform to the shape at its point of contact with the piece; the pins are then mechanically locked against the part. The device is compact, has high stiffness, and is reconfigurable.

Adjustable-force Clamping. A schematic illustration of this type of system is shown in Fig. 37.28. The strain gage mounted on the clamp senses the magnitude of the clamping force; the system then adjusts this force to keep the workpiece securely clamped for the particular application. It can also prevent excessive clamping forces that otherwise may damage the workpiece surface, particularly if it is soft or has a slender design.

Phase-change Materials. Other than by hard tooling, there are two methods capable of holding irregular-shaped or curved workpieces:

1. A *low-melting-point metal* is used as the clamping medium. Typically, an irregular-shaped workpiece is dipped into molten lead and allowed to set (e.g., the wooden stick in a popsicle), a process that is similar to *insert molding* (Section 19.3). After setting, the solidified lead block is clamped in a simple fixture. However, the possibly adverse effect of such materials as lead on the workpiece to be clamped (due to *liquid–metal embrittlement*; see Section 1.5.2) must be considered.

2. The supporting medium is a *magnetorheological* (MR) or *electrorheological* (ER) fluid. In the MR method, the particles are ferromagnetic or paramagnetic particles of micrometer size in a nonmagnetic fluid; surfactants are added to prevent the particles from settling. After the workpiece is immersed in the fluid, an external magnetic field is applied, whereby the particles are polarized and the behavior of the fluid changes from a liquid to that of a solid. The workpiece is later retrieved by removing the external magnetic field. In the ER method, the fluid is a suspension of fine dielectric particles in a liquid of low dielectric constant. Upon application of an electrical field, the liquid becomes a solid.

37.9 Assembly Systems

The individual parts and components produced by various manufacturing processes must be *assembled* into finished products. The total assembly operation is usually broken into individual assembly operations (*subassemblies*), with an operator assigned to carry out each step. Traditionally, assembly has involved much manual work and thus has contributed significantly to product cost.

Depending on the type of product, assembly costs can vary widely. For example, Apple iPhones cost \$12.50 to \$30.00 in total labor, with a total cost of \$450 to \$1000. The assembly cost for automobiles is around 10% of the sales cost. Assembly costs are generally 10-50% of the total cost of manufacturing, with the percentage of workers involved in assembly operations ranging from 20% to 60%. In developed countries with high productivity and associated automation, the number of workers involved in assembly is on the low end of this range; in countries with inexpensive labor, the percentage is higher. As production costs and quantities of products to be assembled began to increase, the necessity for *automated assembly* became obvious. Beginning with the hand assembly of muskets with *interchangeable parts* in the late 1700s and the early 1800s, assembly methods have vastly been improved upon over the years.

The first large-scale efficient application was the assembly of flywheel magnetos for the Model T Ford automobile. This experience eventually led to mass production of the automobile. The choice of an assembly method and system depends on the required production rate, the total quantity to be produced, the product's life cycle, the availability of labor, and cost.

Automated Assembly. Recall that parts are manufactured within certain dimensional tolerance ranges. Taking ball bearings as an example, it is well known that, although they all have the same *nominal* dimensions, some balls in a lot will be smaller than others, although by a very small amount. Likewise, some bearing races will be smaller than others in the lot. There are two methods of assembly for such high-volume products.

In **random assembly**, the components are put together by selecting them randomly from the lots produced. In **selective assembly**, the balls and races are segregated by groups of sizes, from smallest to largest. The parts are then selected to mate properly. Thus, the smallest diameter balls are mated with inner races having the largest outside diameter and, likewise, with outer races having the smallest inside diameters.

Methods and Systems of Assembly. There are three basic methods of assembly: manual, high-speed automatic, and robotic; they can be used individually or, as is the case in most applications, in combination. As shown in Fig. 37.29, an analysis of the product design must first be made to determine an appropriate and economical method of assembly:

1. **Manual assembly** uses relatively simple tools and generally is economical for small lots. Because of the dexterity of the human hand and fingers, and their capability for feedback through various senses, workers can manually assemble even complex parts without much difficulty. In spite of the use of sophisticated mechanisms, robots, and computer controls, however, aligning and placing of

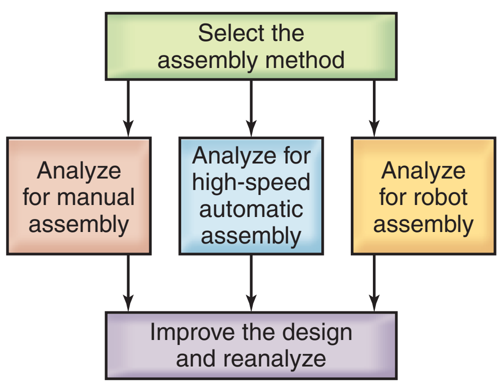


Figure 37.29: Stages in the design for assembly analysis. *Source:* After G. Boothroyd and P. Dewhurst.

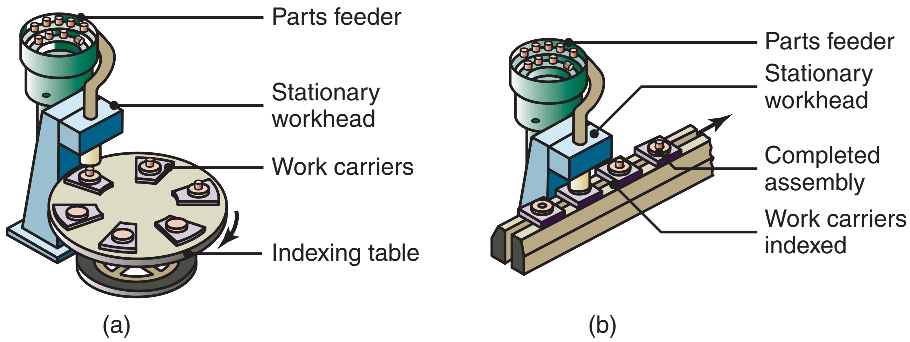


Figure 37.30: Transfer systems for automated assembly: (a) rotary indexing machine and (b) in-line indexing machine. *Source:* After G. Boothroyd.

a simple square peg into a square hole with small clearances can be a difficult task in automated assembly.

2. **High-speed automated assembly** utilizes *transfer mechanisms* designed specially for assembly. Two examples are shown in Fig. 37.30, in which individual assembly is carried out on products that are *indexed* for proper positioning.

3. In **robotic assembly**, one or more general-purpose robots operate at a single workstation (Fig. 37.31) or they operate at a multistation assembly system.

There are three basic types of assembly systems:

1. **Synchronous systems.** In these *indexing* systems, individual parts are supplied and assembled at a constant rate at fixed individual stations. The rate of movement of the parts in this system is based on the station that takes the longest time to complete its portion of the assembly. The synchronous system is used primarily for high-volume, high-speed assembly of small products.

   Transfer systems move the partial assemblies from workstation to workstation by various mechanical means; two typical transfer systems (*rotary indexing* and *in-line indexing*) are shown in Fig. 37.30. These systems can operate in either a fully automatic or a semiautomatic mode. Note, however, that a breakdown of one station will shut down the whole assembly operation.

   Part feeders supply the individual parts to be assembled and place them on other components, which are mounted on work carriers or fixtures. The feeders move the individual parts by vibratory or other means through delivery chutes and ensure their proper orientation by various ingenious means, some of which are shown in Fig. 37.32. Orienting parts properly and avoiding jamming are essential in all automated assembly operations.

2. **Nonsynchronous systems.** Each station operates independently, and any imbalance is accommodated in *buffer* (storage) between stations. The station continues operating until the next buffer is full or the previous buffer is empty. Also, if one station becomes inoperative, the assembly line continues to operate until all the parts in the buffer have been used up. Nonsynchronous systems are suitable for large assemblies with many parts to be assembled. Note that, if the times required for the individual assembly operations vary significantly, the output will be constrained by the slowest station.

3. **Continuous systems.** The product is assembled while moving at a constant speed on pallets or similar workpiece carriers. The components to be assembled are brought to the product by various means, and their movements are synchronized with the continuous movement of the product. Typical applications of this system are in bottling and packaging plants, although the method also has been used on mass-production lines for automobiles and appliances.

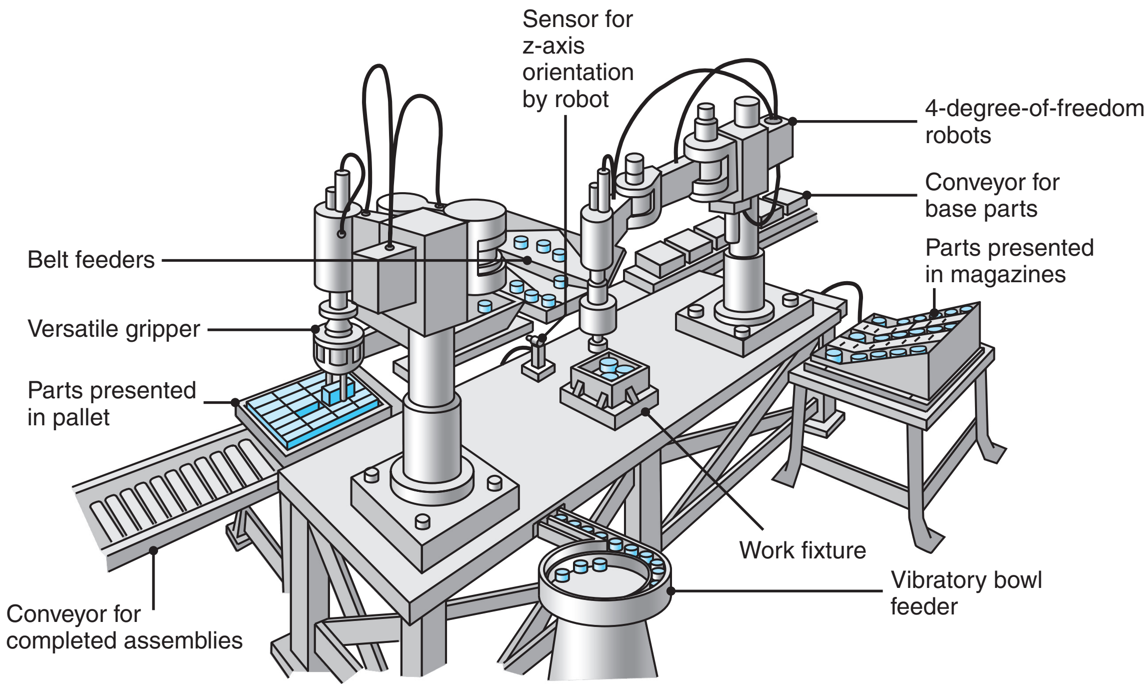


Figure 37.31: A two-arm robot assembly station. *Source:* After G. Boothroyd and P. Dewhurst.

Flexible Assembly Systems. Assembly systems generally are set up for a specific product line. They can, however, be modified for increased flexibility in order to assemble product lines that have a variety of product models. *Flexible assembly systems* (FAS) utilizes computer controls, interchangeable and programmable workheads and feeding devices, coded pallets, and automated guiding devices. This system is capable of, for example, assembling up to a dozen different transmission and engine combinations and power steering and air-conditioning units.

37.10 Design Considerations for Fixturing, Assembly, Disassembly, and Servicing

As in many aspects of production, design of the devices and systems described above is an integral part of the total manufacturing operation.

37.10.1 Design for Fixturing

The proper design, construction, and operation of flexible work-holding devices and fixtures are essential to the efficient operation of advanced manufacturing systems. The major design issues involved are the following:

- Work-holding devices must position the workpiece automatically and accurately. They must maintain its location precisely and with sufficient clamping force to withstand the requirements for a particular manufacturing operation. Fixtures also should be able to accommodate parts repeatedly in the same position.

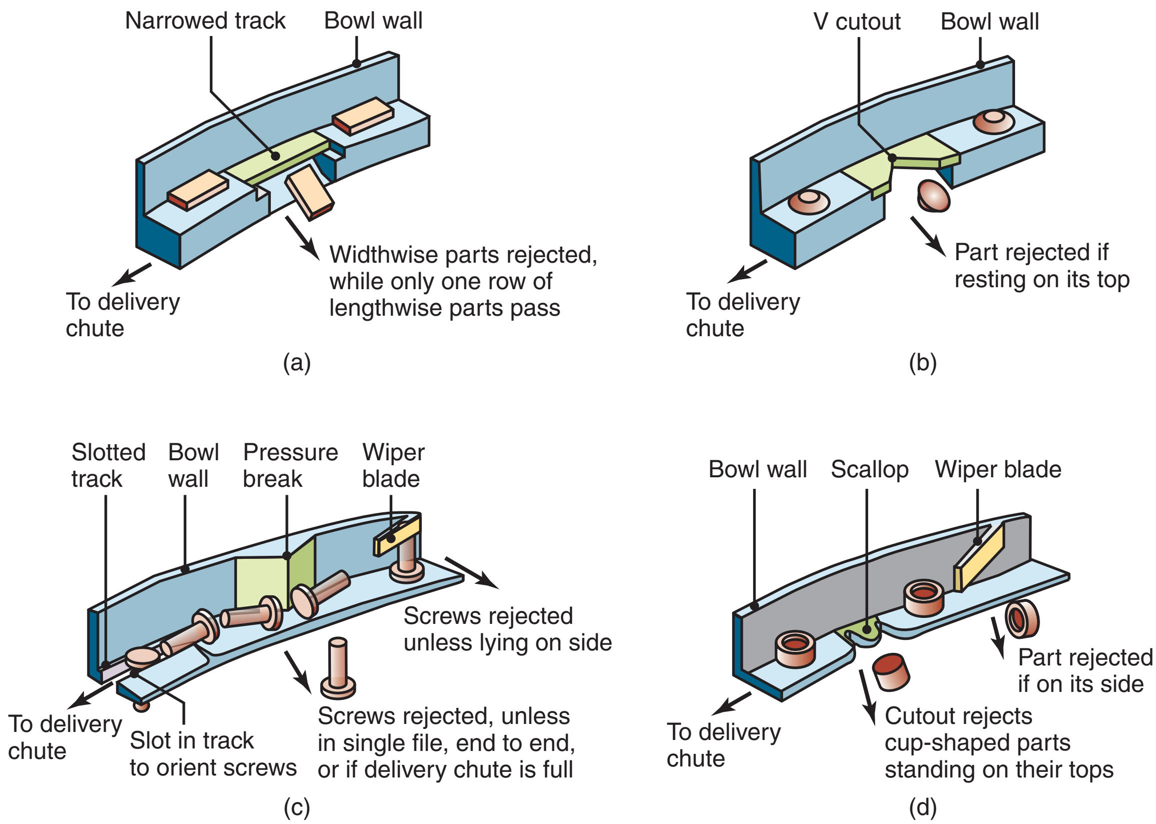


Figure 37.32: Examples of guides to ensure that parts are properly oriented for automated assembly. *Source:* After G. Boothroyd.

- Fixtures must have sufficient stiffness to resist, without excessive distortion, the normal and shear forces developed at the workpiece–fixture interfaces.
- The presence of loose machining or grinding chips and various other debris between the locating surfaces of the workpiece and the fixture can be a serious problem. Chips are most likely to be present where cutting fluids are used, as they tend to adhere to the wet surfaces due to surface-tension forces.
- A flexible fixture must be able to accommodate the parts to be made by different processes and ones with dimensions and surface features that vary from part to part. These considerations are even more important when the workpiece (a) is fragile or made of a brittle material; (b) is made of a relatively soft and flexible material, such as thermoplastics and elastomers; or (c) has a relatively soft coating on its contacting surfaces.
- Clamps and fixtures must have low profiles to avoid collision with cutting tools. Avoiding collisions is an important consideration in programming tool paths in machining operations.
- Flexible fixturing must meet special requirements in manufacturing cells and flexible manufacturing systems.
- Workpieces must be designed so as to allow locating and clamping within the fixture. Flanges, flats, or other locating surfaces should be incorporated into product design to simplify fixture design and to aid in part transfer into various machinery.

37.10.2 Design for Assembly, Disassembly, and Servicing

Design for Assembly. While product design for manufacture is *design for assembly* (DFA) has attracted special attention (particularly design for automated assembly), because of the continued need to reduce assembly costs. In *manual assembly*, a major advantage is that humans can easily pick the correct parts from bulk, such as from a nearby bin, orient them properly, and insert them as necessary. In *high-speed automated assembly*, however, automatic handling generally requires that parts be separated from the bulk, conveyed by hoppers or vibratory feeders (Fig. 37.32), and assembled in their proper locations and orientations.

The principle of **poka-yoke** has often been applied in assembly. Poka-yoke is a Japanese term that refers to "mistake-proofing" or "fail-safing." This method is also applied to lean manufacturing systems (Section 39.7). With respect to assembly, this principle suggests that assembly operations must be designed so that operator-caused errors are unlikely or even impossible to occur. This approach requires a review of assembly operations and identification of potential problems, as well as corrective actions to minimize assembly errors.

Some of the general guidelines for design for assembly may be summarized as:

1. Reduce the number and variety of parts in a product. Simplify the product design and incorporate multiple functions into a single part, and design parts for easy insertion. Use common parts as much as possible. Consider subassemblies that would serve as modules.

2. Ensure that parts have a high degree of symmetry, such as round or square, or a high degree of asymmetry, such as oval or rectangular, so that they cannot be installed incorrectly and do not require locating, aligning, or adjusting.

3. Designs should allow parts to be assembled without any obstructions. There should be a direct line of sight. Assemblies should not have to be turned over for insertion of components.

4. Consider methods such as snap fits (see Fig. 32.19) to avoid the need for fasteners such as bolts, nuts, and screws. If fasteners are used, their variety should be minimized and they should be spaced and located so that tools can be used without obstruction.

5. Part designs should consider such factors as size, shape, weight, flexibility, abrasiveness, and possible entanglement with other parts.

6. Assembly from two or more directions can be difficult. Parts should be inserted from a single direction, preferably vertically and from above in order to take advantage of gravity.

7. Products must be designed, or existing products redesigned, so that there are no physical obstructions to the free movement of parts during assembly. Sharp external and internal corners, for example, should be replaced with chamfers, tapers, or radii.

8. Color codes should be used on parts that may appear to be similar but are different. Letters or other symbols also can be provided to ensure proper part identification.

Robotic Assembly. Design guidelines for robotic assembly include the following additional considerations:

- Parts should be designed so that they can be gripped and manipulated by the same gripper of the robot. Parts should be made available to the gripper in the proper orientation.

- Assembly that involves threaded fasteners (bolts, nuts, and screws) may be difficult to perform by robots. One exception is the use of self-threading screws for sheet metal, plastics, and wooden parts. Note that robots easily can handle snap fits, rivets, welds, and adhesives. The advances in compliant end effectors and dexterous manipulators has made robotic assembly even more attractive.

Evaluating Assembly Efficiency. To evaluate *assembly efficiency*, each component of an assembly is evaluated with respect to its features that can affect both assembly itself and a baseline estimated time required to incorporate the part into the assembly. Note that assembly efficiency can also be measured for existing products. The assembly efficiency, η, is given by

$$\eta = \frac{Nt}{t_{\text{tot}}}, \tag{37.1}$$

where N is the number of parts, t_{tot} is the total assembly time, and t is the ideal assembly time for a small part that presents no difficulties in handling, orientation, or assembly; t is commonly taken to be 3 seconds. On the basis of Eq. (37.1), competing designs can thus be evaluated with respect to design for assembly. It has been observed that products that are in need of redesign to facilitate assembly usually have assembly efficiencies around 5–10%, while well-designed parts have efficiencies around 25%.

Design for Disassembly. The manner and ease with which a product may be taken apart for maintenance or replacement of its parts is another important consideration in product design. Recall, for example, the difficulties one has in removing certain components from under the hood of some automobiles; similar difficulties exist in the disassembly of appliances and numerous other products.

The general approach to *design for disassembly* requires the consideration of factors that are similar to those for design for assembly. Analysis of computer or physical models of products and their components with regard to disassembly can generally indicate any potential problems, such as obstructions, size of passageways, lack of a line of sight, and the difficulty of firmly gripping and guiding components.

An important aspect of design for disassembly is how, after its life cycle (see Section 40.4), a product is to be taken apart for recycling, especially to salvage its more valuable components. Note, for example, that depending on their design and location, the type of tools used, and whether manual or power tools are used (a) rivets will take longer to remove than screws or snap fits and (b) a bonded layer of valuable material on a component would be very difficult, if not impossible, to remove for recycling or reuse.

Obviously, the longer it takes to take components apart, the higher is the cost of doing so. It is then possible that this cost becomes prohibitive; consequently, the time required for disassembly also has to be studied and measured. Although the time depends on the manner in which disassembly is performed, some examples are: (a) cutting wire at 0.25 s, (b) disconnecting wire at 1.5 s, (c) effecting snap fits and clips at 1 to 3 s, and (d) loosening machine screws and bolts at 0.15 to 0.6 s per revolution. These quantities will of course greatly depend on the level of automation employed.

Design for Servicing. *Design for servicing* is essentially based on the concept that the elements that are most likely to need servicing are at the outer layers of the product. In this way, individual parts are easier to reach and service, without the need to remove various other parts in order to do so. Thus, designing for assembly and disassembly should take into account the ease with which a product can be serviced and, if necessary, repaired.

37.11 Economic Considerations

As described in greater detail in Chapter 40, and as seen throughout many chapters in this book, there are numerous considerations involved in determining the overall *economics of production*. Because all production systems are essentially combinations of machines and people, important factors influencing the final decisions include

- Type and cost of machinery, equipment, and tooling
- Cost of operation of the machinery
- Skill level and amount of labor required
- Production quantity desired.

Recall also that lot size and production rate greatly influence the economics of production. Small quantities per year can be produced in job shops. However, the type of machinery in job shops generally requires skilled labor and the production quantity and rate are low; as a result, the cost per part can be high. Similarly, rapid prototyping facilities can be used for low production runs, and can be more economical if material requirements are compatible with the processing sequence.

At the other extreme is the production of very large quantities, using conventional flow lines and transfer lines and involving special-purpose machinery and equipment, specialized tooling, and computer-control systems. Although all of these components constitute major investments, both the level of skill required and the labor costs are relatively low because of the high level of automation implemented. These production systems are, however, organized for a specific type of product and hence they lack flexibility.

Because most manufacturing operations are between the preceding two extremes, an appropriate decision must be made regarding the optimum level of automation to be implemented. In many situations, selective automation rather than total automation of a facility has been found to be cost effective.

Summary

- Automation has been implemented in manufacturing processes, material handling, inspection, assembly, and packaging at increasing rates. There are several levels of automation, ranging from simple automation of machines to untended manufacturing cells

- True automation began with the numerical control of machines, which offers flexibility of operation, lower cost, and ease of making different parts with less operator skill. Production quantity and rate are important factors in determining the economic levels of automation.

- Manufacturing operations are optimized further, both in quality and in cost, by adaptive control techniques, which continuously monitor an operation and quickly make necessary adjustments in the processing parameters.

- Major advances have been made in material handling, particularly with the implementation of industrial robots and automated guided vehicles.

- Sensors are essential in the implementation of modern technologies; a wide variety of sensors based on various principles have been developed and installed.

- Additional advances include flexible fixturing and automated assembly techniques that reduce the need for worker intervention and lower manufacturing costs. Their effective and economic implementation requires that design for assembly, disassembly, and servicing be recognized as an important factor in the total design and manufacturing operations.

- Efficient and economic implementation of these techniques also requires that design for assembly, disassembly, and servicing be recognized as important elements in manufacturing.

Key Terms

Adaptive control
Assembly
Automated guided vehicle
Automation
Buffer
Closed-loop control
Compliant end effectors
Computer numerical control
Computer vision
Continuous path
Contouring
Control systems

Dedicated machines
End effector
Feedback
Flexible assembly systems
Flexible fixturing
Hard automation
Hardwired controls
Industrial robot
Intelligent robot
Interpolation
Machine vision
Manipulators
Material handling
Mechanization
Numerical control
Open-loop control
Part programming
Poka-Yoke
Positioning
Power-head production units
Productivity
Programmable controller
Programming language
Random assembly
Repeat accuracy
Resolution
Robot
Selective assembly
Selective automation
Sensor fusion
Sensors
Smart sensors
Soft automation
Stand-alone machines
Tactile sensing
Tombstone fixture
Total productive maintenance
Transfer lines
Visual sensing
Work envelope

Bibliography

Batchelor, B.G. (ed.), **Machine Vision Handbook**, Springer, 2012.
Beyerer, J., and Leon, F.P., **Machine Vision: Automated Inspection**, Springer, 2015.
Blum, R.S., and Liu, Z., **Multi-Sensor Image Fusion and its Applications**, CRC, 2005.
Bomberg, A., **Handbook of Machine and Computer Vision**, 2nd ed., Wiley, 2017.
Booothroyd, G., Dewhurst, P., and Knight, W.A., **Product Design for Manufacture and Assembly**, 3rd ed., 2010.
Batchelor, B.G., **Handbook of Machine Vision**, Springer, 2012.
Corke, P., **Robotics, Vision and Control**, Springer, 2013.
Craig, J.J., **Introduction to Robotics: Mechanics and Control**, 4th ed., Prentice Hall, 2017.
Davies, E.R., **Computer and Machine Vision**, 4th ed., Academic Press, 2012.
Fraden, J., **Handbook of Modern Sensors: Physics, Designs, and Applications**, 5th ed., Springer, 2015.
Hornberg, A., **Handbook of Machine Vision**, Wiley, 2007.
Kandray, D., **Programmable Automation Technologies: An Introduction to CNC, Robotics and PLCs**, Industrial Press, 2010.
Kurfess, T.R. (ed.), **Robotics and Automation Handbook**, CRC Press, 2004.
Niku, S.B., **Introduction to Robotics, Analyses, Control, Applications**, Wiley, 2011.
Quesada, R., **Computer Numerical Control Machining and Turning Centers**, Prentice Hall, 2004.
Rekiek, B., and Delchambre, A., **Assembly Line Design**, Springer, 2005.
Ripka, P., and Tipek, A., **Modern Sensors Handbook**, ISTE Publishing Co., 2007.
Smid, P., **CNC Programming Handbook,** 2nd ed., Industrial Press, 2002.
Snyder, W.E., and Qi, H., **Machine Vision**, Cambridge University Press, 2010.

Soloman, S., **Sensors Handbook**, McGraw-Hill, 2009.
Stenerson, J., and Curran, K.S., **Computer Numerical Control: Operation and Programming**, 3rd ed., Prentice Hall, 2006.
Ulsoy, A.G., Peng, H., and Cakmakci, M., **Adaptive Control Systems**, Cambridge University Press, 2012.
Valentino, J.V., and Goldenberg, J., **Introduction to Computer Numerical Control**, 5th ed., Prentice Hall, 2012.
Zhang, J., **Practical Adaptive Control: Theory and Applications**, VDM Verlag, 2008.

Review Questions

37.1. Describe the differences between mechanization and automation.

37.2. Explain the difference between hard and soft automation. Why are they so called?

37.3. What is productivity? Why is it important?

37.4. Explain the difference between a flexible manufacturing line and a transfer line.

37.5. Describe the principle of numerical control of machines.

37.6. Explain open-loop and closed-loop control circuits.

37.7. Describe the principle and purposes of adaptive control.

37.8. What factors have led to the development of automated guided vehicles?

37.9. What is a point-to-point control system? How is it different from a contouring system?

37.10. Describe the features of an industrial robot. Why are these features necessary?

37.11. List and describe the principles of various types of sensors.

37.12. Describe the concept of design for assembly. Why has it become an important factor in manufacturing?

37.13. Is it possible to have partial automation in assembly? Explain.

37.14. Explain the advantages of flexible fixturing.

37.15. How are robots programmed to follow a certain path?

37.16. What kind of end effectors are available?

37.17. What is a cobot? How is it different from a robot?

Qualitative Problems

37.18. Why is automation generally regarded as evolutionary rather than revolutionary?

37.19. Explain why it would be difficult to justify automation for small production runs.

37.20. Are there activities in manufacturing operations that cannot be automated? Explain.

37.21. What is a programmable logic controller? Why are they popular?

37.22. Explain the factors that have led to the development of numerical control.

37.23. Giving specific examples, discuss your observations concerning Fig. 37.2.

37.24. What are the relative advantages and limitations of the two arrangements for power heads shown in Fig. 37.4?

37.25. Discuss methods of online gaging of workpiece diameters in turning operations other than that shown in Fig. 37.15.

37.26. Are drilling and punching the only applications for the point-to-point system shown in Fig. 37.10a? Explain.

37.27. If three points were known on a flat line, is it better to use linear or circular interpolation? Explain.

37.28. What determines the number of robots in an automated assembly line such as that shown in Fig. 37.22?

37.29. Describe situations in which the shape and size of the work envelope of a robot (Fig. 37.21) can be critical.

37.30. Explain the difference between an automated guided vehicle and a self-guided vehicle.

37.31. What are the two kinds of robot joints? Give applications for each.

37.32. Explain why sensors have become so essential in the development of automated manufacturing systems.

37.33. Table 37.2 shows a few examples of typical products for each category. Add several other examples to the table.

37.34. List applications for industrial robots.

37.35. What is meant by the term *sensor fusion*?

37.36. Describe applications of machine vision for specific parts that are similar to the examples shown in Fig. 37.26.

37.37. What is a tombstone fixture?

37.38. Sketch the workspace (envelope) of each of the robots shown in Fig. 37.20.

37.39. List the advantages and disadvantages of modular filtering.

Quantitative Problems

37.40. A spindle–bracket assembly uses the following parts: a steel spindle, two nylon bushings, a stamped steel bracket, and six screws and six nuts to attach the nylon bushings to the steel bracket and thereby support the spindle. Compare this assembly with the spindle–bracket assembly shown in Problem 16.73, and estimate the assembly efficiency for each design.

37.41. Disassemble a simple ballpoint pen. Carefully measure the time it took for you to reassemble the pen, and calculate the assembly efficiency. Repeat the exercise for a mechanical pencil.

37.42. Examine Fig. 37.11b, and obtain an expression for the maximum error in approximating a circle with linear increments as a function of the radius of the circle and the number of increments on the circumference of the circle.

37.43. Develop open- and closed-loop control system equations in order to have the worktable achieve a sinusoidal position given by $x = \sin \omega t$, where t is time.

37.44. A 35 kg worktable is controlled by a motor/gear combination that can develop a maximum force of 500 N. It is desired to move the worktable from it's current location ($x = 0$) to a new location ($x = 100$ mm). Plot the resultant force and position as a function of time for (a) an open-loop control system and (b) a closed-loop control system for $k_p = 5$ N/mm and $k_v = 4.5$ Ns/mm.

37.45. Assume that you are asked to give a quiz to students on the contents of this chapter. Prepare five quantitative problems and five qualitative questions, and supply the answers.

Synthesis, Design, and Projects

37.46. Refer to Part III of this book, and give an example of a metal-forming operation that is suitable for adaptive control.

37.47. Describe possible applications for industrial robots not discussed in this chapter.

37.48. Design two different systems of mechanical grippers for two widely different applications.

37.49. Give some applications for the systems shown in Fig. 37.26a and c.

37.50. For a system similar to that shown in Fig. 37.28, design a flexible fixturing setup for a lathe chuck.

37.51. Give examples of products that are suitable for the three types of production shown in Fig. 37.3.

37.52. Describe situations in which tactile sensors would not be suitable. Explain why.

37.53. Are there situations in which machine vision cannot be applied properly and reliably? Explain.

37.54. Choose one machine each from Parts II through IV, and design a system in which sensor fusion can be used effectively.

37.55. Think of a product, and design a transfer line for it which is similar to that shown in Fig. 37.5. Specify the types and the number of machines required.

37.56. Describe your thoughts on the usefulness and applications of modular fixturing consisting of various individual clamps, pins, supports, and attachments mounted on a base plate.

37.57. Inspect several household products and describe the manner in which they have been assembled. Comment on any product design changes you would make so that assembly, disassembly, and servicing are simpler and faster.

37.58. Inspect Table 37.1 on the history of automation, and describe your thoughts as to what new developments might be added to the bottom of the list in the near future.

37.59. Design a robot gripper that will pick up and place the following: (a) eggs, (b) an object made of foam rubber, (c) a metal ball with a very smooth and polished surface, (d) a newspaper, and (e) tableware, such as knives, spoons, and forks.

37.60. Design an end effector that can pick up and place (a) an egg, (b) a marshmallow in the shape of an egg, (c) a steel casting in the shape of an egg.

37.61. Review the specifications of various numerical-control machines, and make a list of typical numbers for their (a) positioning accuracy, (b) repeat accuracy, and (c) resolution. Comment on your observations.

37.62. Describe the sensors that you use in a simple act such as walking or throwing a ball.

37.63. Give an example of each of the design rules in Section 37.10.2.

37.64. Obtain an old toaster and disassemble it. Explain how you would go about reassembling it by automated assembly.

37.65. Conduct an Internet search and obtain product literature for four kinds of 3D scanners. Compare the price and capabilities of each.

37.66. Assume that you are asked to give a quiz to students on the contents of this chapter. Prepare five quantitative problems and five qualitative questions, and supply the answers.

37.67. Design a guide that operates in the same manner as those shown in Fig. 37.32, but to align U-shaped parts so that they are inserted with the open end down.

Chapter 38

Computer-aided Manufacturing

- Computers have fundamentally and pervasively changed the product design and manufacturing enterprise; powerful computer software is now available to assist and integrate all engineering tasks.
- This chapter opens with a description of computer-aided design, in which the graphic description of parts is created and stored in software.
- The use of computers in the direct control of manufacturing processes and in computer-aided manufacturing is then discussed.
- The chapter then describes how software can allow the simulation of manufacturing processes and systems.
- Finally, a description of group technology is presented an approach that is often built into CAD software, allowing the rapid recovery of previous design and manufacturing experience, and an essential tool for production flow analysis.

38.1 Introduction

The importance of product quality was emphasized in Chapter 36, along with the necessity for the commitment of a company to total quality management. Recall also the statements that *quality must be built into the product*, that high quality does not necessarily mean high costs, and that marketing poor-quality products can indeed be very costly to the manufacturer.

High quality is far more attainable and less expensive if design and manufacturing activities are properly integrated, rather than treated as separate activities. Integration can be performed successfully and effectively through *computer-aided design*, *engineering, manufacturing, process planning*, and *simulation of processes and systems*, as described throughout this chapter. The widespread availability of high-speed computers and powerful software has allowed computers to proliferate into all areas of manufacturing.

Computer technology is pervasive and exists at many levels. A part geometry can be programmed in CAD software, which is in itself a fairly complex computer program. The manufacture of a part can be achieved, for example, by programming it into G-code (Section 37.3.6), which uses another fairly complex computer program to translate geometric instructions into machine actions. Software is currently available and sufficiently powerful to integrate design with CNC programming activities and, thereby, streamline the design and manufacturing process. Indeed, every aspect of the modern manufacturing enterprise is currently associated with computers and software, and integration of the entire business through communication standards is now possible.

38.2 Manufacturing Systems

Manufacturing is a complex system, because it consists of many diverse physical and human elements. Some of these elements are difficult to predict and control due to such factors as the supply and cost of raw materials, the impact of continually developing technologies, global market changes, and human behavior and performance. Ideally, a manufacturing system should be represented by mathematical and physical models that show the nature and extent of the interdependence of all relevant variables. In this way, the effects of a change or a disturbance that occurs anywhere in the system can be analyzed and necessary, timely adjustments can be made.

The supply of a particular raw material may, for example, decrease significantly due to global demands or for geopolitical reasons. Because the raw material cost will rise as a result (supply and demand), alternative materials have to be considered and selected. The selection must be made after a careful consideration of several factors, because such a change may have adverse effects on product quality, production rate, and manufacturing costs. The material selected, for example, may not be as easy to form, machine, or weld, and thus product integrity may suffer.

In a constantly changing global marketplace, the demand for a product also may fluctuate randomly and rapidly for a variety of reasons. As examples, note the downsizing of automobiles in response to rising fuel costs and the increasing popularity of gas–electric hybrids, fuel cells, and electrically-powered vehicles. The manufacturing system must be able to produce the modified product in a relatively short lead time while minimizing large expenditures in new machinery and tooling that will be required. Lead time is defined as the length of time between the creation of the product as a concept, or receipt of an order for a product, and the time that the product first becomes available in the marketplace.

Such a complex system can be difficult to analyze and model, largely because of a lack of comprehensive and reliable data on all of the variables involved. Moreover, it is difficult to correctly predict and control some of these variables, because (a) raw-material costs are difficult to predict accurately; (b) machine-tool characteristics, their performance, and their response to random external disturbances cannot be precisely modeled; and (c) human behavior and performance are even more difficult to model.

38.3 Computer-integrated Manufacturing

Computer-integrated manufacturing (CIM) involves the computerized integration of all aspects of product design, process planning, production, and distribution, as well as the management and operation of the whole manufacturing organization. The effectiveness of CIM greatly depends **integrated communications system**, involving computers, machines, equipment, and their controls (see Section 39.7). Because CIM ideally should involve the total operation of an organization, it requires an extensive *database* concerning the technical and business aspects of the operation. Consequently, if planned all at once, CIM can be prohibitively expensive, particularly for small and medium-size companies.

Implementation of CIM in existing manufacturing plants may begin with the use of modules in *selected phases* of a company's operation. For new plants, comprehensive and long-range strategic planning, covering all phases of the operation, is essential. Such planning and the level of integration must take into account considerations such as (a) the mission, goals, and culture of the organization; (b) the availability of financial, technical, and human resources; and (c) the existing, as well as emerging, technologies in the areas of the products to be manufactured.

Subsystems. Computer-integrated manufacturing systems comprise the following *subsystems*, which are integrated into a whole (Fig. 38.1):

1. Business planning and support
2. Product design
3. Manufacturing process planning
4. Process automation and control
5. Production-monitoring systems.

The subsystems are designed, developed, and implemented in such a manner that the output of one subsystem serves as the input of another. Organizationally, the subsystems generally are divided into two functions:

- **Business-planning functions:** Forecasting, scheduling, material-requirements planning, invoicing, and accounting.
- **Business-execution functions:** Production and process control, material handling, testing, and inspection of the system.

If implemented properly, the major benefits of CIM are the following:

- Emphasis on *product quality* and uniformity, through better process control
- *Efficient* use of materials, machinery, and personnel and a major reduction in work-in-progress inventory, all of which improve productivity and lower product cost.
- *Total control* of the production, schedules, and management of the entire manufacturing operation
- *Responsiveness* to shorter product life cycles, changing market demands, and global competition.

38.3.1 Databases

An effective CIM system requires a single, large database that is shared by the entire organization. Databases consist of up-to-date, detailed, and accurate information relating to designs, products, processes, materials, machinery, production, finances, purchasing, sales, and marketing. This vast array of information

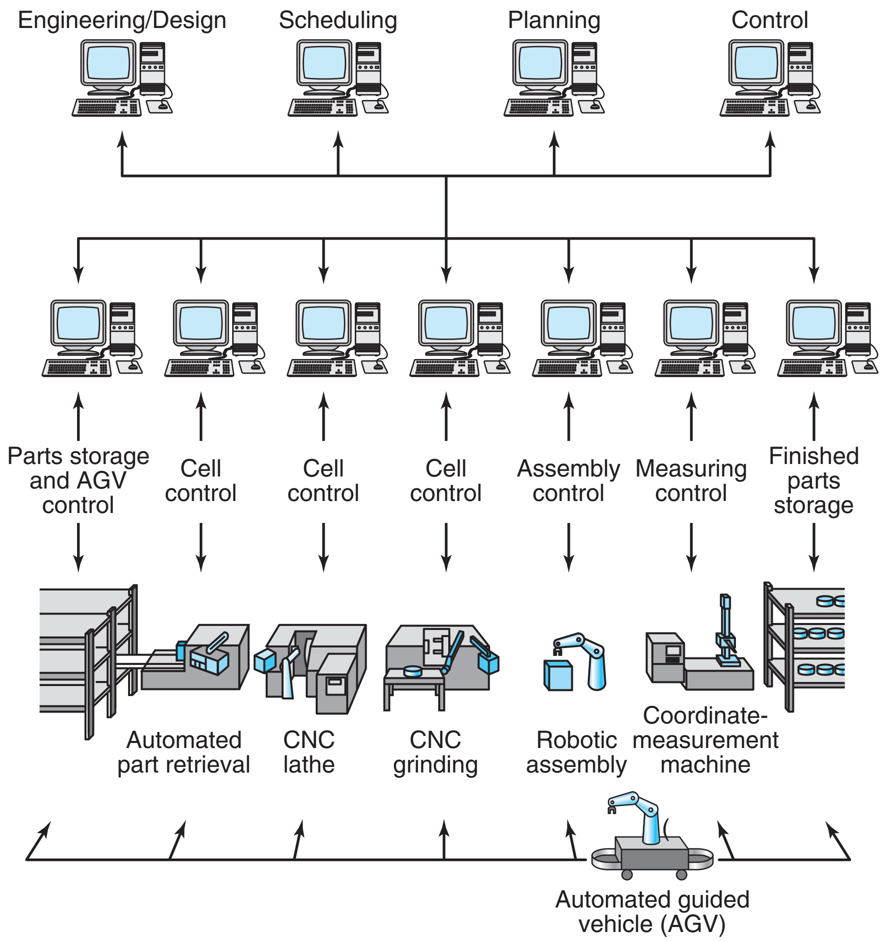


Figure 38.1: A schematic illustration of a computer-integrated manufacturing system. The manufacturing cells and their controls shown at the lower left are described in Section 39.2. *Source:* After U. Rembold.

is stored in computer memory and recalled or modified as necessary, either by individuals in the organization or by the CIM system itself. Data storage on remote servers using Internet tools for data retrieval is referred to as *cloud storage*, and is increasingly popular for the large databases typical of large organizations. However, cybersecurity and protection of the data remains a serious concern.

A database typically consists of the following items, some of which are classified as technical and others as nontechnical:

- **Product data:** Part shape, dimensions, and specifications
- **Data-management attributes:** Part number and revision level, including descriptions or keywords to assist in retrieving data
- **Production data:** Manufacturing processes employed
- **Operational data:** Scheduling, lot sizes, and assembly requirements
- **Resources data:** Capital, machines, equipment, tooling, personnel, and their capabilities.

Databases are compiled by individuals in the organization, with input from various sensors in production machinery and equipment. Data are automatically collected by a **data-acquisition system** (DAS), which can track the number of parts being produced per unit of time and their dimensional accuracy, surface finish, weight, and other characteristics at specified rates of sampling. The components of DAS include microprocessors, transducers, and analog-to-digital converters (ADC). Data-acquisition systems also are capable of analyzing data and transferring them to other computers for such purposes as statistical analysis, data presentation, and the forecasting of product demand.

Several factors are important in the use and implementation of databases:

1. They should be timely, accurate, easily accessible, easily shared, and user friendly.

2. Because they are used for a variety of purposes and by many people in an organization, databases must be flexible and responsive to the needs of different users.

3. CIM systems can be accessed by designers, manufacturing engineers, process planners, financial officers, and the management of the company through appropriate access codes; companies must protect data against tampering or unauthorized use.

4. If problems arise with data accuracy or loss of data, the correct data should be recovered and restored.

38.4 Computer-aided Design and Engineering

Computer-aided design (CAD) involves the use of software to create design drawings and product models (see also Fig. I.11 in the General Introduction). CAD is generally associated with *interactive computer graphics*, known as a *CAD system*. *Computer-aided engineering* (CAE) simplifies the creation of the database by allowing several applications to share the information in the database. These applications include, for example, (a) finite-element analysis of stresses, strains, deflections, and temperature distribution in structures and load-bearing members; (b) the generation, storage, and retrieval of NC data; and (c) the design of integrated fixtures and machinery used to produce the design.

In CAD, the user can generate drawings or sections of a drawing on a computer. The design can be printed if desired, but usually it is stored as a digital file to be accessed as needed, often on networked computers throughout the organization. When using a CAD system, the designer can conceptualize the object to be designed, can consider alternative designs, or quickly modify a particular design to meet specific requirements.

There are several powerful commercially available programs to aid designers in geometry description and engineering analysis, such as SolidWorks, CATIA, AutoCAD, Solid Edge, and VectorWorks. The software can help identify potential problems, such as excessive loads, deflections, or interference at mating surfaces when encountered during assembly. Information, such as a list of materials, specifications, and manufacturing instructions, also is stored in the CAD database. Using this information, the product designer can then analyze the manufacturing economics of alternative designs.

Modern CAD software allows the use of parametric design, as described in Section 38.4.2. In a parametric design, instead of specifying dimensions explicitly, relations between dimensions are set, so that a different sized part of the same design can easily be generated.

38.4.1 Exchange Specifications

Because of the availability of a wide variety of CAD systems with different characteristics and supplied by different vendors, effective communication and exchange of data between these systems is essential. **Drawing exchange format** (DFX) was developed for use with *Autodesk*® and is still supported, but has been superseded by the DWG file format. Stereolithography (STL) formats are used to export three-dimensional geometries, initially only to *rapid-prototyping systems* (Chapter 20), but they now have become a format for data exchange between different CAD systems.

The necessity for a single, neutral format for better compatibility and for the transfer of more information than geometry alone is currently filled mainly by the *Initial Graphics Exchange Specification* (*IGES*). This format is used for translation in two directions (in and out of a system), and is also widely used for the translation of three-dimensional line and surface data. There are several variations of IGES in existence; the latest is version 5.3, published in 1996.

Another useful format is a solid-model-based standard, called the *Product Data Exchange Specification* (*PDES*), which is based on the Standard for the Exchange of Product model data (STEP) and developed by the International Standards Organization. PDES allows information on shape, design, manufacturing, quality assurance, testing, maintenance, etc., to be transferred between CAD systems. The increasing popularity of PDES and STEP have led to less use of IGES.

38.4.2 Elements of CAD Systems

The design process in a CAD system consists of four stages: Geometric modeling, design analysis and optimization, design review and optimization, and database.

Geometric Modeling. In *geometric modeling*, a physical object or any of its parts is described mathematically. The designer first constructs a geometric model by giving commands that create or modify lines, surfaces, solids, dimensions, and text. Together, these elements present an accurate and complete two- or three-dimensional representation of the object. The results are displayed and can be moved around on the screen, and any section can be magnified to view details.

The models in a CAD system can be presented in three ways:

1. In *line representation*, also called **wire-frame** representation (Fig. 38.2), all of the *edges* of the model are visible as solid lines. This image can, however, be ambiguous or difficult to visualize, particularly for complex shapes.

   The three types of wire-frame representations are *two, two-and-one-half,* and *three dimensional*. A two-dimensional image shows the profile of the object, and a two-and-one-half-dimensional image can be obtained by a *translational sweep* (by moving the two-dimensional object along the z-axis). For round objects, a two-and-one-half-dimensional model can be generated simply by *rotating* a two-dimensional model around its axis.

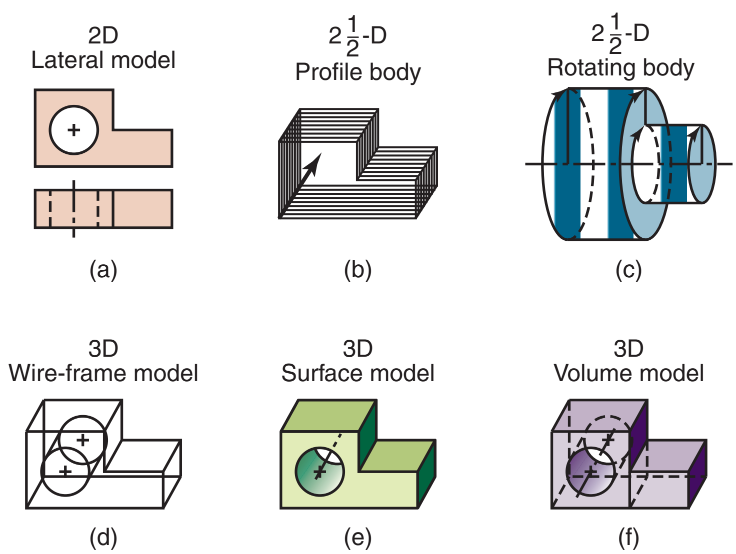


Figure 38.2: Various types of modeling for CAD.

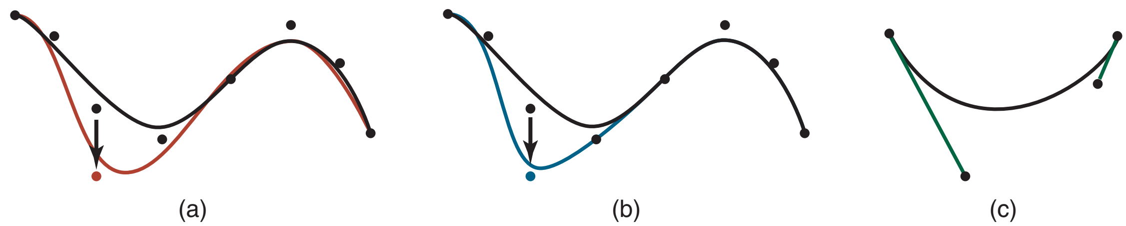


Figure 38.3: Types of splines. (a) A Bezier curve passes through the first and last control point, but it generates a curve from the other points; changing a control point modifies the entire curve. (b) A B-spline is constructed piecewise, so that changing a vertex affects the curve only in the vicinity of the changed control point. (c) A third-order (cubic) piecewise Bezier curve is constructed through two adjacent control points, with two other control points defining the slope of the curve at the endpoints. A third-order piecewise Bezier curve is continuous, but its slope may be discontinuous.

2. In the **surface model**, all visible *surfaces* are shown. These models define surface features and edges of objects. CAD programs now use *Bezier* curves, B-splines, or nonuniform rational B-splines (NURBS) for surface modeling. Each of these approaches uses control points to define a polynomial curve or surface. A Bezier curve passes through the first and last vertex and uses the other control points to generate a blended curve. The drawback to Bezier curves is that any modification of one control point will affect the entire curve.

   B-splines are blended piecewise polynomial curves where the modification of a control point affects only the curve in the area of the modification. Figure 38.3 shows examples of two-dimensional Bezier curves and B-splines. A *NURBS* is a special type of B-spline such that each control point has a weight associated with it.

3. In the **solid model**, all surfaces are shown, but the data describes the interior volume. Solid models can be constructed from (a) *swept volumes* (Fig. 38.2b and c) or by the techniques shown in Fig. 38.4; (b) *boundary representation* (B-rep), where surfaces are combined to develop a solid model (Fig. 38.4a); and (c) *constructive solid geometry* (CSG), where simple shapes such as spheres, cubes, blocks, cylinders, and cones (called *primitives of solids*) are combined to develop a solid model (Fig. 38.4b).

The standard for rapid prototyping machinery, the **STL file format** (an abbreviation for *Standard Tessellation Language* but often associated with stereolithography for which it was developed), allows for three-dimensional part descriptions. Basically, an STL file consists of a number of triangles that define the exterior surface (Fig. 38.5). With a sufficiently large number of triangles, the surface can be defined within a prescribed tolerance, although with requiring a larger file size. With additive manufacturing, a part cross-section can be obtained at any height, and the resulting polygon is then used to plan the part (see Fig. 20.3). The use of STL in rapid prototyping, along with its easy implementation, has led to the use of this format in other applications as well, such as computer graphics and general CAD data transfer. A number of new file formats, such as 3MF, are under development, which in general migrate away from triangular tessellations and include other information, such as color and surface texture. The 3MF format also allows the transfer of additional data, including proprietary data, within the same CAD file.

A special kind of solid model is a **parametric model**, where a part is stored not only in terms of a B-rep or CSG definition, but is derived from the dimensions and constraints that define the features (Fig. 38.6). Whenever a change is made, the part is re-created from these definitions, a feature that allows for simple and straightforward updates and changes to be made to the models.

Design Analysis and Optimization. After the geometric features of a particular design have been determined, the design is subjected to engineering analysis. This phase may, for example, consist of analyzing stresses, strains, deflections, vibrations, heat transfer, temperature distribution, or dimensional tolerances.

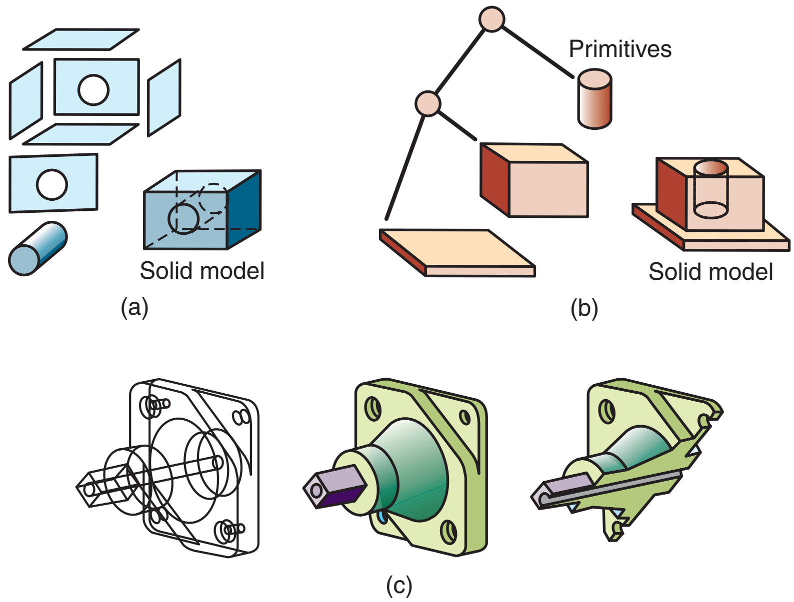


Figure 38.4: (a) Boundary representation of solids, showing the enclosing surfaces of the solid model and the generated solid model. (b) A solid model represented as compositions of solid primitives. (c) Three representations of the same part by CAD. *Source:* After P. Ranky.

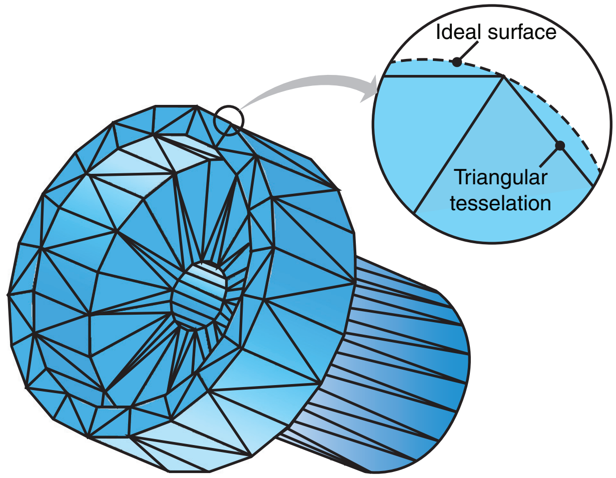


Figure 38.5: An example of an STL part description. Note that the surface is defined by a tessellation of triangles, and that there is an inherent error of form that occurs with curved surfaces; however, this can be brought to any desired tolerance by incorporating more triangles in the surface.

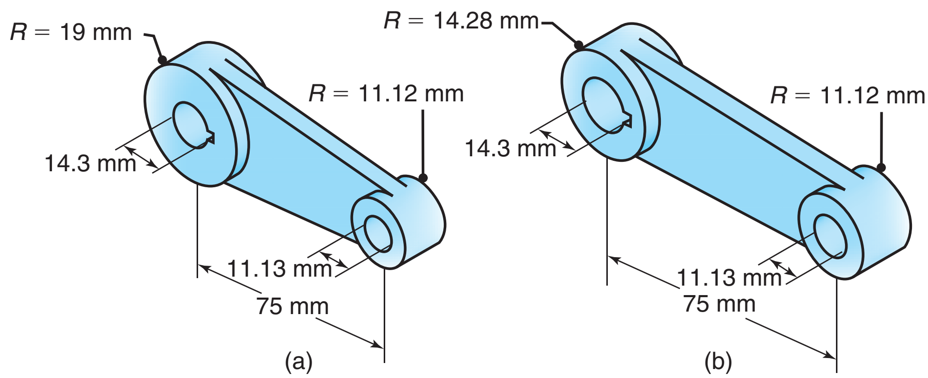


Figure 38.6: An example of parametric design; the dimensions of part features can be modified easily to quickly obtain an updated solid model.

Various software packages are now available, such as the finite element-based programs ABAQUS, ANSYS, NASTRAN, LS-DYNA, MARQ, and ALGOR, each having the capabilities to compute these quantities accurately and rapidly.

Because of the relative ease with which such analyses can be carried out, designers increasingly are willing to analyze a design more thoroughly before it is moved on to production. Experiments and measurements in the field nonetheless may be necessary to determine the actual effects of loads, temperature, and various other variables on the designed components.

Design Review and Evaluation. An important design stage is the design review and evaluation in order to check for any interference or excess gap between various components. The review is done in order to avoid difficulties either during assembly or in the use of the part and to determine whether moving members, such as linkages, are going to operate as intended. Software is available with animation capabilities to identify potential problems with moving members and other dynamic situations. During this stage, the part is dimensioned and toleranced precisely to the full degree required for manufacturing it.

Database. Many components, such as bolts and gears, either are standard components mass produced according to a given design specification or are identical to the parts used in previous designs. CAD systems thus have a built-in database management system that allows designers to locate, view, and adopt parts from a stock part library. These parts can be modeled parametrically to allow cost-effective updating of the part geometry. Some databases are available commercially, with extensive parts libraries; many vendors make their part libraries also available on the Internet.

38.5 Computer-aided Manufacturing

Computer-aided manufacturing (CAM) involves the use of computers to assist in all phases of manufacturing a product; it encompasses many of the technologies described in Chapter 37 and in this chapter. Because of their joint benefits, CAD and CAM are often combined into *CAD/ CAM systems*. This combination allows the transfer of information from the design stage to the stage of planning for manufacture, without the necessity to reenter the data on part geometry manually. The database developed during CAD is stored and further processed by CAM into the relevant data and instructions for such purposes as operating and controlling production machinery, material-handling equipment, and automated testing and inspection for product quality. CAD/CAM systems also are capable of coding and classifying parts into groups that have similar design or manufacturing attributes, as described in Section 38.8.3.

Typical applications of CAD/CAM include the following:

- Programming for numerical control and industrial robots
- Design of dies and molds for casting in which, for example, shrinkage allowances are preprogrammed
- Dies for metalworking operations, such as complex dies for sheet forming and progressive dies for stamping
- Design of tooling and fixtures and EDM electrodes
- Quality control and inspection, such as coordinate-measuring machines programmed on a CAD/CAM workstation
- Process planning and scheduling
- Plant layout.

An important feature of CAD/CAM in machining operations is the capability to calculate and describe the *tool path* (see Figs. 23.11, 23.12, 24.2, 25.9, 26.12, and 26.20). The instructions (*programs*) are computer generated, and they can be modified by the programmer to optimize the tool path. The engineer or technician can then display and visually check the tool path for possible tool collisions with clamps, fixtures, or other interferences.

By standardizing product development and reducing design effort, tryout, and prototype work, CAD/CAM has made possible significantly reduced manufacturing costs and improved productivity. The two-engine Boeing 777 passenger airplane, for example, was designed completely by computer (known as **paperless design**), with 2000 workstations linked to eight computers. The plane was constructed directly from the CAD/CAM software that was developed (an enhanced CATIA system), and no prototypes or mock-ups were built, as were required for previous models. The cost for this development was on the order of $6 billion.

38.6 Computer-aided Process Planning

Process planning is basically concerned with selecting methods of production: tooling, fixtures, machinery, sequences of operations, and assembly; all of these diverse activities must be planned, which traditionally has been done by process planners. The sequence of processes and operations to be performed, the machines to be used, the standard time for each operation, and similar information all are documented in a computer file. **Routing sheets** (Fig. 38.7) are the traditional means for storing manufacturing data, and are useful to demonstrate the type of data required.

Computer-aided process planning (CAPP) accomplishes the complex task of process planning by viewing the total operation as an *integrated* system, so that the individual processing steps are coordinated and performed efficiently and reliably. CAPP is particularly effective in small-volume, high-variety parts production. Although extensive software and good coordination with CAD/CAM and other aspects of integrated manufacturing systems (described throughout the rest of this chapter) are required, CAPP is a powerful tool for efficiently planning and scheduling manufacturing operations.

38.6.1 Elements of CAPP Systems

There are two types of computer-aided process-planning systems:

Variant System. Also called the *derivative system*, these computer files contain a standard process plan for the part to be produced. On the basis of its shape and manufacturing characteristics, a search for a standard plan is then conducted in the database, using a specific code number for the part. The plan is retrieved, displayed for review, and printed as a routing sheet.

ROUTING SHEET		
CUSTOMER'S NAME: Midwest Valve Co.		PART NAME: Valve body
QUANTITY: 15		PART NO.: 302
Operation No.	**Description of operation**	**Equipment**
10	Inspect forging, check hardness	Rockwell tester
20	Rough machine flanges	Lathe No. 5
30	Finish machine flanges	Lathe No. 5
40	Bore and counterbore hole	Boring mill No. 1
50	Turn internal grooves	Boring mill No. 1
60	Drill and tap holes	Drill press No. 2
70	Grind flange end faces	Grinder No. 2
80	Grind bore	Internal grinder No. 1
90	Clean	Vapor degreaser
100	Inspect	Ultrasonic tester

Figure 38.7: An example of a simple, traditional routing sheet. *Operation sheets* may include additional information on materials, tooling, the estimated time for each operation, processing parameters (such as cutting speeds and feeds), and other information. The routing sheet travels with the part from operation to operation. Current practice is to store all relevant data in computers and to affix to the part a bar code, radio frequency ID, or equivalent label that serves as a key into the database of parts information.

The *variant-process plan* includes such information as the types of tools and machines required, the sequence of operations to be performed, and the speeds, feeds, and time required for each sequence. Minor modifications of an existing process plan, which usually are necessary, also can be made. If the standard plan for a particular part is not in the computer files, a plan that is close to it and that has a similar code number and an existing routing sheet is retrieved. If a routing sheet does not exist for a new part, one is made and stored in computer memory.

Generative System. In this system, a process plan is automatically generated on the basis of the same logical procedures that would be followed by a traditional process planner in making that particular part. However, the generative system is complex because it must contain comprehensive and detailed information about (a) the part shape and dimensions; (b) process capabilities; (c) selection of manufacturing methods, machinery, and tools; and (d) the sequence of operations to be performed.

The generative system can create a new plan instead of having to use and modify an existing plan, as the variant system must do. Although generally used less commonly than the variant system, the generative system has such advantages as (a) flexibility and consistency in process planning for new parts and (b) higher overall planning quality, because of the capability of the decision logic in the system to optimize the planning and utilize up-to-date manufacturing technology.

The process-planning capabilities of computers also can be integrated into the planning and control of production systems. These activities are a subsystem of computer-integrated manufacturing, as described in Section 38.3. Several functions can be performed, such as **capacity planning** for plants to meet production schedules, control of inventory, purchasing, and production scheduling.

38.6.2 Material-requirements Planning and Manufacturing Resource Planning

Computer-based systems for managing inventories and delivery schedules of raw materials and tools are called *material-requirements planning* (MRP) systems. Also regarded as a method of *inventory control*, MRP involves the keeping of complete records of inventories of materials, supplies, parts in various stages of production (called *work in progress*, WIP), orders, purchasing, and scheduling. Several files of data usually are involved in a master production schedule. These files pertain to the raw materials required (listed on a **bill of materials**), product structure levels (individual items that compose a product, such as components, subassemblies, and assemblies), and scheduling.

Manufacturing resource planning (MRP-II) controls all aspects of manufacturing planning through feedback. Although the system is complex, it is capable of final production scheduling, monitoring actual results in terms of performance and output, and comparing those results against the master production schedule.

38.6.3 Enterprise Resource Planning

Enterprise resource planning (ERP) is basically an extension of MRP-II. Although there are several variations, it is also a method for effective planning and control of all the resources needed in a business enterprise to take orders for products, produce them, ship them to the customer, and service them. ERP thus attempts to coordinate, optimize, and dynamically integrate all information sources and the widely diverse technical and financial activities in a manufacturing organization.

Modern packages plan and manage availability and utilization of employees, machines, tools, fixtures, and many other resources in order to maximize shop-floor efficiency. Moreover, ERP systems are increasingly sophisticated regarding the management of global supply chains. It has therefore been suggested that a fully functional ERP will be able to activate a whole factory as soon as an order comes in to the company's website, thereby enabling *pull* (see Section 39.7). Every ERP is integrated with a **manufacturing execution system** (MES), which gives specific scheduling and routing instructions to the factory, organizes production, and even generates a CNC plan for each part.

Effective implementation of ERP can be a challenging task because of:

- The difficulties encountered in timely, effective, and reliable communication among all parties involved, especially in a global business enterprise.
- The need for changing and evolving business practices in an age where information systems and **e-commerce** (defined as buying and selling of products or services over electronic systems) have become highly relevant and important to the success of business organizations.
- The need to meet extensive and specific hardware and software requirements for ERP. ERP-II is a recent development that uses web-based tools to perform the tasks of ERP. These systems are intended to extend the ERP capabilities beyond the host organization, and allowing interaction and coordination across corporate entities.

38.7 Computer Simulation of Manufacturing Processes and Systems

With increasing sophistication of computer hardware and software, *computer simulation of manufacturing processes and systems* has advanced rapidly. Simulation takes two basic forms:

1. It is a model of a specific operation, intended to determine the viability of a process or to optimize and improve its performance.
2. It models multiple processes and their interactions, to help process planners and plant designers in the layout of machinery and facilities.

Individual processes can be modeled using various mathematical schemes (see, for example, Fig. 16.51). Typical problems addressed are (a) *process viability*, such as the formability of sheet metal in a certain die and

(b) *process optimization*, such as material flow in forging a given die to identify potential defects, or mold design in casting to eliminate hot spots, promote uniform cooling, and minimize defects. Finite-element analysis is increasingly being applied in software packages (called **process simulation**) that are available commercially and are inexpensive.

Mathematical models can take several forms. Finite element simulation can involve stresses and strains but ignores accelerations (*quasi-static*), or simulation can involve vibrations and chatter. Similarly, a *design of experiments* approach can be used to experimentally characterize a system. A current trend is to incorporate process simulations into machine control programs to monitor and, if necessary, modify operations in real time.

38.8 Group Technology

Group technology (GT) is a methodology that seeks to take advantage of the **design** and **processing similarities** among the parts to be produced. As illustrated in Fig. 38.8, these characteristics clearly suggest that major benefits can be obtained by **classifying** and **coding** the parts into *families*. One company found that by disassembling each product into its individual components and then identifying the similar parts, 90% of the 3000 parts made fell into only five major families of parts.

A pump, for example, can be broken down into its basic components, such as the motor, housing, shaft, flanges, and seals. It will be noted that each of these components in a pump is basically the same in terms of its design and manufacturing characteristics. Consequently, all shafts can be placed in one family of shafts or a family of seals, and so on. Group technology becomes especially attractive because of the ever-growing variety of products, which are often produced in batches, since nearly 75% of manufacturing today is batch production.

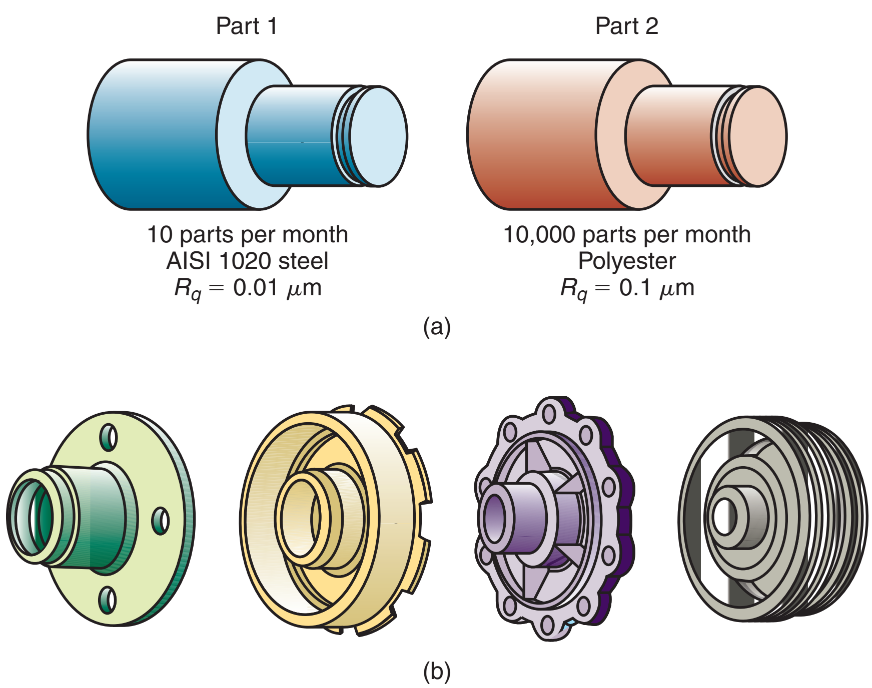
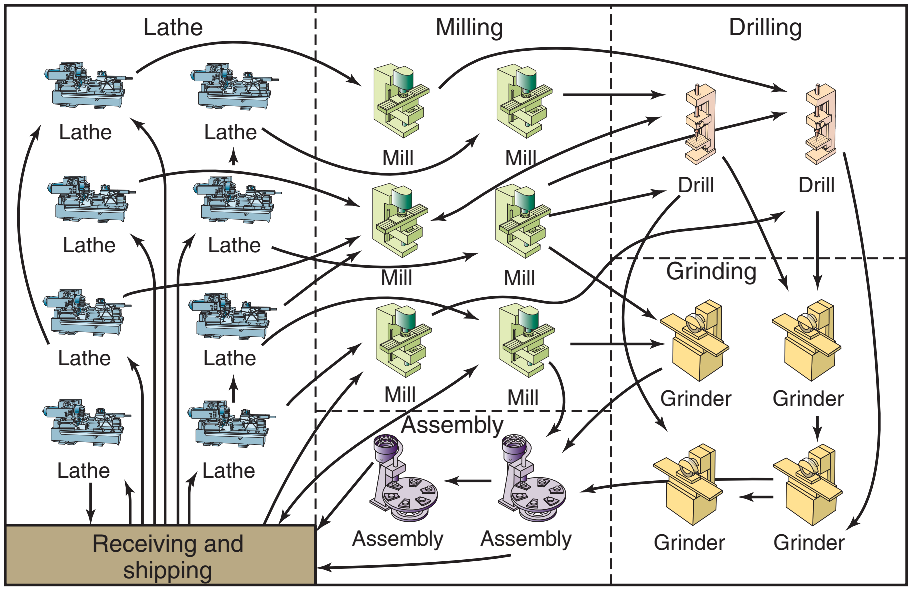
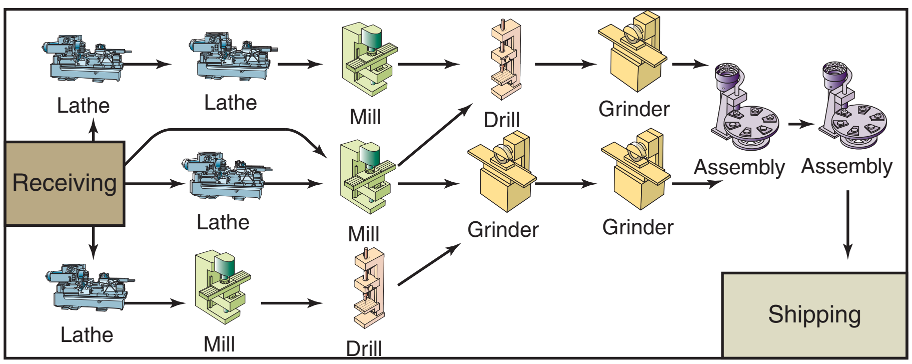


Figure 38.8: Grouping parts according to their (a) geometric similarities and (b) manufacturing attributes.



(a)



(b)

Figure 38.9: (a) Functional layout of machine tools in a traditional plant; arrows indicate the flow of materials and parts in various stages of completion. (b) Group-technology (cellular) layout. Legend: L = lathe, M = milling machine, D = drilling machine, G = grinding machine, A = assembly. *Source:* After M.P. Groover.

Plant Layout. A traditional product flow in batch manufacturing, called **functional layout**, is shown in Fig. 38.9a. Note that machines of the same type are arranged in groups, that is, groups of lathes, milling machines, drill presses, and grinders. It will be noted that in such a layout there is considerable random movement, as shown by the arrows indicating movement of materials and parts. Because it wastes time and effort, such an arrangement is obviously not efficient. A better product flow line that would take advantage of group technology is the **group layout**, shown in Fig. 38.9b (see also *cellular manufacturing*, Section 39.2). Note the greater simplicity and the decrease in the number of paths and movements among the machines.

38.8.1 Advantages of Group Technology

The major advantages of group technology are the following:

- It makes possible the standardization of part designs and the minimization of design duplication. New part designs can be developed from similar, yet previously used, designs, thus saving a significant amount of time and effort. Moreover, the product designer quickly can determine whether data on a similar part already exists in the computer files of the company.
- Data that reflect the experience of the designer and the manufacturing process planner are stored in the database; thus, a new and less experienced engineer can quickly benefit from that experience by retrieving any of the previous designs and process plans.
- Manufacturing costs can be more easily estimated and the relevant statistics on materials, processes, number of parts produced, and other factors can be more easily obtained.
- Process plans can be standardized and scheduled more efficiently, orders can be grouped for more efficient production, and thus machine utilization is improved. Setup times are reduced and parts are produced more efficiently, and with better and more consistent product quality. Similar tools, fixtures, and machinery are shared in the production of a family of parts.
- With the implementation of CAD/CAM, cellular manufacturing, and CIM, group technology is capable of greatly improving productivity and reducing costs, with batch production approaching the benefits of mass production. Depending on the level of implementation, savings in each of the various design and manufacturing phases can range from 5% to 75%.

38.8.2 Classification and Coding of Parts

In group technology, parts are identified and grouped into families by **classification and coding** (C/C) **systems**. This process is a critical and complex first step and is done according to the part's design attributes and manufacturing attributes (see Fig. 38.8).

Design Attributes. These attributes pertain to similarities in geometric features and consist of:

- External and internal shapes and dimensions
- Aspect ratios, such as length-to-width or length-to-diameter ratios
- Dimensional tolerances
- Surface finish
- Part functions.

Manufacturing Attributes. Group technology uses the similarities in the methods and sequence of the manufacturing operations performed on the part. Because the selection of a manufacturing process or processes depends on numerous factors, manufacturing and design attributes are interrelated. The manufacturing attributes of a part consist of:

- Primary processes
- Secondary and finishing processes
- Dimensional tolerances and surface finish

- Sequence of operations performed
- Tools, dies, fixtures, and machinery
- Production quantity and production rate.

Coding can be time consuming and considerable experience is required. It can be done simply by viewing the shapes of the parts in a generic way and then classifying the parts accordingly; for example, parts having rotational symmetry, parts having rectilinear shape, and parts having large surface-to-thickness ratios. A more thorough approach is to review all of the data and drawings concerning the design *and* manufacture of all of the parts.

Parts also may be classified by studying their production flow during the manufacturing cycle, an approach called **production flow analysis** (PFA). Recall from Section 38.6 that routing sheets clearly show process plans and the sequence of operations to be performed. One drawback to PFA, however, is that a particular routing sheet does not necessarily indicate that the total manufacturing operation is optimized.

38.8.3 Coding

The code for parts can be based on a company's own system of coding or it can be based on one of several classification and coding systems available in commercial software. Often, coding is incorporated into CAD/CAM packages. Whether it was developed in-house or it was purchased, the coding system must be compatible with the company's other systems, such as NC machinery and CAPP systems. The code structure for part families typically consists of numbers, letters, or a combination of the two. Each specific component of a product is assigned a code. This code may pertain to design attributes only (generally, fewer than 12 digits) or to manufacturing attributes only; most advanced systems include both, using as many as 30 digits. Coding may be done without input from the software user and displayed only if the information is requested. Commonly, design or manufacturing data retrieval can be based on keyword searches.

The three basic levels of coding vary in degree of complexity:

1. **Hierarchical coding.** Also called *monocode*, hierarchical coding interprets each succeeding digit on the basis of the value of the preceding digit. Each symbol amplifies the information contained in the preceding digit; therefore, a digit in the code cannot be interpreted alone. The advantage of this method is that a short code can contain a large amount of information; however, the method is difficult to apply in a computerized system.

2. **Polycodes.** In this method, also known as *chain-type* coding, each digit has its own interpretation, which does not depend on the preceding digit. This structure tends to be relatively long, but it allows the identification of specific part attributes and is well suited to computer implementation.

3. **Decision-tree coding.** This type of coding, also called *hybrid coding*, is the most advanced and combines both design and manufacturing attributes (Fig. 38.10).

38.8.4 Coding Systems

There are numerous coding systems in use, including Opitz, Multiclass, KK-3, Brisch, CODE, CUTPLAN, and DCLASS. Each has its particular strengths and more or less integration with commercial database and CAD software. Three of the most common industrial coding systems are the following:

1. The **Opitz system** (after H. Opitz, 1905–1977) was the first comprehensive coding system developed. It consists basically of nine digits (12345 6789) representing design and manufacturing data (Fig. 38.11); four additional codes (1234) may be used to identify the type and sequence of production operations

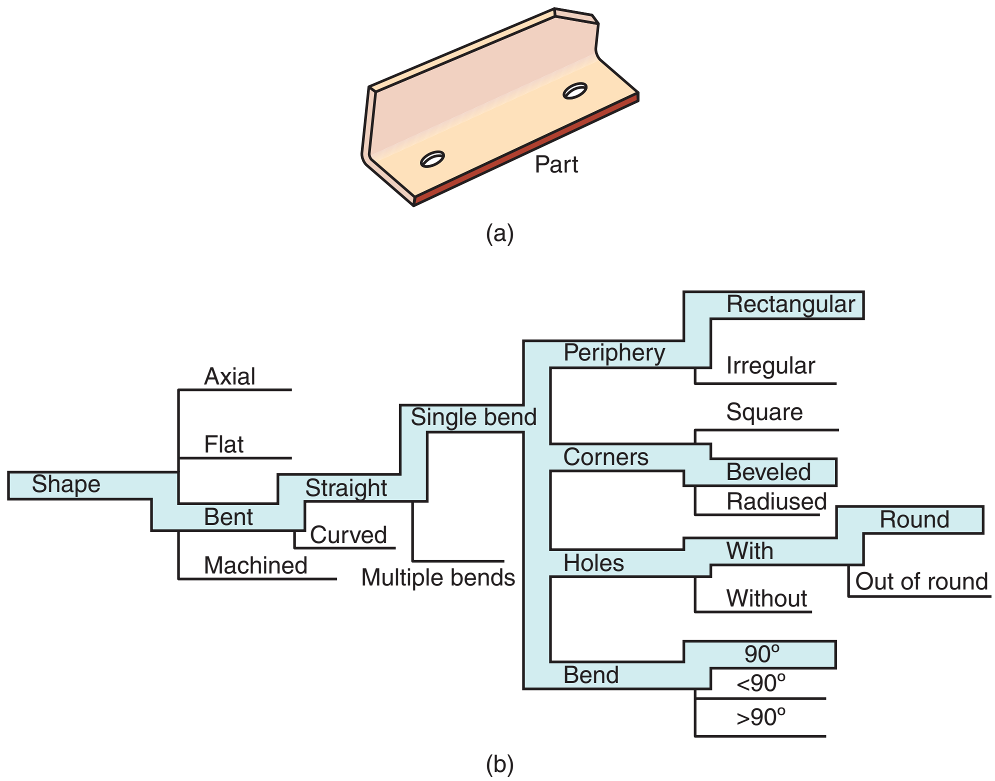


Figure 38.10: Decision-tree classification for a sheet-metal bracket. *Source:* After G.W. Millar.

to be performed. The Opitz system has two drawbacks: (a) It is possible to have different codes for parts that have similar manufacturing attributes and (b) a number of parts with different shapes can have the same code.

2. The *multiClass system* was developed to help automate and standardize several design, production, and management functions. It involves up to 30 digits (Fig. 38.12). This system is used interactively with a computer that asks the user a number of questions; on the basis of the answers given, the computer automatically assigns a code number to the part.

3. The *KK-3 system* is a general-purpose system for parts that are to be machined or ground; this code uses a 21-digit decimal system. The code is much greater in length than the two previous codes described, but it classifies dimensions and dimensional ratios, such as the length-to-diameter ratio of the part. The structure of a KK-3 system for rotational components is shown in Fig. 38.13.

Production Flow Analysis. One of the main benefits of GT is the design of manufacturing cells (Section 39.2). Consider the situation in Fig, 38.14, which is intended to show a highly simplified list of parts that are to be produced and the required machinery. As can be seen in this figure, the types of machines and variety of parts do not lend themselves to groupings of machines. By using GT, parts can be classified and codified, then sorted into logical groupings. Combined with production requirements, this approach allows the design of manufacturing cells to achieve the associated benefits of flexibility and utility.

For example, in Fig. 38.14a, a spreadsheet shows a collection of parts produced by a pump manufacturer; the machines required to produce these parts are also indicated. From this spreadsheet, no intuitive machine groupings are apparent. With a reorganization of the data, the parts suggest logical groupings of machines, as shown in Fig. 38.14b. It should be recognized that these figures are highly simplified in order to demonstrate the significance of group technology in organizing cells; usually there are many more parts and processes (including metrology) that must be considered.

		Form Code					Supplementary Code			
		1st digit Part class	2nd digit Main shape	3rd digit Rotational surface machining	4th digit Plane surface machining	5th digit Auxiliary holes, gear teeth, and forming	Digit 1	2	3	4
0	Rotational parts	$\frac{L}{D} < 0.5$	External shape, external shape elements	Internal shape, internal shape elements	Plane surface machining	Auxiliary holes	Dimension	Material	Original shape of raw material	Accuracy
1		$0.5 < \frac{L}{D} < 3$								
2		$\frac{L}{D} > 3$				Gear teeth				
3		$\frac{L}{D} < 2$ with deviation	Main shape	Rotational machining, internal and external shape elements		Auxiliary holes, gear teeth, and forming				
4		$\frac{L}{D} \leq 2$ with deviation								
5		Special								
6	Nonrotational parts	$\frac{A}{B} < 3, \frac{A}{C} > 4$ Flat parts	Main shape	Principal bores	Plane surface machining	Auxiliary holes, gear teeth, and forming				
7		$\frac{A}{B} > 3$ Long parts	Main shape							
8		$\frac{A}{B} < 3, \frac{A}{C} > 4$ Cubic parts	Main shape							
9		Special								

Figure 38.11: Classification and coding system according to Opitz, consisting of a form code of five digits and a supplementary code of four digits.

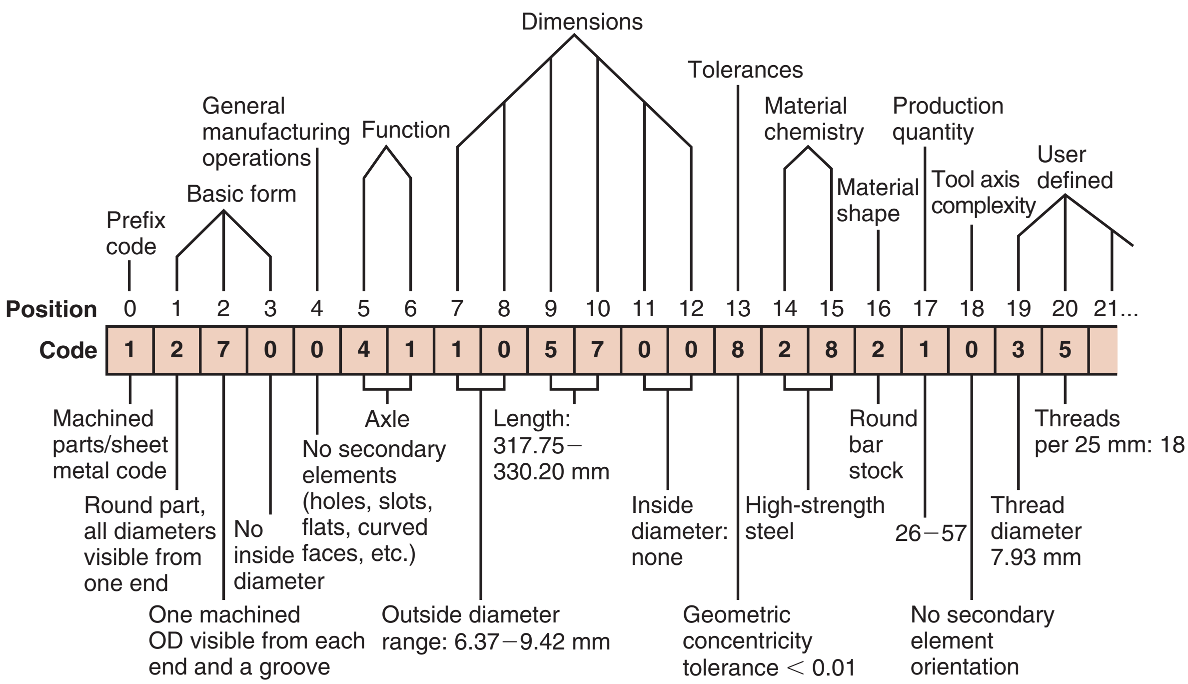


Figure 38.12: Typical multiClass code for a machined part. The routing sheet travels with the part from operation to operation. Source: Courtesy of Organization for Industrial Research.

Digit	Items		(Rotational component)
1	Part name		General classification
2			Detail classification
3	Materials		General classification
4			Detail classification
5	Major dimensions		Length
6			Diameter
7	Primary shapes and ratio of major dimensions		
8	Shape details and kinds of processes	External surface	External surface and outer primary shape
9			Concentric screw-threaded parts
10			Functional cutoff parts
11			Extraordinary shaped parts
12			Forming
13			Cylindrical surface
14		Internal surface	Internal primary shape
15			Internal curved surface
16			Internal flat and cylindrical surface
17		End surface	
18		Nonconcentric holes	Regularly located holes
19			Special holes
20		Noncutting process	
21	Accuracy		

Figure 38.13: The structure of a KK-3 system for rotational components. *Source:* Courtesy of Japan Society for the Promotion of Machine Industry.

Part	Broach	CNC Lathe	Hor. Machining Center	Hob	Manual Lathe	Hor. Lathe	Screw machine
Cover bearing		X					
Pinion gear, 56 teeth, 8 pitch	X	X		X			
Shaft coupling							X
Impellor	X	X					
6” Gland					X		
Driven gear, 26 teeth, 8 pitch	X	X		X			
Bushing							X
Body casing			X			X	
Relief valve			X				
Bearing spacer							X
Intake manifold			X		X	X	
Yoke bar							X
Spring seat							X
Yoke		X		X			
Pump head					X	X	

(a)

Part	Manual Lathe	Hor. Lathe	Hor. Machining Center	Broach	CNC Lathe	Hob	Screw machine
6” Gland	X						
Pump head	X	X					
Intake manifold	X	X	X				
Body casing		X	X				
Relief valve			X				
Impellor				X	X		
Yoke					X	X	
Cover bearing					X		
Pinion gear, 56 teeth, 8 pitch				X	X	X	
Driven gear, 26 teeth, 8 pitch				X	X	X	
Bearing spacer							X
Shaft coupling							X
Bushing							X
Yoke bar							X
Spring seat							X

Turning/ milling cell

CNC turning/ broach/hob cell

(b)

Figure 38.14: The use of group technology to organize manufacturing cells. (a) Original spreadsheet of parts; no logical organization is apparent. (b) After grouping parts based on manufacturing attributes, logical machinery groupings are discernible.

Summary

- Integrated manufacturing systems can be implemented to various degrees to optimize operations, improve product quality, and reduce production costs.
- Computer-integrated manufacturing operations have become the most important means of improving productivity, responding to rapidly changing market demands, as well as improving the control of both manufacturing and the management functions of an organization.
- Advanced software and powerful computers now allow the description of part geometry in several different formats, including wire-frame, surface models, solid models, skeletons, and boundary representations.
- Computers also are used to simulate manufacturing operations and systems, and to aid in the selection of manufacturing processes.
- Group technology is a powerful approach that allows the rapid recovery of previous design and manufacturing experiences by encoding a part on the basis of its geometric features or manufacturing attributes. Several group-technology coding systems are available.

Key Terms

Business-execution function

Business-planning function

Classification and coding systems

Coding

Computer-aided design and engineering
Computer-aided manufacturing
Computer-aided process planning
Computer-integrated manufacturing
Computer simulation
Data-acquisition system
Database
Design attributes
Enterprise resource planning
Exchange specifications
Functional layout
Group layout
Group technology
Manufacturing attributes
Manufacturing resource planning
Material-requirements planning
Paperless design
Parametric model
Process planning
Production flow analysis
Routing sheet
Solid model
Surface model
Wire frame

Bibliography

Amirouche, F.M.L., **Principles of Computer-Aided Design and Manufacturing,** 2nd ed., Prentice Hall, 2004.

Chang, T.-C., Wysk, R.A., Wang, H.P., and Rembold, U., **Computer-Aided Manufacturing**, 3rd ed., Prentice Hall, 2007.

Chryssolouris, G., **Manufacturing Systems: Theory and Practice**, Springer, 2006.

Curry, G.L., and Feldman, R.M., **Manufacturing Systems Modeling and Analysis**, Springer, 2008.

Groover, M., **Automation, Production Systems and Computer-Integrated Manufacturing**, 4th ed., Prentice Hall, 2014.

Parsaei, H., Leep, H., and Jeon, G., **The Principles of Group Technology and Cellular Manufacturing Systems**, Wiley, 2006.

Rehg, J.A., and Kraebber, H.W., **Computer-Integrated Manufacturing**, 3rd ed., Prentice Hall, 2004.

Rong, Y., and Huang, H., **Advanced Computer-Aided Fixture Design**, Academic Press, 2005.

Review Questions

38.1. In what ways have computers had an impact on manufacturing?

38.2. Describe the benefits of computer-integrated manufacturing operations.

38.3. What is a database? Why is it necessary?

38.4. What is an STL file? What is it popularly used for?

38.5. What are the differences between the terms “computer-aided” and “computer-integrated”?

38.6. What are the advantages of CAD systems over traditional methods of design? Are there any limitations?

38.7. What do the following abbreviations mean: NURB, DAS, DFX, PDES?

38.8. What are the main advantages of CIM?

38.9. Describe the purposes of process planning. How are computers used in such planning?

38.10. Describe the features of a routing sheet. Why is it necessary?

38.11. What are the advantages of simulation of manufacturing lines?

38.12. What is group technology? Why was it developed? Explain its advantages.

38.13. Explain the three types of GT coding: hierarchical, polycode, and decision-tree.

38.14. Describe what is meant by the term "manufacturing system." What are its benefits?

38.15. What does classification and coding mean in group technology?

38.16. What do the terms Opitz and KK-3 have in common?

38.17. Examine a cardboard box. Is this an example of a boundary representation, a solid model, or a profile body?

Qualitative Problems

38.18. Describe your observations regarding Figs. 38.1 and 38.2.

38.19. Some software packages displace STL files as collections of surface quadrilaterals, not triangles. Explain why this can be done.

38.20. Give examples of primitives of solids other than those shown in Fig. 38.4a and b.

38.21. Explain the logic behind the arrangements shown in Fig. 38.9b.

38.22. What are the advantages of hierarchical coding?

38.23. What is the difference between a variant system and a generative system?

38.24. Referring to Fig. 38.3, explain the advantages of a third-order piecewise Bezier curve over a B-spline or a conventional Bezier curve.

38.25. Sketch a sphere using (a) wire fame; (b) 3D surface models; and (c) STL formats.

38.26. Describe situations that would require a design change at its larger end of the part in Fig. 38.6.

38.27. Describe your thoughts on the differences between e-commerce and traditional business practices.

Synthesis, Design, and Projects

38.28. How would you describe the principle of computer-aided manufacturing to an older worker in a manufacturing facility who is not familiar with computers?

38.29. Review various manufactured parts illustrated in this book, and group them in a manner similar to those shown in Fig. 38.8. Explain.

38.30. Think of a simple part and make a decision-tree chart similar to that shown in Fig. 38.10.

38.31. Review the machine arrangements in Fig. 38.9 and suggest changes that may improve the flow of materials and parts.

38.32. Think of a simple product and make a routing sheet, similar to that shown in Fig. 38.7. If the same part is given to another person, what is the likelihood that the routing sheet developed will be the same? Explain.

38.33. Review Fig. 38.7, and prepare a routing sheet for one of the following: (a) a spur gear, (b) a turbine blade, (c) a glass bottle, (c) an automotive connecting rod, and (d) a forging die.

38.34. Consider a large-scale kitchen, such as will be used to prepare meals for a school, hospital, or large office building. Make a list of equipment necessary, arrange the equipment in a functional layout and then arrange the equipment in a cellular layout. Describe your observations for flow lines when different foods are produced.

38.35. Using drawing tools, carefully create a sketch of the end view of the part in Fig. 38.6a. Make the same sketch on a CAD program. Compare the time required to produce each sketch.

Chapter 39

Computer-integrated Manufacturing Systems

- This chapter describes the effects of computer systems and communications networks on product development and manufacturing, and the integration of all related activities.
- The chapter begins by describing the principles of manned and untended manufacturing cells and their features, and how cells can be integrated into flexible manufacturing systems. The concept of holonic manufacturing and its applications also are reviewed.
- Just-in-time production, lean manufacturing, and communication systems are then described.
- The chapter also discusses the emerging concept of mass customization, driven by new process capabilities and design tools, allowing every part to be optimized for a user or application.
- The chapter concludes with a discussion of artificial intelligence and ESs as applied to manufacturing.

39.1 Introduction

This chapter focuses on the **computer integration of manufacturing operations**, where *integration* means that manufacturing processes, operations, and their management are all treated as a *system*. A major advantage of such an approach is that machines, tooling, and manufacturing operations now acquire a built-in flexibility, called **flexible manufacturing**. As a result, the system is capable of (a) *rapidly responding* to changes in product types and fluctuating demands and (b) ensuring *on-time delivery* of products to the customer. Failure of on-time delivery in a highly competitive global environment can have major adverse effects on a company's operations and its success.

This chapter first describes the key elements that enable the execution of the functions necessary to achieve flexible manufacturing, beginning with *cellular manufacturing*, which is the basic unit of flexibility in the production of goods. Manufacturing cells can be broadened into *flexible manufacturing systems*, with major implications for production capabilities. Described next is *holonic manufacturing*, which is a strategy of organizing manufacturing units such as to achieve higher efficiency in operations.

The important concept and the benefits of *just-in-time production* is then examined, in which parts are produced "just in time" to be made into subassemblies, assemblies, and final products. This is a method that eliminates the need for inventories (a major financial burden on a company), as well as significantly saving on space and storage facilities.

Because of the necessity for and extensive use of computer controls, hardware, and software in all the activities just outlined, the planning and effective implementation of *communications networks* constitute a critical component of the overall manufacturing operations. The chapter concludes with a review of *artificial intelligence*, which consists of expert systems, natural-language processing, machine vision, and artificial neural networks.

39.2 Cellular Manufacturing

A **manufacturing cell** is a small unit consisting of one or more workstations. A workstation usually contains either one machine (called a *single-machine cell*) or several machines (*group-machine cell*), with each machine performing a different operation. The machines can be modified, retooled, and regrouped for different product lines within the same family of parts.

Cellular manufacturing has thus far been utilized primarily in machining, finishing, and sheet-metal-forming operations. The machine tools commonly used in the cells are lathes, milling machines, drills, grinders, and electrical-discharge machines; for sheet forming, the equipment typically consists of shearing, punching, bending, and other forming machines. The equipment may include special-purpose machines and CNC machines; automated inspection and testing equipment also are generally a part of a manufacturing cell.

The capabilities of cellular manufacturing typically involve the following operations:

- Loading and unloading raw materials and workpieces at workstations
- Changing tools at workstations
- Transferring workpieces and tooling between workstations
- Scheduling and controlling the total operation in the cell.

In *manned* or attended machining cells, raw materials and parts can be moved and transferred manually by the operator (unless the parts are too heavy or the movements are too hazardous) or by an industrial robot (Section 37.6) located centrally in the cell.

Flexible Manufacturing Cells. Manufacturing cells can be made flexible by using CNC machines and machining centers (Section 25.2) and by means of industrial robots or other mechanized systems for handling materials and parts in various stages of completion (Section 37.6).

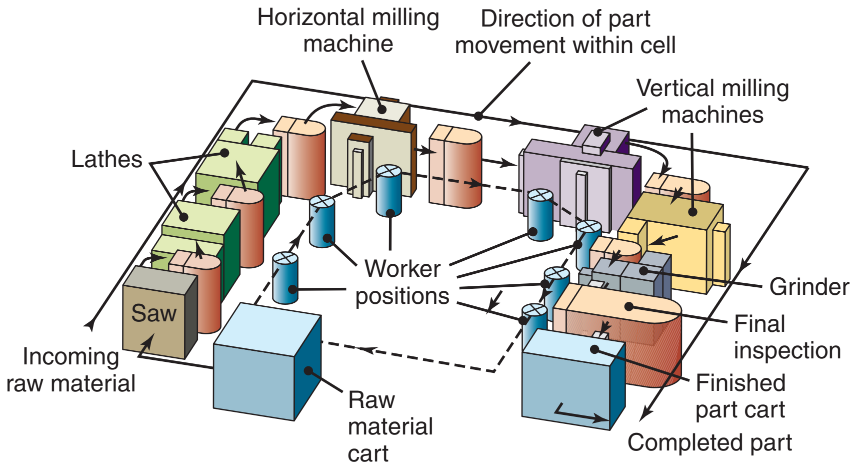


Figure 39.1: Schematic illustration of a manned flexible manufacturing cell, showing various machine tools and an inspection station. *Source:* After JT. Black.

An example of an attended *flexible manufacturing cell* (FMC) that involves several machining operations is illustrated in Fig. 39.1. Note that machining centers, fitted with automatic tool changers and tool magazines, have the capability of performing a wide variety of operations (see Section 25.2). A computer-controlled inspection station with a coordinate-measuring machine or other flexible metrology equipment can similarly inspect dimensions on a wide variety of parts. Thus, the organization of these machines into a cell can allow the manufacture of very different parts. With computer integration, a manufacturing cell can produce parts in batch sizes as small as one part, with negligible delay between batches.

Flexible manufacturing cells are usually *unmanned* or unattended and their design and operation are more exacting than those for other cells. The selection of the machines and industrial robots, including the types and capabilities of their end effectors and control systems, is critical to the effective functioning of the FMC. As with other flexible manufacturing systems (Section 39.3), the cost of FMCs is very high (see Table 40.6); however, this disadvantage is outweighed by increased productivity, flexibility, and controllability.

Cell Design. Because of the unique features of manufacturing cells, their design and placement requires efficient layout and organization of the plant and the consideration of product flow lines (see *production flow analysis*, Section 38.8). The machines may be arranged along a line or in a U-shape, an L-shape, or a loop. Selecting the best machine and arrangement for material-handling equipment also involves taking into account such factors as the type of product, its shape, size, weight, and production rate. In designing these cells, the likelihood of a significant change in demand for a particular part families must be considered, in order to ensure that the equipment involved has the required flexibility and capacity.

39.3 Flexible Manufacturing Systems

A *flexible manufacturing system* (FMS) integrates all of the major elements of production into a highly automated system (Fig. 39.2). Utilized in the late 1960s, an FMS consists of (a) a number of manufacturing cells, each containing an industrial robot serving several CNC machines and (b) an automated material-handling system. In Fig. 39.2, an automated guided vehicle (AGV; Section 37.5) moves parts between machines and inspection stations, which could easily involve an industrial robot or integral transfer device. All of these activities are interfaced with a central computer and different computer instructions can be downloaded

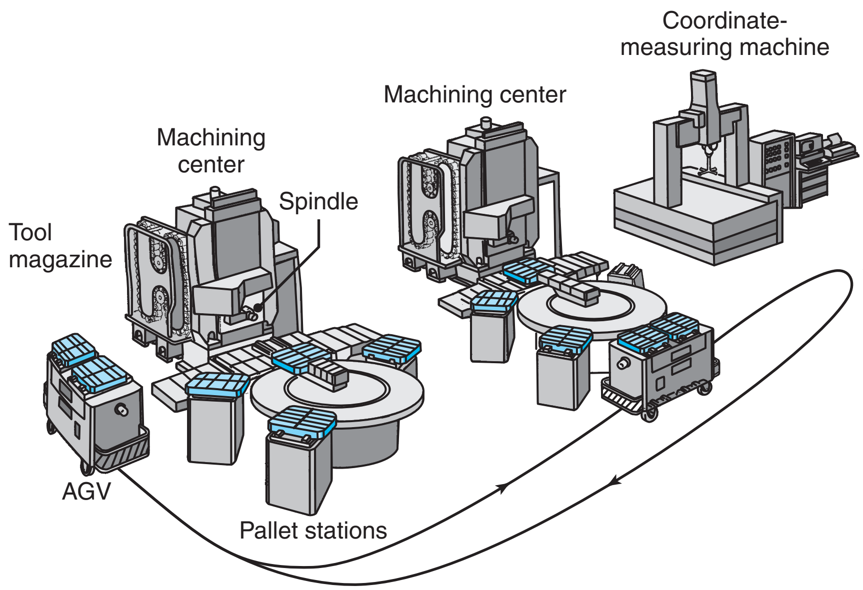


Figure 39.2: A schematic illustration of a flexible manufacturing system, showing two machining centers, a measuring and inspection station, and two automated guided vehicles. *Source:* After JT. Black.

Figure 39.3: A Robotic arm production line. *Source:* Shutterstock/Andrey Armyagov.

Table 39.1: Comparison of General Characteristics of Transfer Lines and Flexible Manufacturing Systems.

Characteristic	Transfer line	FMS
Part variety	Few	Infinite
Lot size	> 100	1–50
Part-changing time	Long	Very short
Tool change	Manual	Automatic
Adaptive control	Difficult	Available
Inventory	High	Low
Production during breakdown	None	Partial
Justification for capital expenditure	Simple	Difficult

for each successive part passing through a particular workstation. The system can handle a *variety of part configurations* and produce them *in any order*.

An FMS is capable of optimizing each step of the total operation. These steps may involve (a) one or more processes and operations, such as machining, grinding, cutting, forming, powder metallurgy, heat treating, and finishing; (b) the handling of raw input materials and parts; (c) measurement and inspection; and (d) assembly. The most common applications of FMS to date have been in machining and assembly operations.

FMS can be regarded as a system that combines the benefits of two systems: (1) the highly productive, but inflexible, transfer lines and (2) job-shop production, which can produce large product variety on stand-alone machines, but is inefficient (see also Fig. 37.2). The relative characteristics of transfer lines and FMS are shown in Table 39.1. Note that with an FMS, the time required for a changeover to a different part is very short. The quick response to product and market-demand variations is a major attribute of FMS.

Compared with conventional manufacturing systems, FMS have the following major benefits:

- Parts can be produced in any order, in batch sizes as small as one, and at lower unit cost
- Direct labor and inventories are reduced or eliminated
- The lead times required for product changes are shorter
- Because the system is self-correcting, production is more reliable and product quality is uniform.

However, FMS has the following drawbacks:

- Equipment costs are higher, and computer integration requires more highly trained operating personnel
- FMS is generally not suitable for very low or very high production rates.

Elements of FMS. The basic elements of a flexible manufacturing system are the following:

- Computers and software for tracking of parts and materials through a plant and for control of supply chains;
- Automated handling and transport of materials and parts
- Control systems.

The workstations are arranged to yield the highest efficiency in production, with an orderly flow of materials and work in progress (WIP) through the system. The types of machines in workstations depend on the type of production. For machining operations, for example, they usually consist of a variety of

3- to 5-axis machining centers, CNC lathes, milling machines, drill presses, and grinders (described in Part IV). Also included are various pieces of equipment, such as that for automated inspection (including coordinate-measuring machines), assembly, and cleaning. In addition to machining, other types of manufacturing operations suitable for FMS are sheet-metal forming, punching, shearing, and forging. FMS also may incorporate various equipment, trimming presses, heat-treating facilities, and cleaning equipment.

The flexibility of FMS requires that material-handling systems be computer-controlled and performed by automated guided vehicles, conveyors, and various transfer mechanisms. Thus, FMS are capable of transporting raw materials, blanks, and parts in various stages of completion to any machine, in random order, and at any time. *Prismatic parts* are usually moved on specially designed **pallets**, whereas those having *rotational symmetry* are usually moved by robots and various mechanical devices.

Scheduling. In FMS, efficient machine utilization is essential; machines must not stand idle, thus proper scheduling and process planning are crucial. Unlike that in job shops where a relatively rigid schedule is followed to perform a set of operations, scheduling for FMS is *dynamic*, meaning that it is capable of responding to quick changes in product type; thus it is responsive to real-time decisions. The scheduling system in FMS (a) specifies the types of operations to be performed on each part and (b) identifies the machines or manufacturing cells on which these operations are to take place.

No setup time is wasted in switching between manufacturing operations. The characteristics, performance, and reliability of each unit in the system must, however, be monitored to ensure that parts are of acceptable quality and dimensional accuracy before they move on to the next workstation.

Economic Justification of FMS. FMS installations are very capital intensive, costing millions of dollars. Consequently, a thorough cost-benefit analysis must be conducted that must include such factors as (a) the costs of capital, energy, materials, and labor; (b) the expected markets for the products to be manufactured; (c) any anticipated fluctuations in market demand and product type; and (d) the time and effort required for installing and debugging the whole system.

As can be seen in Fig. 37.2, the most effective FMS applications are in medium-quantity batch production. When a variety of parts is to be produced, FMS is suitable for production quantities typically up to 15,000 to 35,000 aggregate parts per year. For individual parts with the same configuration, production may, however, reach 100,000 parts per year. In contrast, high-volume, low-variety parts production is best obtained from transfer machines (*dedicated equipment*). Finally, low-volume, high-variety parts production can best be done on conventional standard machinery (with or without numerical control) or by using machining centers.

Example 39.1 Flexible Manufacturing Systems in Large and Small Companies

Many manufacturers have long considered implementing a large-scale system in their facilities. After detailed review and on the basis of the experience of other similar companies, however, most decide on smaller, simpler, modular, and less expensive systems that are more cost effective. These systems include flexible manufacturing cells (the cost of which would be on the order of a few hundred thousand dollars), stand-alone machining centers, and various CNC machine tools that are significantly easier to control than an FMS.

39.4 Mass Customization

Mass customization implies that a manufacturing facility can produce a large number of parts, each of which can be dramatically different from the other. The important technologies that enable mass customization are:

1. Easy to use computer-aided design (CAD) software, along with design tools and part libraries to allow rapid description of parts or products.

2. Machinery capable of efficiently producing part quantities of one. This is commonly thought to be the exclusive characteristic of additive manufacturing approaches (Chapter 20), but this is also the case for machining using modern software to generate machine instructions.

In theory, mass customization has the same benefits of just-in-time (see Section 39.6). In practice, these attributes are secondary to the advantages of mass customization, including the following:

1. Supply chains are greatly simplified, since raw materials are limited to the options for the machine-selected polymers or metal powders for additive manufacturing, materials with sufficient machinability for CNC milling, etc.

2. Replacement parts do not need to be maintained in storage or ordered when needed; they can instead be produced on demand. This reduces or eliminates the time needed to obtain replacement parts.

3. Parts that are produced specifically for their intended application are generally better at delighting the customer (see Section 36.5).

To date, mass customization has not been economically competitive with conventional manufacturing approaches where parts are produced in batches or in larger part quantities. However, it is attractive for some applications. For example, consider a large ocean-going ship or warship that is at sea and cannot obtain a replacement part. Instead of maintaining an enormous spare part inventory, a replacement part can be produced at sea. It is recognized that in this case, the replacement part need not have as high a performance as the original, but if it can be sufficient to bring the ship to port it represents a significant advantage.

39.5 Holonic Manufacturing

Holonic manufacturing describes a unique organization of manufacturing units, whereby each component in a holonic manufacturing system is, at the same time, an independent or *whole* entity and a subservient *part* of a hierarchical organization. The word *holonic* is from the Greek *holos* (meaning "whole") and the suffix *on* (meaning "a part of").

Holonic organizational systems have been studied since the 1960s, and there are a number of examples in biological systems. Three fundamental observation features about these systems are the following:

1. Complex systems will evolve from simple systems much more rapidly if there are stable intermediate forms than if there are none. Also, stable and complex systems require a hierarchical system for their evolution.

2. Holons are simultaneously self-contained wholes of their subordinated parts and dependent parts of other systems. Holons are autonomous and self-reliant units that have a degree of independence and can handle contingencies without asking higher levels in the hierarchical system for instructions. At the same time, holons are subject to control from multiple sources of higher system levels.

3. A *holarchy* consists of (a) autonomous wholes in charge of their parts and (b) dependent parts controlled by higher levels of a hierarchy. Holarchies are coordinated according to their local environment.

In biological systems, hierarchies have the characteristics of (a) stability in the face of disturbances, (b) optimum use of available resources, and (c) a high level of flexibility when their environment changes. A *manufacturing holon* is an autonomous and cooperative building block of a manufacturing system for the production, storage, and transfer of objects or information. It consists of a control part and an optional physical-processing part. For example, a holon can be a combination of a CNC milling machine and an operator interacting via a suitable interface. A holon can also consist of other holons that provide the necessary processing, information, and human interfaces to the outside world, such as a group of manufacturing cells. Holarchies can be created and dissolved dynamically, depending on the current needs of the particular manufacturing process.

A holonic system's view of the manufacturing operation is one of creating a working manufacturing environment from the bottom up. Maximum flexibility can be achieved by providing intelligence within holons to (a) support all production and control functions required to complete production tasks and (b) manage the underlying equipment and systems. The manufacturing system can dynamically reconfigure itself into operational hierarchies to optimally produce the desired products, with holons or elements being added or removed as needed.

Holarchical manufacturing systems rely on fast and effective communication among holons, as opposed to traditional hierarchical control where individual processing power is essential. Although a large number of specific arrangements and software algorithms have been proposed for holarchical systems, the general sequence of events can be outlined as follows:

1. A factory consists of a number of *resource holons*, available as separate entities in a resource pool. Available holons may, for example, consist of (a) a CNC milling machine and operator, (b) a CNC grinder and operator, or (c) a CNC lathe and operator.

2. Upon receipt of an order or a directive from higher levels in the factory hierarchical structure, an *order holon* is formed and begins communicating and negotiating with the available resource holons.

3. The negotiations lead to a self-organized grouping of resource holons, which are assigned on the basis of product requirements, resource holon availability, and customer requirements. For example, a given product may require a CNC lathe, a CNC grinder, and an automated inspection station to organize it into a *production holon*.

4. In case of a breakdown, the unavailability of a particular machine or due to changing customer requirements, other holons from the resource pool can be added or subtracted as needed, allowing a reorganization of the production holon. Production bottlenecks can be identified and eliminated through communication and negotiation among the holons in the resource pool. Step 4 has been referred to as *plug and play*, a term borrowed from the computer industry, where hardware components are seamlessly integrated into a system.

39.6 Just-in-time Production

The *just-in-time* (JIT) *production* concept originated in the United States, with early contributions by Ford Motor Co. and John Deere Corporation, and was first implemented on a large scale in 1953 at the Toyota Motor Company in Japan to eliminate waste of materials, machines, capital, manpower, and inventory throughout the manufacturing system. The JIT concept has the following goals:

- Receive supplies just in time to be used
- Produce parts just in time to be made into subassemblies
- Produce subassemblies just in time to be assembled into finished products
- Produce and deliver finished products just in time to be sold.

In traditional manufacturing, the parts are made in batches, placed in inventory, and then used whenever necessary. This approach is known as a **push system**, meaning that parts are made *according to a schedule* and are placed in inventory, to be used whenever they are needed. In contrast, JIT is a **pull system**, meaning that parts are *produced to order* and thus the production is matched with demand for the final assembly of products.

There are no stockpiles in JIT and the ideal production quantity is one, hence it is also called *zero inventory*, *stockless production*, and *demand scheduling*. In this manner, it is similar to the concept of *mass customization*, Section 39.4, but the JIT approach applies equally well to any production volume.

In JIT, parts are inspected as they are manufactured and are used within a short period of time. In this way, a worker maintains continuous production control, immediately identifying defective parts and reducing process variation to produce quality products.

Implementation of the JIT concept requires that all aspects of manufacturing operations be monitored and reviewed so that all operations and resources that do not add value are eliminated. This approach emphasizes (a) pride and dedication in producing high-quality products; (b) elimination of idle resources; and (c) teamwork among workers, engineers, and management to promptly solve any problems that arise during production or assembly.

The ability to detect production problems as parts are being produced has been likened to what happens to the level of water (representing *inventory levels*) in a lake covering a bed of boulders (representing *production problems*). When the water level is high (analogous to the high levels of inventory associated with push production), the boulders are not exposed. By contrast, when the level is low (analogous to the low inventories associated with pull production), the boulders are exposed and thus can promptly be identified and removed. This analogy indicates that high inventory levels can mask quality and production problems with parts that are already made and stockpiled.

The just-in-time concept requires the timely and reliable delivery of all supplies and parts from outside sources and/or from other divisions of a company, called *supply chain management*. Effective supply chain management, which can include global considerations, significantly reduces or eliminates in-plant inventory. Suppliers are expected to deliver, often on a daily basis, preinspected goods as they are needed for production and assembly. The reliability of supply can become a major concern, particularly in generally unpredictable major natural events, such as earthquakes, flooding, and nuclear disasters.

The JIT approach requires (a) reliable suppliers, (b) close cooperation and trust among the company and its vendors, and (c) a reliable system of transportation. Also important for smoother operation is a reduction in the number of its suppliers.

Advantages of JIT. The major advantages of just-in-time production are the following:

- Low inventory-carrying costs
- Fast detection of defects in the production or the delivery of supplies and, hence, low scrap loss
- Reduced inspection and reworking of parts
- High-quality products made at low cost.

Although there can be significant variations in performance, JIT production has resulted in reductions of 20–40% in product cost, 60–80% in inventory, up to 90% in rejection rates, 90% in lead times, and 50% in scrap, rework, and warranty costs. Increases of 30–50% in direct-labor productivity and of 60% in indirect-labor productivity also have been attained.

Kanban. The implementation of JIT in Japan involved *kanban*, meaning "visible record." These records originally consisted of two types of cards (called *kanbans*, now replaced by bar-coded tags and other devices):

- The *production card*, which authorizes the production of one container or cart of identical, specified parts at a workstation
- The *conveyance card* or *move card*, which authorizes the transfer of one container or cart of parts from a particular workstation to the workstation where the parts will be used.

The cards or tags contain information on the (a) type of part, (b) location where the card was issued, (c) part number, and (d) number of items in the container. The number of containers in circulation at any time is completely controlled and can be scheduled as desired for maximum production efficiency.

39.7 Lean Manufacturing

In a modern manufacturing environment, companies must be responsive to the needs of the customers and their specific requirements and to fluctuating global market demands. At the same time, to ensure competitiveness the manufacturing enterprise must be conducted with a minimum amount of wasted resources. This realization has lead to *lean production* or *lean manufacturing* strategies.

Lean manufacturing involves the following steps:

1. **Identify value.** The critical starting point for lean thinking is a recognition of *value*, which can be done only by a customer and considering a customer's product (see also Section 40.10.1). The goal of any organization is to produce a product that a customer wants, at a desired price, location, time, and volume. Providing the wrong good or service produces waste, even if it is delivered efficiently. It is important to identify all of a manufacturers activities from the *viewpoint of the customer* and *optimize* processes to maximize added value. This viewpoint is critically important because it helps identify whether or not a particular activity:
    a. Clearly adds value
    b. Adds no value, but cannot be avoided
    c. Adds no value, and can be avoided.
2. **Identify value streams.** *Value stream* is the set of all actions required to produce a product, including
    a. Product design and development tasks, involving all actions from concept, to detailed design, and to production launch
    b. Information management tasks, involving order taking, detailed schedule, and delivery
    c. Physical production tasks, by means of which raw materials progress to a finished product delivered to the customer.
3. **Make the value stream flow.** *Flow* is achieved when parts encounter a minimum of idle time between any successive operation. It has been noted that flow is easiest to achieve in mass production, but it is more difficult for small-lot production. However, production in batches inherently involves part idle time; this should be avoided, thus just-in-time approaches are essential. In such cases the solution is to use manufacturing cells , where minimum time and effort are required to switch from one product to another; a product being manufactured encounters continuous flow regardless of the production rate.

    In addition to JIT approaches, establishing product flow through factories requires:

    - Eliminating waiting time, which may be caused by unbalanced workloads, unplanned maintenance, or quality problems; thus, the efficiency of workers must be maximized at all times.
    - Leveling or balancing production. Production may vary at different times of the day or the day of the week. Uneven production occurs when machines are underutilized, invariably leading to waiting time and waste. If production can be leveled, inventory is reduced, productivity increases, and process flow is easier to achieve. It should also be noted that balancing can occur on cell, line, and plant scales.
    - Eliminating unnecessary or additional operations and steps, because they represent costs.

- Minimizing or eliminating movements of parts or products in plants, because it represents an activity that adds no value. This waste can, for example, be either eliminated or minimized by using machining cells or better plant layouts.
- Performing motion and time studies to avoid unnecessary part or product movements or to identify inefficient workers.
- Eliminating part defects.
- Avoiding a single-source supplier, especially in the unforeseen events of natural disasters and regional conflicts.

4. **Establish pull.** It has been observed that once value streams are flowing, significant savings are gained in terms of inventory reduction, as well as product development, order processing, and production. In some cases, up to 90% savings in actual production have been obtained. Under these circumstances, it is possible to establish *pull manufacturing* (Section 39.6), where products are produced upon order by a customer or upstream machine, and not in batches that ultimately are unwanted and do not create value.

5. **Achieve perfection.** As described in Section 36.1, *kaizen* is used to signify continuous improvement, and clearly, there is a need for continuous improvement in all organizations. With lean manufacturing approaches, it has been found that continuous improvement can be accelerated, so that production without waste is possible. Moreover, upon the adoption of lean manufacturing principles, companies encounter an initial benefit, referred to as *kaikaku* or "radical improvement."

39.8 Communications Networks in Manufacturing

In order to maintain a high level of coordination and efficiency of operation in integrated manufacturing, an extensive, high-speed, and interactive *communications network* is essential. The **local area network** (LAN) is a hardware-and-software system in which logically related groups of machines and equipment communicate with each other. A LAN links these groups to each other, bringing different phases of manufacturing into a unified operation.

A LAN can be very large and complex, linking hundreds or even thousands of machines and devices in several buildings of an organization. Various network layouts of fiber-optic or copper cables are typically used, over distances ranging from a few meters to as much as 32 km; for longer distances, **wide area networks** (WANs) are used. Different types of networks can be linked or integrated through "gateways" and "bridges," often with the use of secure *file transfer protocols* (FTPs) over Internet connections. Several advanced network protocols, including *ipV6* (Internet protocol version 6) and *Internet2*, have been implemented and applied to manufacturing networks. A *carrier-sense multiple access with collision detection* (CSMA/CD) system was developed in the 1970s and implemented in *Ethernet*, which is now the industry standard.

Conventional LANs require the routing of wires, often through masonry walls or other permanent structures, and require computers and machinery to remain stationary. **Wireless local area networks** (WLANs) allow equipment such as mobile test stands or data-collection devices (e.g., bar-code readers) to easily maintain a network connection. A communication standard (IEEE 802.11) currently defines frequencies and specifications of signals, two radio-frequency methods, and one infrared method for WLANs. Although wireless networks are slower than those which are hardwired, their flexibility makes them desirable, especially in situations where slow tasks, such as machine monitoring, are the main application.

Personal area networks (PANs) are used for electronic devices, such as cellular telephones and personal data assistants, but are not as widespread for manufacturing applications. PANs are based on communications standards (such as Bluetooth, IrDA, and HomeRF) and are designed to allow data and voice communication over short distances. For example, a short-range Bluetooth device will allow communication over a 10-m distance. PANs are undergoing major changes, and communications standards are continually being refined.

Communications Standards. Typically, one manufacturing cell is built, with machines and equipment purchased from one vendor, another cell, with machines purchased from another vendor, and a third, purchased from yet another vendor. As a result, a variety of programmable devices are involved and are driven by several computers and microprocessors purchased at various times from different vendors and having various capacities and levels of sophistication.

In 2008, **MTConnect** was demonstrated for the first time. *MTConnect* is a machine tool communications protocol, developed by the Association for Manufacturing Technology, which has quickly become an industry standard (aided by its royalty-free availability). MTConnect uses *hypertext transfer protocols* (http) for communication of data suitable for all machine-tool manufacturers. Software is widely available for retrieving data from machine tools, called agents, and is even available for tablet computers, smartphones, or specially designed Bluetooth-based devices. Collection of manufacturing data, tracking of machine utilization and production rates, and other forms of data is greatly enabled by this protocol, allowing for real-time plant management. MTConnect is undergoing rapid development to extend the software to a wide variety of processes and systems.

39.8.1 The Internet of Things

The **Internet of Things** (IoT) is a concept that refers to objects being connected over the Internet; it has already progressed significantly in several applications. A person may have, for example, a number of devices that communicate either with each other or with other devices on the Internet; cell phones, laptop computers, television sets, and gaming systems are well-known to be Internet capable. Automobiles, security cameras or web cams, and building environmental controls are also examples where Internet communication capability is now commonplace, but is not usually noticed; such integration is said to be transparent in that it is not noticed by the user.

In manufacturing, IoT refers to a state where all machines, sensors, fixtures, and even manufactured products are able to communicate with each other or with computers. A number of technological advances have been combined to enable much greater integration of design and manufacturing:

1. Communications protocols, such as MTConnect, have enabled machine tools to communicate with each other and with computers through Internet protocols.

2. Remote servers are available (**cloud-based storage**) to store data (see Fig. 39.4). Cloud-based storage can combine multiple storage devices as needed, providing essentially unlimited storage capability.

3. Sensors of all types have been integrated into machinery to allow monitoring of manufacturing processes. Roll-to-roll printing of integrated circuits and sensors (see Section 28.14) promises to speed the proliferation of sensors with integrated communication capability.

4. The development of Internet protocol version 6 has resulted in 2^{128} discrete web addresses, allowing hundreds of trillions of web addresses for every person and company in the world.

5. *Big Data* approaches have been developed to *mine* information from large data files. Big Data refers to the circumstances where the available data is too large to analyze with conventional approaches, and unique and dedicated software tools have been developed to extract or *mine* data. Internet search engines, such as Google, exemplify big data mining; it is common to find information using a few search terms, and to have valuable and relevant results based on a search of the entire Internet, produced almost immediately.

Combining these technologies has led to a manufacturing environment where every machine, part, tool, and fixture are connected with communications tools, and data from the entire manufacturing enterprise

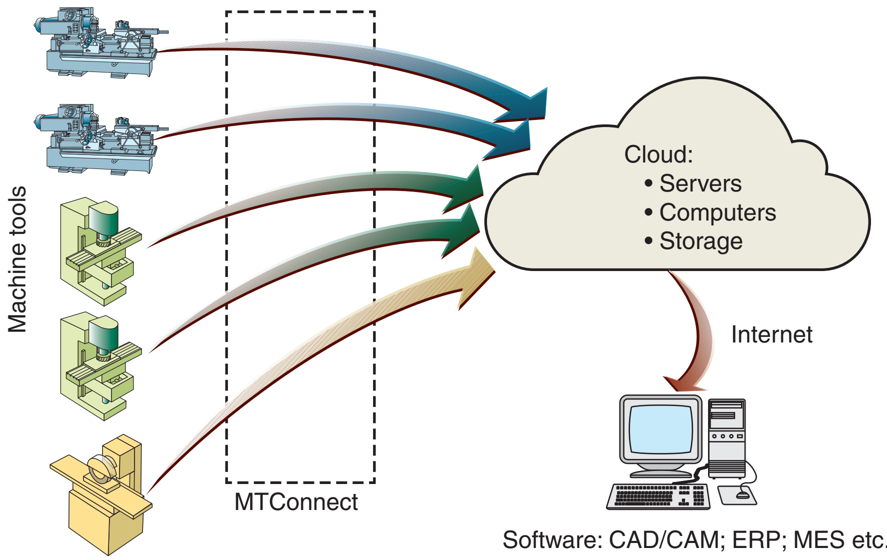


Figure 39.4: Schematic illustration of the use of cloud computing to capture manufacturing data and make it available wherever needed. A critical technology for cloud computing remains data security.

(including historical and real time) is available to anyone in an organization from cloud servers. This leads to an unprecedented access to information that can help guide management decisions.

A **digital twin** is a computer representation of a part, consisting of all the data collected from the time it was first manufactured and during its service. This can include data from a large number of sensors in production equipment, metrology results, hardness test results, loads and temperatures applied over time, etc. Since these sensors can provide data effectively continuously, big data approaches, combined with advanced material models, allow predictions of reliability in service and the need for preventative maintenance.

The availability of data also allows new approaches to machine control, quality control, and process validation, as well as extended capabilities for supply-chain management and ERP. Although The Internet of Things in manufacturing is still in its infancy, it continues to be developed at a rapid pace.

39.9 Artificial Intelligence and Machine Learning

Artificial intelligence (AI) and *machine learning* relate to that part of computer science concerned with systems that exhibit some characteristics that are usually associated with intelligence in humans: learning, reasoning, problem solving, recognizing patterns, and understanding language. The goal of AI is to simulate such behaviors on the computer. The art of bringing relevant principles and tools of AI to bear on difficult application problems is known as **knowledge engineering**.

Artificial intelligence has had major effect on the design, automation, and overall economics of the manufacturing operation, largely because of advances made in computer-memory expansion (see VLSI chip design; Chapter 28) and decreasing costs. Artificial intelligence packages costing as much as a few thousand dollars have been developed, many of which can be run on inexpensive personal computers.

Expert Systems. An *expert system* (ES), also called a **knowledge-based system**, generally is defined as an intelligent computer program that has the capability to solve difficult real-life problems by using **knowledge-based** and **inferential** procedures (Fig. 39.5) or Big Data to infer system behavior. The goal of an expert system is to conduct an intellectually demanding task in the way that a human expert would.

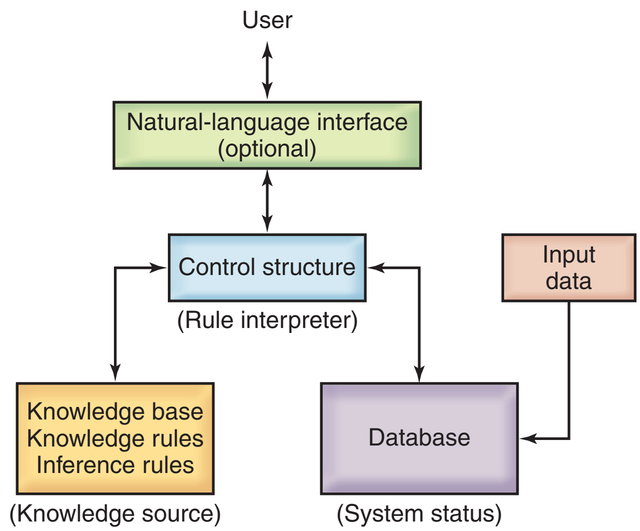


Figure 39.5: Basic structure of an expert system. The knowledge base consists of knowledge rules (general information about the problem) and inference rules (the way conclusions are reached). The results may be communicated to the user through the natural-language interface.

The field of knowledge required to perform the task in question is called the **domain** of the expert system. Expert systems utilize a knowledge base containing models, facts, data, definitions, and assumptions. They also have the capability of adopting a **heuristic** approach: making good judgments on the basis of discovery and revelation and making high-probability guesses, just as a human expert would. Expert systems can communicate with other computer software packages.

To construct expert systems for solving the complex design and manufacturing problems one encounters in real life, one needs a great deal of knowledge and a mechanism for manipulating that knowledge to create solutions. Because of the difficulty involved in (a) accurately modeling the many years of experience of an expert or a team of experts and (b) the complex inductive reasoning and decision-making capabilities of humans (including the capability to learn from mistakes), developing knowledge-based systems requires considerable time and effort. The increasing availability in Big Data is driving the development of machine learning and artificial intelligence algorithms for machinery applications.

Natural-language Processing. Traditionally, retrieving information from a database in computer memory has required the utilization of computer programmers to translate questions in natural language into “queries” in some machine language. Natural-language interfaces with database systems, which are in continuous development, allow a user to retrieve information by entering English or other language commands in the form of simple, typed questions.

Software shells are available and are used in such applications as scheduling material flow in manufacturing plants and analyzing information in databases. Significant progress continually is being made on computer software that will have speech synthesis and recognition (**voice recognition**) capabilities, in order to eliminate the need to type commands on keyboards.

Machine Vision. In machine vision (Section 37.7.1), computers and software, implementing artificial intelligence, are combined with cameras and other optical sensors (Fig. 39.6). These machines then perform such operations as inspecting, identifying, and sorting parts, as well as guiding robots (see *intelligent robots*, Section 37.6.2), operations that otherwise would require human intervention.

Artificial Neural Networks. Although computers are much faster than the human brain at performing sequential tasks, humans are much better at pattern-based tasks that can be performed with parallel

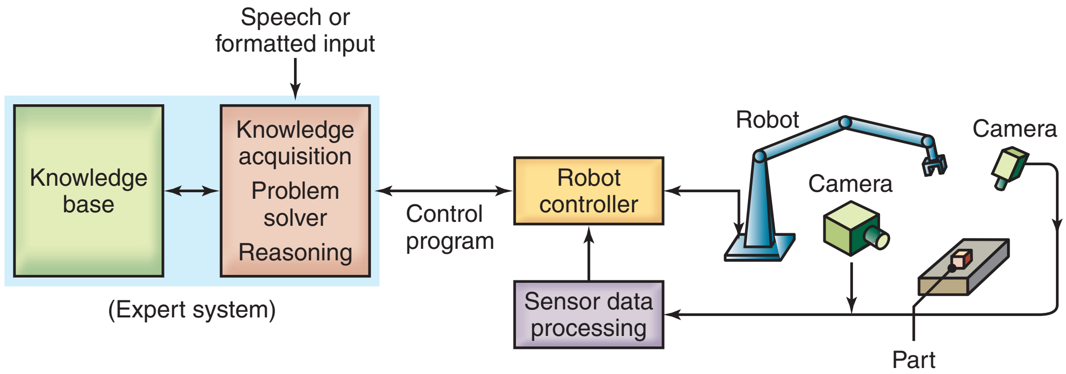


Figure 39.6: Illustration of an expert system as applied to an industrial robot guided by machine vision.

processing, such as recognizing features (on faces and in voices, even under noisy conditions), assessing situations quickly, and adjusting to new and dynamic conditions. These advantages also are due partly to the ability of humans to use all five senses (sight, hearing, smell, taste, and touch) in real time and simultaneously, called *data fusion*. The branch of AI called *artificial neural networks* (ANN) attempts to gain some of these capabilities through computer imitation of the way that data are processed by the human brain.

The human brain has about 100 billion linked **neurons** (cells that are the fundamental functional units of nerve tissue) and more than a thousand times that many connections. Each neuron performs exactly one simple task: it receives input signals from a fixed set of neurons; when those input signals are related in a certain way (specific to that particular neuron), it generates an electrochemical output signal that goes to a fixed set of neurons. It is now believed that human learning is accomplished by changes in the strengths of these signal connections among neurons.

Artificial neural networks are used in such applications as noise reduction (in telephones), speech recognition, and process control. For example, they can be used for predicting the surface finish of a workpiece obtained by end milling on the basis of input parameters such as cutting force, torque, acoustic emission, and spindle acceleration.

39.10 Economic Considerations

The economic considerations in implementing the various computer-integrated activities described in this chapter are critical in view of their many complexities and the high costs involved. Flexible manufacturing system installations are very capital intensive, thus requiring a thorough cost–benefit analysis, including the following:

- Costs of capital, energy, materials, and labor
- Expected markets for the products to be made
- Anticipated fluctuations in market demand and in the type of product
- Time and effort required for installing and debugging the system.

Typically, an FMS system can take two to five years to install and a few months to debug . Although it requires few, if any, machine operators, the personnel in charge of the total operation must be trained and highly skilled; these personnel also include manufacturing engineers, computer programmers, and maintenance engineers. The most effective FMS applications have been in medium-volume batch production.

When a variety of parts is to be produced, FMS is suitable for production volumes of 15,000–35,000 parts per year; for individual parts that are of the same configuration, production may reach 100,000 units per year. In contrast, high-volume, low-variety parts production is best obtained from transfer machines (see *dedicated equipment*, Section 37.2.4). Low-volume, high-variety parts production can best be done on conventional standard machinery, with or without NC, or by using machining centers (Section 25.2).

Summary

- Integrated manufacturing systems are implemented to various degrees to optimize operations, improve product quality, and reduce costs.
- Computer-integrated manufacturing systems have become the most important means of improving productivity, responding to changing market demands, and enhancing the control of manufacturing and management functions.
- Widespread applications of powerful computers and software have led to advanced manufacturing system control, allowing mass customization to be considered. In such systems, large production rates can be achieved, with each part having custom features.
- Advances in holonic manufacturing, just-in-time production, and communications networks are all essential elements in improving productivity.
- Lean manufacturing is intended to identify and eliminate waste, leading to improvements in product quality, customer satisfaction, and decreasing product cost.
- The Internet of Things is a continuing trend, where objects will be connected to the Internet. In manufacturing, this involves connecting machines and sensors through the MTConnect communication protocol and using Big Data approaches.
- Artificial intelligence continues to create new opportunities in all aspects of manufacturing engineering and technology.
- Economic considerations in the design and implementation of computer-integrated manufacturing systems, especially flexible manufacturing systems, are particularly crucial because of the major capital expenditures required.

Key Terms

Artificial intelligence
Artificial neural networks
Big Data
Cellular manufacturing
Communications network
Communications standard
Computer-integrated manufacturing systems
Digital Twin
Environments
Ethernet
Expert systems
Flexible manufacturing
Flow
Framework systems
Holonic manufacturing
If-then rules
Inference engine
Internet of Things
Internet tools
Just-in-time production
Kanban
Knowledge engineering

Knowledge-based system
Lean manufacturing
Local area network
Machine vision
Manufacturing cell
Mass customization
Natural-language processing
Pallet
Pull system
Push system
Technical and office protocol
Value
Waste
Wireless local area networks
Zero inventory

Bibliography

Black, J.R., **Lean Production**, Industrial Press, 2008.
Black, JT., and Hunter, S.L., **Lean Manufacturing Systems and Cell Design**, Society of Manufacturing Engineers, 2003.
Conner, G., **Lean Manufacturing for the Small Shop**, 2nd ed., Society of Manufacturing Engineers, 2008.
Ertel, W., and Black, N.T., **Introduction to Artificial Intelligence**, Springer, 2017.
Geng, H., (ed.), **Internet of Things and Data Analytics**, Wiley, 2017.
Hall, J., and Scott, T., **Lean Six Sigma**, Amazon Digital, 2016.
Liker, J., and Convys, G.L., **The Toyota Way to Lean Leadership**, McGraw-Hill, 2011.
Mann, D., **Creating a Lean Culture**, 3rd ed., Productivity Press, 2014.
Meyers, F.E., and Stewart, J.R., **Motion and Time Study for Lean Manufacturing**, 3rd ed., Prentice Hall, 2001.
Myerson, P., **Lean Supply Chain and Logistics Management**, McGraw-Hill, 2012.
Ortiz, C.A., **Kaizen Assembly: Designing, Constructing, and Managing a Lean Assembly Line**, CRC Press, 2006.
Poole, D.L., and Mackworth, A.K., **Artificial Intelligence**, 2nd ed., Cambridge, 2017.
Rehg, J.A., **Introduction to Robotics in CIM Systems**, 5th ed., Prentice Hall, 2002.
Rehg, J.A., and Kraebber, H.W., **Computer-Integrated Manufacturing**, 3rd ed., Prentice Hall, 2004.
Russell, S., and Norvig, P., **Artificial Intelligence: A Modern Approach**, 3rd ed., Prentice Hall, 2009.
Wilson, L., **How to Implement Lean Manufacturing**, McGraw-Hill, 2009.
Zhang, Y., and Tai, F., **Optimization of Manufacturing Systems Using the Internet of Things,** Academic Press, 2016.

Review Questions

39.1. What is a manufacturing cell? Why was it developed?

39.2. Describe the basic principle of flexible manufacturing systems.

39.3. Why is a flexible manufacturing system capable of producing a wide range of lot sizes?

39.4. What are the benefits of just-in-time production? Why is it called a pull system?

39.5. Explain the function of a local area network.

39.6. What is an expert system?

39.7. What are the advantages of a communications standard?

39.8. What is MTConnect?

39.9. What is a WLAN? A PAN?

39.10. Describe your understanding of holonic manufacturing.

39.11. What is Kanban? Explain.

39.12. What is lean manufacturing?

39.13. What is pull?

39.14. In the lean manufacturing concept, what is the meaning of a value stream?

39.15. Describe the elements of artificial intelligence. Is machine vision a part of it? Explain.

Qualitative Problems

39.16. In what ways have computers had an impact on manufacturing? Explain.

39.17. What advantages are there in viewing manufacturing as a system? What are the components of a manufacturing system?

39.18. One restaurant makes sandwiches as they are ordered by customers. Another competing restaurant makes sandwiches in advance and sells them to customers from their inventory. Which is a pull, and which is a push system? Explain which restaurant makes the better sandwiches.

39.19. Discuss the benefits of computer-integrated manufacturing operations.

39.20. Why is just-in-time production required in lean manufacturing?

39.21. What drawbacks are there to just-in-time?

39.22. Would machining centers be suitable in just-in-time production? Explain.

39.23. Give an example of a push system and of a pull system. Indicate the fundamental difference between the two methods.

39.24. What are the advantages to having level production across lines and with respect to time?

39.25. Is there a minimum to the number of machines in a manufacturing cell? Explain.

39.26. Are robots always a component of an FMC? Explain.

39.27. Are there any disadvantages to zero inventory? Explain.

39.28. Review Table 36.1 and identify the points that are consistent with lean manufacturing.

39.29. Give examples in manufacturing processes and operations in which AI could be effective.

Synthesis, Design, and Projects

39.30. Think of a product line for a commonly used household item and design a manufacturing cell for making it. Describe the features of the machines and equipment involved.

39.31. What types of (a) products and (b) production machines would not be suitable for FMC? What design or manufacturing features make them unsuitable? Explain with examples.

39.32. Describe your opinions concerning the voice-recognition capabilities of future machines and controls.

39.33. Can a factory ever be completely untended? Explain.

39.34. Assume that you own a manufacturing company and that you are aware that you have not taken full advantage of the technological advances in manufacturing. However, now you would like to do so, and you have the necessary capital. Describe how you would go about analyzing your company's needs and how you would plan to implement these technologies. Consider technical as well as human aspects.

39.35. How would you describe the benefits of FMS to an older worker in a manufacturing facility whose experience has been running only simple machine tools?

39.36. Artificial neural networks are particularly useful where problems are ill-defined and the data are vague. Give examples in manufacturing where artificial neural networks can be useful.

39.37. It has been suggested by some that artificial intelligence systems ultimately will be able to replace the human brain. Do you agree? Explain.

39.38. Evaluate a process from a lean-production perspective. For example, observe the following closely, and identify, eliminate (when possible), or optimize the steps that produce waste in (a) preparing breakfast for a group of eight; (b) washing clothes or cars; (c) using Internet browsing software; and (d) studying for an exam, writing a report, or writing a term paper.

39.39. Pull can be achieved by working with one supplier and developing a balanced flow of products. However, it was stated that single-source suppliers should be avoided in the unforeseen events of natural disasters. Write a one-page paper explaining this paradox.

39.40. Explain how you can make your study habits more lean.

39.41. Conduct an Internet search and list five software packages that incorporate MTConnect.

39.42. Make a list of the number of devices you own that are connected to the Internet. How many of these existed ten years ago? Make a list of the devices you think you will own in ten years that will be connected to the Internet.

39.43. Evaluate a process from a lean production perspective. For example, closely observe the following and identify, eliminate (when possible), or optimize the steps that produce waste when:

1. Preparing breakfast for a group of eight.
2. Washing clothes or cars.
3. Using Internet browsing software.
4. Studying for an exam, or writing a report or a term paper.

39.44. Conduct a literature and Internet review, and summarize the basic principles of file encryption and cybersecurity.

39.45. In the Internet of Things, how would you connect people? Explain.

Chapter 40

Product Design and Manufacturing in a Competitive Environment

Examples:

- Manufacturing high-quality products at the lowest possible cost is critical in a global economy.
- This chapter describes the interrelated factors in product design, development, and manufacturing; it begins with a discussion of product design and life-cycle assessment in design and manufacturing.
- The importance of energy considerations in manufacturing and their role in production costs are then discussed.

- Material and process selection, together with their effects on design and manufacturing, are then described, followed by a discussion of the factors involved in the costs associated with a product.
- Finally, the principle of value analysis is described, along with a discussion of how it can help optimize manufacturing operations and minimize product cost.

40.1 Introduction

In an increasingly competitive global marketplace, manufacturing high-quality products at the lowest possible cost requires an understanding of the often complex interrelationships among numerous factors. It was indicated throughout this text that:

1. Product design and selection of materials and manufacturing processes are interrelated
2. Designs are periodically modified to (a) improve product performance, (b) strive for zero-based rejection and waste, (c) design products to make them easier and faster to manufacture, and (d) consider new materials and processes that are continually being developed.

Because of the increasing variety of materials and manufacturing processes now available, the task of producing a high-quality product by selecting the best materials and the best processes while, at the same time, minimizing costs, continues to be a major challenge, as well as an opportunity. The term **world class** is widely used to indicate high levels of product quality, signifying the fact that products must meet international standards and be marketable and acceptable worldwide. World-class status, like product quality, is not a fixed target for a company to reach but rather a *moving target*, also known as **continued improvement**, rising to higher and higher levels. The *selection of materials* for products traditionally has required much experience; however, several databases and expert systems are now widely available that facilitate the selection process. Moreover, in reviewing the materials used in existing products, it readily becomes apparent that there are numerous opportunities for the *substitution of materials* for improved performance and, especially, cost savings.

In the production phase of a product, the *capabilities of manufacturing processes* must be properly assessed as an essential guide to the ultimate selection of an appropriate process or sequence of processes. As described throughout various chapters, there usually is more than one method of manufacturing the individual components of a product.

Increasingly important are also *life-cycle* assessment and *life-cycle engineering* of products, services, and systems, particularly regarding their potentially adverse global impact on the environment. The major emphasis now is on *sustainable manufacturing*, with the purpose of reducing or eliminating any and all adverse effects of manufacturing on the environment, while still allowing companies to continue to be profitable.

Although the *economics* of individual manufacturing processes has been described throughout the book, this chapter takes a broader view and summarizes the overall manufacturing cost factors. It also introduces cost-reduction methods, including value analysis, a powerful tool to evaluate the cost of each manufacturing step relative to its contribution to a product's value.

40.2 Product Design

Design for manufacture and assembly (DFMA) and *competitive aspects* of manufacturing have been highlighted throughout this text. Several guidelines for the selection of materials and manufacturing processes are given in the references listed in Table 40.1. Major advances are continually being made in design for manufacture and assembly, for which several software packages are now available. Although their use requires considerable training, these advances greatly help designers develop high-quality products with fewer components, thus reducing production time and assembly as well as product cost.

40.2.1 Product Design Considerations

In addition to the design guidelines regarding individual manufacturing processes, there are general product design considerations (see also *robust design*, Section 36.5.1). Designers often must check and verify whether they have addressed considerations such as:

1. Have all alternative designs been thoroughly investigated?
2. Can the design be simplified and the number of its components minimized without adversely affecting its intended functions and performance?
3. Can the design be made smaller and lighter?
4. Are there unnecessary features in the product and, if so, can they be eliminated or combined with other features?
5. Have modular design and building-block concepts (see, for example, Section 25.2.4) been considered for a family of similar products and for servicing, repair, and upgrading?
6. Are the specified dimensional tolerances and surface finish unnecessarily tight, thereby significantly increasing product cost, and if so, can they be relaxed without any significant adverse effects?
7. Will the product be difficult or excessively time consuming to assemble or to disassemble for maintenance, servicing, or recycling of some or all of its components?
8. Is the use of fasteners minimized, including their quantity and type variety?
9. Have environmental considerations been taken into account and incorporated into product design and material and process selection?
10. Have green design and life-cycle engineering principles been applied, including recycling and cradle-to-cradle considerations?
11. Can any of the design or manufacturing activities be outsourced, and can any currently outsourced activities be reshored?

40.2.2 Product Design and Quantity of Materials

Significant reductions in the *quantity* of materials required can be achieved by approaches such as (a) reducing the component's size or volume or (b) using materials with higher strength-to-weight or stiffness-to-weight ratios (see Fig. 3.2). The latter can be attained by improving and optimizing the product design and by selecting different cross-sections, such as those having a high moment of inertia (e.g., I-beams and channels) or by using tubular or hollow components instead of solid sections.

Implementing design changes may, however, present significant challenges in manufacturing. Consider, for example, the following:

1. Casting or molding thin cross-sections can present difficulties in die and mold filling and in meeting specified dimensional accuracy and surface finish (Section 12.2).
2. Forging of thin sections requires higher forces and can present difficulties (Section 14.3).
3. Impact extrusion of thin-walled parts can be difficult, especially when high dimensional accuracy and part symmetry are required (Section 15.4.1).
4. The formability of sheet metals may be reduced as their thickness decreases, also possibly leading to buckling of the part (Section 16.3).

Table 40.1: References to Various Topics in This Book (Page numbers are in parentheses)

Material Properties
- Tables 2.1 (83), 2.2 (86), and 2.3 (90), and Figs. 2.5, 2.6, 2.7, 2.15, 2.16, 2.17, and 2.17
- Tables 3.1 (125) and 3.2 (126), and Figs. 3.1, 3.2, 3.3
- Tables 5.2 (173), 5.3 (175), 5.4 (177) and 5.5 (178) 5.6 through 5.8 (180-182), Fig. 5.6
- Tables 6.3 through 6.10 (191–199)
- Tables 7.1 (212), 7.2 (220), and 7.3 (225)
- Tables 8.1 (240), 8.2 (244), and 8.3 (248)
- Tables 9.1 (261), 9.2 (262), and 9.3 (270), and Figs. 9.3, 9.5, 9.7
- Table 10.1 (295)
- Table 11.3 (323)
- Tables 12.3 (353), 12.4 (354), and 12.5 (354), and Fig. 12.5
- Tables 16.2 (454), 16.3 (460), and 16.4 (472), and Fig. 16.14
- Tables 17.3 (528), 17.4 (529), and 17.5 (529), and Fig. 17.11
- Table 20.2 (604)
- Tables 22.1 (676), 22.2 (677), 22.3 (677), and 22.5 (682), and Figs. 22.1 and 22.9
- Table 26.1 (816)
- Table 32.3 (1046)

Manufacturing Characteristics of Materials
- Table I.4 (17)
- Table 4.1 (156)
- Table 5.8 (182)
- Table 6.2 (190)
- Tables 12.1 (346) and 12.6 (356)
- Table 14.3 (402)
- Table 16.3 (460) and Fig. 16.34
- Tables 17.1 (517) and 17.2 (524)
- Tables 21.1 (639) and 21.2 (651)
- Fig. 22.1

Dimensional Tolerances and Surface Finish
- Table 11.2 (323)
- Table 23.1 (702) and Figs. 23.14 and 23.15
- Fig. 25.17
- Fig. 33.5
- Figs. 35.20 and 35.21
- Figs. 40.4 and 40.5

Capabilities of Manufacturing Processes
- Tables 11.1 (272) and 11.2 (275)
- Table III.1 (351)
- Tables 14.1 (390) and 14.4 (407)
- Table 16.1 (444)
- Section 17.7 and Fig. 17.15
- Table 18.1 (541)
- Tables 19.1 (561) and 19.2 (594)
- Table 20.1 (603)
- Tables 23.1 (702), 23.6 (711), 23.8 (720), and 23.9 (721)
- Tables 26.3 (830) and 26.4 (831)
- Table 27.1 (859)
- Tables 28.1 (901), 28.2 (910), and 28.3 (911), and Fig. 28.20
- Table 29.1 (954)
- Table VI.1 (1006)
- Table 30.1 (973)
- Table 32.4 (1047)
- Table 34.1 (1106)
- Table 37.2 (1173) and Fig. 37.3
- Table 39.1 (1238)
- Tables 40.3 (1263) and 40.5 (1270), and Figs. 40.2, 40.3, 40.4, and 40.5

Design Considerations in Processing
- Abrasive processes: Section 26.5
- Advanced machining: Various sections in Chapter 27
- Casting: Section 12.2
- Ceramics shaping: Section 18.5
- Forging: Section 14.6
- Heat Treating: Section 4.13
- Joining processes: Various sections in Chapters 30–32
- Machining: Sections in Chapters 23–24
- Polymers processing: Section 19.15
- Powder metallurgy: Section 17.6
- Sheet-metal forming: Section 16.14

Costs and Economics
- Tables I.5 (32), I.6 (33), I.7 (34), and Section I.10
- Table 6.1 (189)
- Section 12.4
- Section 14.9
- Section 16.16
- Table 17.6 (532) and Section 17.7
- Table 19.2 (594) and Section 19.16
- Section 25.8
- Section 26.9 and Fig. 26.37
- Section 27.11
- Section 31.8
- Section 32.7
- Section 37.11
- Section 39.10
- Table 40.6 (1274) and Section 40.10

5. Machining and grinding of thin workpieces may lead to part distortion, poor dimensional accuracy, and vibration and chatter (Section 25.4); consequently, advanced machining processes may have to be considered (Chapter 27).

6. Welding thin sheets or slender structures can cause significant distortion of parts and structures (Section 30.10).

Conversely, making parts with *thick* cross-sections can have their own adverse effects:

1. The production rate in die casting (Section 11.4.5) and injection molding (Section 19.3) can become slower because of the increased cycle time required to allow sufficient time for the thicker regions to cool before removing the part from the die or mold.

2. Porosity can develop in thicker regions of castings (Fig. 10.16).

3. In die-cast parts, thinner sections will have lower strength per unit thickness as compared with strength in thinner sections (Section 11.4.5).

4. Processing plastic parts requires increased cycle times as their thickness or volume increases, because of the longer time required for the parts to cool sufficiently to be removed from their molds (Chapter 19).

5. The bendability of sheet metals decreases as their thickness increases (Section 16.5).

6. In parts made by powder metallurgy, there can be significant variations in density, hence properties, throughout regions with varying thickness (Section 17.6).

7. Welding thick sections can present difficulties in the quality of the welded joint (Section 30.10).

40.3 Product Quality

Recall that the word *quality* (Section 36.2) is difficult to define precisely, largely because it includes not only well-defined technical characteristics but also human, and hence subjective, opinions. Generally, however, a high-quality product is considered to have at least the following characteristics:

1. High reliability

2. Perform the required functions well and safely

3. Good appearance

4. Inexpensive

5. Upgradeable

6. Available in the quantities desired when needed

7. Robust over their intended life.

A major priority in product quality is the concept of **continuous improvement**, as exemplified by the Japanese term **kaizen**, meaning *never-ending improvement* (Section 36.1). Recall also that the *level* of quality a manufacturer chooses to impart to a particular product depends on the market for which the product is intended. Low-quality, low-cost products, such as hand tools, have their own market niche, just as there is a market for high-quality and high-end products, such as a Rolls-Royce automobile, a diamond-studded gold watch, precision machine tools and measuring instruments, and high-performance sports equipment.

Return on Quality. In implementing quality into products, it is important to understand the concept of *return on quality* (ROQ), because of the following considerations:

- Quality must be viewed as an investment because of its long-term major influence on customer satisfaction.

- Any improvement in quality must be investigated with respect to the additional costs involved.

- There must be a certain limit to how much a manufacturer should spend on incremental quality improvements, especially when quality can be rather subjective.

Although customer satisfaction can be difficult to assess and include in cost calculations, when their satisfaction increases, customers are more likely to be retained and become repeat customers when there are no significant defects in the products they purchase. It is also important to note the fact that the relative costs involved in identifying and repairing defects in products grow by orders of magnitude, in accordance with the *rule of ten*, as shown in Table I.5.

40.4 Life-cycle Assessment and Sustainable Manufacturing

Life-cycle assessment (LCA) is defined, according to the ISO 14000 standard, as "a systematic set of procedures for compiling and examining the inputs and outputs of materials and energy, and the associated environmental impacts or burdens directly attributable to the functioning of a product, process, or service system throughout its entire life cycle." **Life cycle** involves consecutive and interlinked stages of a product or a service, from the very beginning of design and manufacture to its recycling or disposal; it includes:

1. Extraction of natural resources

2. Processing of raw materials

3. Manufacturing of products

4. Transportation and distribution of the product to the customer

5. Use, maintenance, and reuse of the product

6. Recovery, recycling, and reuse of components of the product or their disposal, which also include metalworking fluids, cleaning solvents, and various liquids used in heat-treating and plating processes.

A product typically has several components made of a variety of metallic and nonmetallic materials, processed into individual parts, and then assembled. Thus, each component has its own life cycle, such as the tires in an automobile, rubber washers in faucets, bulbs in light fixtures, and belts in vacuum cleaners. Moreover, (a) some products, particularly those made of paper, cardboard, glass, or inexpensive plastics, are intentionally made to be *disposable* but nonetheless are now recyclable and (b) numerous other products are completely reusable.

Life-cycle Engineering. The major aim of *life-cycle engineering* (LCE), also called **green design** or **green engineering**, is to consider *reusing* and *recycling* the components of a product, beginning with the earliest stage of product design (see Fig. I.2). Although life-cycle analysis and engineering are comprehensive and powerful tools, their implementation can be challenging, time consuming, and costly, largely because of (a) uncertainties regarding materials, processes, and long-term effects, (b) in the input data and (c) the time required to collect reliable data to properly assess the often-complex interrelationships among the numerous components of the whole system. Software is being developed to expedite these analyses, especially for the chemical and manufacturing industries, because of the higher potential for environmental and ecological impact. Examples of such software include FeaturePlan and Teamcenter, which run in a ProEngineer environment.

Cradle-to-cradle. The terms and concepts of *cradle-to-cradle* design (coined by W.R. Stahel in the 1970s and also called *CRC* or *regenerative design*) and *cradle-to-grave* (which ends with the disposal phase of products) are described in Section I.4. These are basically a holistic model for human activity with due respect for life and the well-being of future generations. (See also Braungart and McDonough in the Bibliography at the end of this chapter.)

Sustainable Manufacturing. As universally acknowledged, the natural resources on the planet Earth are limited, necessitating conservation of both materials and energy. The concept of *sustainable manufacturing* emphasizes the need for conserving resources, particularly through proper maintenance and reuse. Sustainable manufacturing is meant to meet the main purposes of (a) increasing the life cycle of products, (b) eliminating harm to the environment and the ecosystem, and (c) ensuring our collective well-being.

Example 40.1 Sustainable Manufacturing of Nike Athletic Shoes

Among numerous examples, the production of Nike shoes illustrates the benefits of sustainable manufacturing. Athletic shoes are assembled using adhesives (Section 32.4). Up to around 1990, the adhesives used contained petroleum-based solvents, which pose health hazards to humans and contribute to petrochemical smog. The company cooperated with chemical suppliers to successfully develop a water-based adhesive technology, now used in the majority of shoe-assembly operations. As a result, solvent use in all manufacturing processes in Nike's subcontracted facilities in Asia has been greatly reduced.

The rubber outsoles of the athletic shoes are made by a process that results in significant amounts of extra rubber around the periphery of the sole, called *flashing* (similar to the *flash* shown in Figs. 14.6d and 19.17). With about 40 factories using thousands of molds and producing over a million outsoles a day, the flashing constitutes the largest chunk of waste in manufacturing of these shoes.

In order to reduce this significant waste, the company developed a technology that grinds the flashing into 500-μm rubber powder, which is then added back into the rubber mixture needed to make the outsole. With this approach, waste was reduced by 40%. Moreover, it was found that the mixed rubber had better abrasion resistance and durability, and its overall performance was higher than the best premium rubber.

40.5 Energy Consumption in Manufacturing

The manufacturing sector consumes approximately one-fourth of the annual global energy production; this number has fallen since it peaked at around 50% in the 1970s, because of major efforts to reduce waste and improve the efficiency of machinery and manufacturing operations. By far, the most common source of energy in manufacturing is electrical, produced from oil, natural gas, biofuels, coal, nuclear, wind, wave, and solar. Given such a large percentage and varied sources of energy, all of the concerns regarding energy availability and conservation must be considered and addressed in manufacturing. Indeed, it is not likely that viable national energy policies can be developed and implemented without a central consideration of the manufacturing sector.

Process Energy Demand. The energy required to produce a particular part or component is determined, to a large extent, by its design, material, and the manufacturing operation, as well as the quality, condition, and age of the machinery and equipment. The energy requirement is relatively easy to calculate for a manufacturing process (see, for instance, Table 21.2), but becomes difficult when ancillary equipment is included in the final calculation. For example, pumps, fans, blowers, furnaces, and lights are all involved in manufacturing operations. Moreover, the efficiency of operation typically varies depending on a particular plant's practices and procedures or from company to company. For example:

- Some manufacturing operations are more demanding from an energy standpoint than others, as shown in Fig. 40.1.

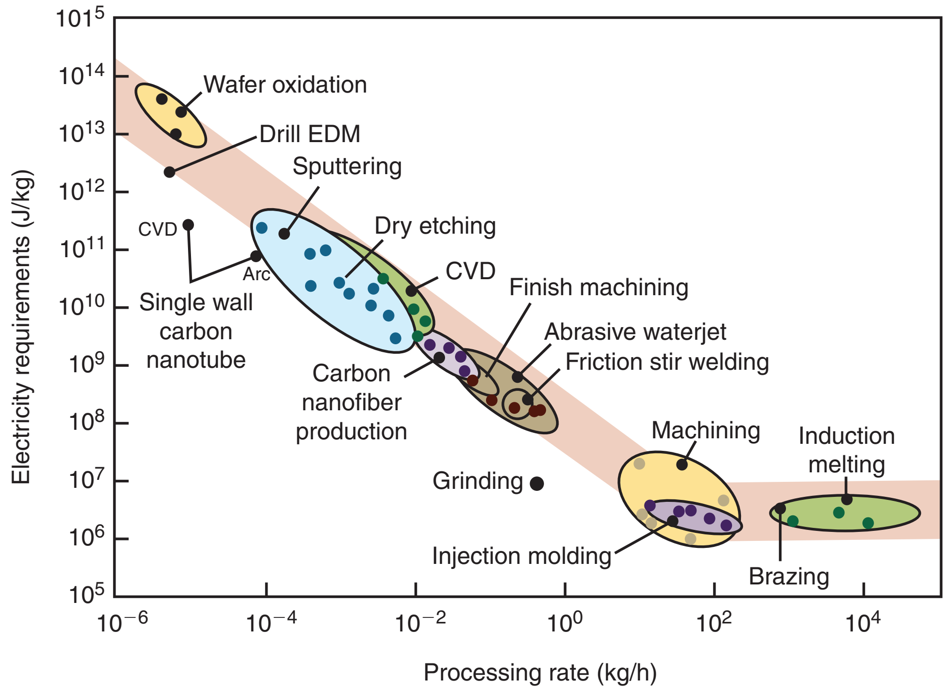


Figure 40.1: Specific energy requirements for various manufacturing processes, including ancillary equipment. Note that most manufacturing processes lie within the dark band. Also note the increased energy intensity for slower processes. *Source:* After T. Gutowski.

- Each manufacturing process has a range of performance; processing rates can vary greatly depending on, for example, the specified tolerances and surface finish specified, with tighter tolerances and smoother surfaces being the most time consuming and energy intensive (Figs. 40.2 and 40.3).
- Energy requirements for some processes are strongly related to the sequence of operations performed, as when, for example, a machine is "ready" but not yet actively processing the material. Examples are (a) injection molding (Section 19.3), when the mold has been installed in the machine but it has not yet been filled with the polymer or (b) when a workpiece is being repositioned in a fixture on a machining center (Section 25.2).
- Some machinery have continuously-operating hydraulic pumps, while others will shut the pumps off during periods of inactivity.
- Workpieces in hot-working operations can be cooled by blowing cool air over them and, in practice, the hot air is often vented to the atmosphere. The resulting hot air can, however, be used to preheat stock material for processing, heat the facility, or provide hot water.

Effect of Materials. Whether a material is mined, refined, and cast; synthesized from chemicals, or recycled from discarded products, energy is required to produce stock materials into forms that can be further processed. In many applications, it is noted that certain materials have a performance advantage over others. For example, titanium and aluminum alloys are obviously preferred to steel for aircraft primarily because of their lower strength-to-weight and stiffness-to-weight ratios (see Fig. 3.2), allowing for lighter designs and associated fuel savings. It is important to note, however, that different materials have very different energy

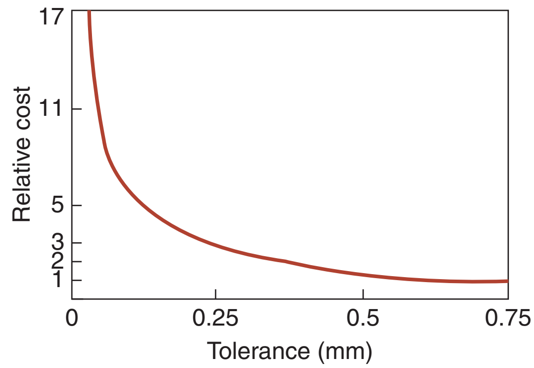


Figure 40.2: Dependence of manufacturing cost on dimensional tolerances.

requirements and recycled materials have significantly less energy needs. Table 40.2 summarizes the energy required to produce various materials, and presents the data by mass and by volume. It has been noted that if energy or *carbon footprint* is divided by mass, there is a natural benefit to using heavier materials that may not be justified in weight-constrained problems. On the other hand, if volume is a constraint, then the energy per unit weight may be a fair measure.

Note in Table 40.2 that metals require significant energy to be produced; they are generally reduced or extracted from their oxides, a process that is energy intensive. It has been estimated that 5% of the total energy consumption in the United States is used to produce aluminum (Section 6.2). It is not uncommon for the energy required to produce a material to be three or four orders of magnitude greater than the energy required to form it into its final shape.

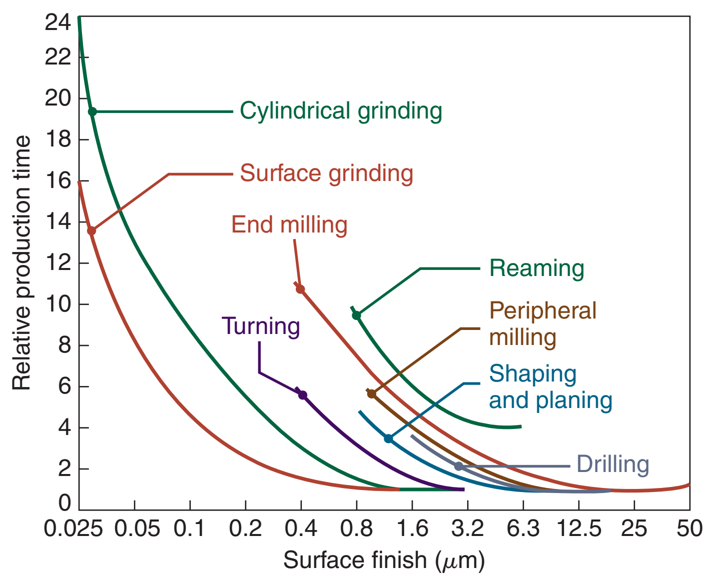


Figure 40.3: Relative production time as a function of surface finish produced by various manufacturing processes (see also Fig. 26.35).

Table 40.2: Energy Content of Selected Materials.

	Energy content	
Material	**MJ/kg**	**GJ/m^3**
Metals		
Aluminum		
From bauxite	300	810
Recycled	42.5	115
Cast Iron	30–60	230–460
Copper		
From ore	105	942
Recycled	55.4	497
Lead	30	330
Magnesium	410	736
Steel		
From ore	55	429
Recycled	9.8	76.4
Zinc	70	380
Polymers		
Nylon 6,6	175	200
Polyethylene		
High density	105–120	100–115
Low density	80–100	75–95
Polystyrene	95–140	95–150
Polyvinyl chloride	67–90	90–150
Ceramics	1–50	4–100
Glasses	10–25	30–60
Wood	1.8–4.0	1.2–3.6

It is therefore understandable that the use of aluminum in automobiles to reduce energy intensity is indeed difficult to justify without the sustained implementation of recycling. Energy savings from weight reductions may not be appreciably higher, and may even be lower, than the energy required to produce aluminum.

40.6 Material Selection for Products

In selecting materials for a product, it is essential to have a clear understanding of the *functional requirements* for each of its individual components. The general criteria for selecting materials were described in Section I.5 of the General Introduction; this section discusses them in more specific detail.

General Properties of Materials. *Mechanical properties* (Chapter 2) include strength, toughness, ductility, stiffness, hardness, and resistance to fatigue, creep, and impact. *Physical properties* (Chapter 3) include density, melting point, specific heat, thermal and electrical conductivity, thermal expansion, and magnetic properties. *Chemical properties* of primary concern in manufacturing are their susceptibility to oxidation and corrosion and their response to various surface-treatment processes (Chapter 34). The cost of corrosion, although largely hidden, is estimated to be on the order of 3% of the U.S. gross domestic product.

Material selection has become easier and faster because of the increasing availability of extensive computer databases that provide greater accessibility and accuracy. To facilitate the selection of materials, **expert-system software** (called **smart databases**, Section 39.9) have been developed. These systems are capable of identifying appropriate materials for a specific application, just as an expert or a team of experts would.

Table 40.3: Shapes of Commercially Available Materials.

Material	Available as
Aluminum	B, F, I, P, S, T, W
Ceramics	B, p, s, T
Copper and brass	B, f, I, P, s, T, W
Elastomers	b, P, T
Glass	B, P, s, T, W
Graphite	B, P, s, T, W
Magnesium	B, I, P, S, T, w
Plastics	B, f, P, T, w
Precious metals	B, F, I, P, t, W
Steels and stainless steels	B, I, P, S, T, W
Zinc	F, I, P, W

Note: B = bar and rod, F = foil, I = ingots, P = plate and sheet, S = structural shapes, T = tubing, and W = wire; all metals are generally available as powders. Lowercase letters indicate limited availability.

Shapes of Commercially Available Materials. After selecting appropriate materials, the next step is to determine the shapes and the sizes in which these materials are available commercially (Table 40.3). Depending on the type, materials generally are available as castings, extrusions, forgings, drawn rod and wire, rolled bars, plates, sheets, foil, and powder metals.

Purchasing materials in shapes that require the least amount of additional processing obviously is a major economic consideration; also relevant are such characteristics as surface finish, dimensional tolerances, straightness, and flatness (see, for example, Figs. 27.4, 23.14, and 33.5, and Table 11.3). The better and the more consistent these characteristics are, obviously the less additional processing will be required. If, for example, simple round shafts with good dimensional accuracy, roundness, straightness, and surface finish are desired, then round bars that are first turned or drawn and centerless-ground (Fig. 26.23) to the dimensions specified could be purchased.

Unless the facilities in a plant have the capability of economically producing round bars, it generally is less expensive to purchase them. If a stepped shaft (having different diameters along its length, as shown in Fig. IV.3) is required, a round bar with a diameter at least equal to the largest diameter of the stepped shaft could be purchased, then turned on a lathe or processed and formed by some other means in order to reduce its diameter.

Each manufacturing operation produces a part that has a specific shape, surface finish, and dimensional accuracy; consider the following examples:

- Castings generally have lower dimensional accuracy and a poorer surface finish than parts made by cold forging, cold extrusion, or powder metallurgy.

- Hot-rolled or hot-drawn products generally have a rougher surface finish and larger dimensional tolerances than cold-rolled or cold-drawn products.

- Extrusions have smaller tolerances than parts made by roll forming of sheet metal.

- Round bars machined on a lathe have a rougher surface finish than similar bars that are ground.

- Seamless tubing made by the tube-rolling process has more thickness variation than that of roll-formed and welded tubing (Chapters 6 and 12).

- The wall thickness of welded tubing is generally more uniform than that of seamless tubing.

Manufacturing Characteristics of Materials. *Manufacturing characteristics* generally include castability, workability, formability, machinability, weldability, and hardenability by heat treatment. Recall also that the quality of the raw material (stock) can greatly influence its manufacturing properties. The following are typical examples:

- A bar with a longitudinal seam or lap will develop cracks during even simple upsetting or heading operations.
- Rods, with internal defects such as hard inclusions, will crack during further processing.
- Porous castings will develop a poor surface finish when subsequently machined for better dimensional accuracy.
- Parts that are heat treated nonuniformly or cold-drawn bars that are not properly stress relieved will distort during subsequent processing.
- Incoming stock that has significant variations in its composition and microstructure cannot be heat treated or machined consistently and uniformly.
- Sheet metal having variations in its cold-worked conditions will exhibit different degrees of springback during subsequent bending and other forming operations.
- If prelubricated sheet-metal blanks are supplied with nonuniform lubricant thickness and distribution, their formability, surface finish, and overall quality in subsequent stamping operations will be adversely affected.

Reliability of Material Supplies. Several factors influence the *reliability of material supplies*: shortages of materials, strikes, geopolitics, and the reluctance of suppliers to produce materials in a particular shape or quality. Moreover, even though raw materials may generally be available throughout a country as a whole, they may not readily be available at a particular plant's location.

Recycling Considerations. *Recycling* may be relatively simple for products such as scrap metal, plastic bottles, and aluminum cans; often, however, the individual components of a product must be taken apart and separated into groups. Obviously, if much effort and time has to be expended in doing so, recycling can indeed become prohibitively expensive. Some general guidelines to facilitate the process during the life cycle of a product are:

- Reduce the number of parts and types of materials in products
- Reduce the variety of product models
- Use a modular design to facilitate disassembly of a product
- For products made of plastic, use single types of polymers as much as possible
- Mark plastic parts for ease of identification
- Avoid using coatings, paints, and plating; instead, use molded-in colors in plastic parts
- Avoid using adhesives, rivets, and other permanent joining and assembly methods in; instead, use fasteners, especially snap-in fasteners.

As one example of such an approach to recycling, one manufacturer of laser-jet printers reduced the number of parts in a cartridge by 32% and the variety of plastic materials used by 55%.

Cost of Materials and Processing. Because of its processing history, the *unit cost* of a raw material (typically per unit weight) depends not only on the material itself, but also on its shape, size, and condition. Also, for

Table 40.4: Approximate Scrap Produced in Various Manufacturing Operations.

Process	Scrap (%)	Process	Scrap (%)
Machining	10–80	Permanent-mold casting	10
Hot forging	20–25	Powder metallurgy	<5
Sheet-metal forming	10–25	Rolling	<1
Hot extrusion	15		

example, (a) because more operations are involved in the production of thin wire than for a round rod, its unit cost is much higher; (b) powder metals generally are more expensive than bulk metals; and (c) the cost of materials typically decreases as the quantity purchased increases, especially for automotive companies that purchase materials in very large quantities.

The cost of a particular material is subject to fluctuations caused by factors as simple as supply and demand or as complex as geopolitics. If a material used in a product is no longer cost competitive, alternative materials have to be selected. For example, (a) the copper shortage in the 1940s led the U.S. government to mint pennies from zinc-plated steel (see Table I.2) and (b) when the price of copper increased substantially during the 1960s, electrical wiring in homes was switched to aluminum; however, this substitution led to the redesign of terminals of switches and outlets in order to avoid excessive heating, because aluminum has a higher contact resistance at the junctions than copper.

Scrap. The value of scrap (Table 40.4) is deducted from a material's cost in order to obtain the net material cost. Its value depends on the type of materials and on demand for it, and typically it is 10–40% of the original cost of the material. Note that in machining, scrap can be very high, whereas such processes as rolling, ring rolling, and powder metallurgy produce the least scrap, and hence they are called net- or near-net-shape processes.

40.7 Material Substitution

Substitution of materials plays a major role in the economics of manufacturing products. Automobile and aircraft manufacturing are typical examples of major industries in which the substitution of materials is an ongoing activity; a similar trend is evident in sporting goods and numerous other products. There are several reasons for substituting materials in products:

1. Reduce the costs of materials and processing them
2. Improve the efficiency of manufacturing and assembly operations
3. Improve the performance of the product, such as by reducing its weight and by improving its resistance to wear, fatigue, and corrosion
4. Increase stiffness-to-weight and strength-to-weight ratios
5. Reduce the need for periodic maintenance and repair
6. Reduce vulnerability to the unreliability of the supply of materials
7. Improve compliance with legislation and regulations prohibiting the use of certain materials, especially for health reasons
8. Improve robustness to reduce variations in performance or environmental sensitivity of the product
9. Improve ease of recycling.

Substitution of Materials in the Automotive Industry.

- Certain components of the metal body replaced with plastic or reinforced-plastic parts
- Metal bumpers, gears, pumps, fuel tanks, housings, covers, clamps, and various other components replaced with plastics or composites
- Carbon-steel chassis pillars replaced by TRIP or TWIP steels
- Structural steel components replaced by aluminum alloys
- Metallic engine components replaced with ceramic and composites parts
- All-metal drive shafts replaced with composite-material drive shafts
- Cast-iron engine blocks changed to cast-aluminum, forged crankshafts to cast crankshafts, and forged connecting rods to cast, powder-metallurgy, or composite-material connecting rods.

Substitution of Materials in the Aircraft and Aerospace Industries.

- Conventional aluminum alloys (particularly 2000 and 7000 series) replaced with aluminum–lithium alloys, titanium alloys, polymer-reinforced composites, and glass-reinforced aluminum because of the higher strength-to-weight ratios of these materials
- Forged parts replaced with powder-metallurgy parts that are manufactured with better control of impurities and microstructure; the powder-metallurgy parts also require less machining and produce less scrap of expensive materials
- Advanced composite materials and honeycomb structures replacing traditional aluminum airframe components, and metal-matrix composites are replacing some of the aluminum and titanium in structural components.

Example 40.2 Material Substitution in Products

In the following list, the commonly available products can be made of either set of materials mentioned:

1. Baseball bat: Metal versus wood
2. Hammer: Metal versus reinforced-plastic or wood handle
3. Engine intake manifold: Metal versus plastic
4. Lawn chair: Cast-iron versus aluminum
5. Light-switch plate: Plastic versus sheet metal
6. Aircraft fuselages: Aluminum versus composite materials
7. Pneumatic fittings: Zinc versus copper.

More recent examples of possible material substitutions are the following, some of which have been built or and some are under various stages of consideration:

8. Automobile windows: Glass versus polycarbonate
9. Pedestrian bridge: Steel versus titanium

10. Bicycle: Metal versus hardwood
11. Automobile bodies: Steel versus aluminum
12. Truck bodies: Steel versus stainless steel
13. Guitar: Wood versus plastic or aluminum
14. Armor plate: Aluminum versus steel or other high-strength alloys
15. Connections in IC packages: Copper versus gold.

40.8 Manufacturing Process Capabilities

Process capability is the ability of a particular manufacturing process to produce, under controlled production conditions, defect-free parts within certain limits of precision (see also Section 36.8.2). The capabilities of several manufacturing processes regarding their dimensional limits are shown in Fig. 40.4. Note, for example, that sand casting cannot produce thin parts, whereas cold rolling is capable of producing very thin materials, as also evidenced by a common product such as aluminum foil.

Equally important are the capabilities of various processes to meet stringent dimensional tolerance and surface-finish requirements, as shown in Fig. 40.5. Note, for example, how sand casting is at the extreme opposite corner of microfabrication, described in Chapters 28 and 29. The importance of emphasizing the term *under controlled conditions* can be appreciated when one views the size of the envelopes in the figure. Note, for instance, the large envelope for machining and finishing operations, with tolerance and roughness boundaries that, for a variety of reasons, span three orders of magnitude. Thus, if a turning operation is carried out on a poorly maintained old lathe and using inappropriate tools and processing parameters, then the tolerances and surface finish will of course be poor.

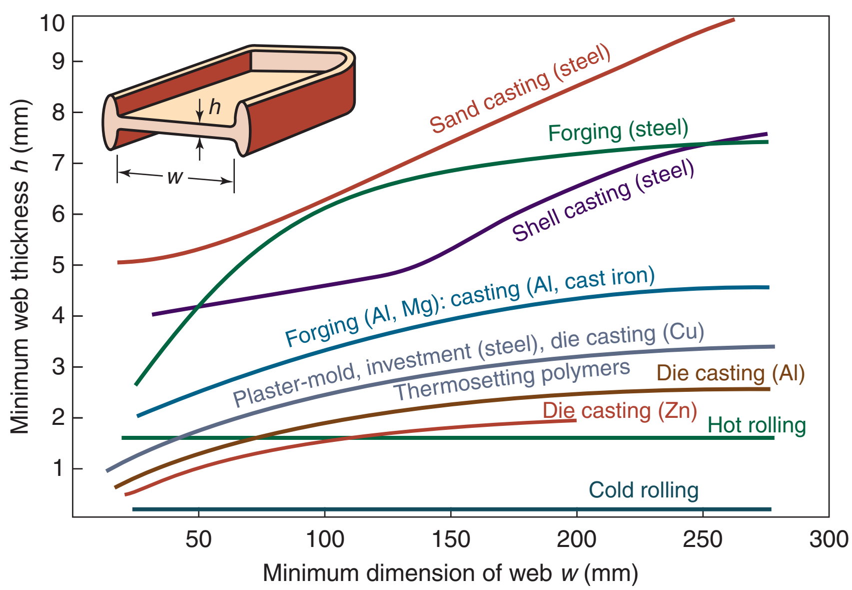


Figure 40.4: Manufacturing process capabilities for minimum part dimensions. *Source:* After J.A. Schey.

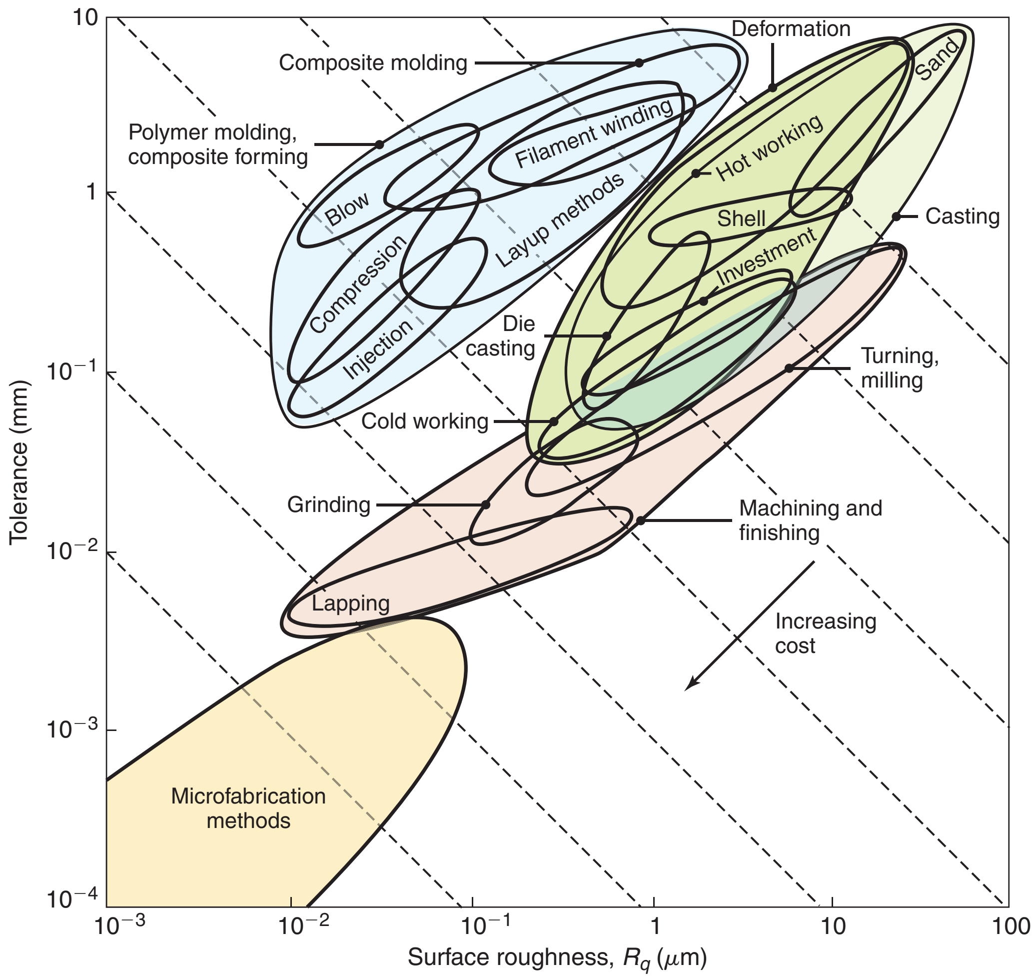


Figure 40.5: A plot of achievable tolerance versus surface roughness for assorted manufacturing operations. The dashed lines indicate cost factors, where an increase in precision corresponding to the separation of two neighboring lines gives an increase in cost for a given process (within a factor of two). *Source:* M.F. Ashby, *Materials Selection in Design*, Butterworth-Heinemann, 2010.

Dimensional Tolerances and Surface Finish. The *dimensional tolerances* and *surface finish* produced are particularly important in (a) subsequent assembly operations, because of possible difficulties in fitting the parts together during assembly and (b) the proper operation of machines and instruments, because their performance will affect tolerances and finish.

Closer tolerances and better surface finish can be achieved by subsequent and additional finishing operations (see Section 26.7) but at higher cost. Moreover, the finer the surface finish specified, the longer is the manufacturing time. For example, it has been observed that in machining aircraft structural members, made of titanium alloys, as much as 60% of the cost of machining may be expended in the final machining pass. Thus, unless otherwise required, parts should be specified with as rough a surface finish and as wide a dimensional tolerance as functionally and aesthetically will be acceptable.

Production Quantity. Depending on the type of product, the production quantity, also known as *lot size*, varies widely. Bearings, bolts and nuts, metal or plastic containers, tires, automobiles, and lawn mowers, for example, are produced in very large quantities, whereas jet engines, diesel engines, locomotives, and medical equipment are produced in much more limited quantities. Production quantity also plays a significant role in process and equipment selection. In fact, an entire manufacturing discipline, called **economic order quantity**, is devoted to mathematically determining the optimum production quantity.

Production Rate. An important factor in manufacturing process selection is the *production rate*, defined as the number of pieces to be produced per unit of time (hour, day, or year). The rate obviously can be increased by using multiple equipment and highly automated machinery. Die casting, powder metallurgy, deep drawing, wire drawing, and roll forming are high-production-rate operations; by contrast, sand casting, electrochemical machining, superplastic forming, adhesive and diffusion bonding, and the processing of reinforced plastics generally are relatively slow operations.

Lead Time. *Lead time* is generally defined as the length of time between the receipt of an order for a product and its delivery to the customer. Depending on part size, material, and shape complexity precision of the dies required, the lead time for such processes as forging, extrusion, die casting, roll forming, and sheet-metal forming can range from weeks to months. By contrast, processes such as machining, grinding, and advanced material-removal have significant built-in flexibility, due to the fact that they utilize machinery and tooling that can readily be adapted to most production requirements. Recall that machining centers, flexible manufacturing cells, and flexible manufacturing systems are all capable of responding rapidly and effectively to product changes and to production quantities (see also *additive manufacturing* in Chapter 20).

Robustness of Manufacturing Processes and Machinery. *Robustness* was described in Section 36.5.1 as characterizing a design, a process, or a system that continues to function within acceptable parameters despite variabilities in its environment. To appreciate its importance in manufacturing processes, consider a situation in which a simple plastic gear is being produced by injection molding, but it has been noticed that there are significant and unpredictable variations in quality as the gears are being produced. There are several well-understood variables and parameters in injection molding, including the effects of raw-material quality, speed of operation, and temperatures within the system. All these are independent variables, and hence they can be controlled.

There are, however, certain other variables, called **noise**, that are largely beyond the control of the operator, such as (a) variations in ambient-temperature and humidity in the plant throughout the day, (b) dust in the air entering the plant from an open door, thus possibly contaminating the pellets being fed into the hoppers of the injection-molding machine, (c) variation in the mold temperature depending on the delay between successive molding shots, and (d) variability in the performance of individual operators during different shifts.

In order to maintain good product quality, it is essential to understand the effects, if any, of each element of noise in the operation, such as, for example: (a) Why and how does the ambient temperature affect the quality and surface characteristics of the molded gears? (b) Why and how does the dust coating on a pellet in the machine's hopper affect its behavior in the molding machine? As a result, certain control features may be instituted in the system.

40.9 Process Selection

Process selection is intimately related to the characteristics of the materials to be processed, as shown in Table 40.5.

Characteristics and Properties of Workpiece Materials. Recall that some materials can be processed at room temperature, whereas others require elevated temperatures, thus the need for certain furnaces, appropriate tooling, and various controls. Some materials are easy to process because they are soft and ductile; others are hard, brittle, and abrasive, thus requiring appropriate processing techniques, tooling, and equipment.

Table 40.5: General Characteristics of Manufacturing Processes for Various Metals and Alloys.

	Carbon steels	Alloy steels	Stainless steels	Tool and die steels	Aluminum alloys	Magnesium alloys	Copper alloys	Nickel alloys	Titanium alloys	Refractory alloys
Casting										
Sand	1	1	1	2	1	1	1	1	2	1
Plaster	3	3	3	3	1	1	1	3	3	—
Ceramic	1	1	1	1	2	2	1	1	2	1
Investment	1	1	1	3	1	2	1	1	1	1
Permanent	2	2	3	3	1	1	1	3	3	—
Die	3	3	3	3	1	1	1	3	3	—
Forging										
Hot	1	1	1	1	1	1	1	1	1	1
Cold	1	1	1	3	1	2	1	3	3	—
Extrusion										
Hot	1	1	1	2	1	1	1	1	1	1
Cold	1	2	1	3	1	3	1	2	3	3
Impact	3	3	3	3	1	1	1	3	3	3
Rolling	1	1	1	3	1	1	1	1	1	2
Powder metallurgy	1	1	1	1	1	1	1	1	1	1
Sheet-metal forming	1	1	1	3	1	1	1	1	1	1
Machining	1	1	1	3	1	1	1	2	1	2
Chemical	1	2	1	2	1	1	1	2	2	2
ECM	3	1	2	1	3	3	2	1	1	1
EDM	3	2	2	1	3	3	2	2	2	1
Grinding	1	1	1	1	1	1	1	1	1	1
Welding	1	1	1	3	1	1	1	1	1	1

Note: 1 = Generally processed by this method, 2 = Can be processed by this method, 3 = Usually not processed by this method.

As can be noted in Table 40.5, few materials have favorable manufacturing characteristics in all categories. A material that is castable or forgeable, for example, may later present difficulties in subsequent processing, such as machining, grinding, and polishing that may be required for an acceptable surface finish and dimensional accuracy.

Moreover, materials have different responses to the rate of deformation (see *strain-rate sensitivity*, Sections 2.2.7 and 7.3) to which they are subjected. Consequently, the speed at which a particular machine is operated can affect product quality, including the development of external and internal defects. For example, impact extrusion or drop forging may not be appropriate for a certain material with high strain-rate sensitivity, whereas the same material will perform well in a hydraulic press or in direct extrusion.

Geometric Features. Requirements for part shape, size, and thickness; dimensional tolerances; and surface finish may greatly influence the selection of a process or processes, as described throughout this and various other chapters in the book.

Production Rate and Quantity. These requirements dictate process selection by way of the productivity of a process, machine, or system (see Section 40.8).

Process Selection Considerations. The major factors involved in process selection may be summarized as follows:

1. Are some or all of the parts or components that are needed in a product commercially available as standard items?
2. Which components of the product have to be manufactured in the plant?
3. Is the tooling that is required available in the plant? If not, can it be purchased as a standard item?
4. Can group technology be implemented for parts that have similar geometric and manufacturing attributes?
5. Have all alternative manufacturing processes been investigated?
6. Are the methods selected economical for the type of material, the part shape to be produced, and the required production rate?
7. Can the requirements for dimensional tolerances, surface finish, and product quality be met consistently or can they be relaxed?
8. Can the part be produced to its final dimensions and surface characteristics without requiring additional processing or finishing operations?
9. Are all processing parameters optimized?
10. Is scrap produced, and if so, is it minimized? What is the value of the scrap?
11. Have all the automation and computer-control possibilities been explored for all phases of the total manufacturing cycle?
12. Are all in-line automated inspection techniques and quality control being implemented?

Example 40.3 Process Substitution in Products

The following list gives some typical choices that can be made in process selection for the products listed. Recall that each pair of process has its own advantages and limitations regarding such considerations as weight, life expectancy, durability, and cost.

1. Crankshaft: Forged versus cast
2. Connecting rod: Forged versus powder metallurgy
3. Hubcap: Sheet metal versus cast
4. Automobile wheel: Forged versus cast
5. Large gear: Machining versus precision forming
6. Spur gear: Forging versus powder metallurgy
7. Threaded fastener: Thread rolling versus machining
8. Frying pan: Casting versus stamping
9. Outdoor furniture: Formed aluminum tubing versus cast iron
10. Machine-tool structures: Welding versus mechanical fastening.

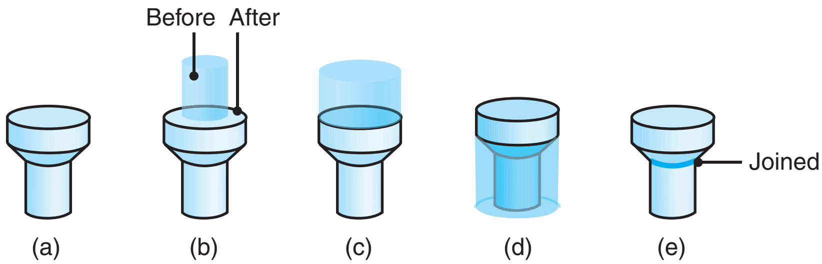


Figure 40.6: Various methods of making a simple part: (a) casting or powder metallurgy, (b) forging or upsetting, (c) extrusion, (d) machining, and (e) joining two pieces.

Example 40.4 Process Selection in Making a Simple Axisymmetric Part

Given: The simple part shown in Fig. 40.6a is to be produced in large quantities. It is 125 mm long and its large and small diameters are, respectively, 38 mm and 25 mm. Assume that this part must be made of metal because of functional requirements such as strength, stiffness, hardness, and resistance to wear at elevated temperatures.

Find: Select a manufacturing process, and describe how you would organize the production facilities to manufacture a cost-competitive, high-quality product.

Solution: Recall that, as much as possible, parts should be produced at or near their final shape (net- or near-net-shape manufacturing), under an approach that largely eliminates much secondary processing and thus reduce the total manufacturing time and cost. Because its shape is relatively simple, this part can be made by (a) casting, (b) powder metallurgy, (c) upsetting or forging, (c) extrusion, (d) machining, or (e) joining two separate pieces together.

For net-shape production, the two suitable processes are casting and powder metallurgy; each of these two processes has its own characteristics, need for specific tooling, labor skill, and costs. The part can also be made by cold, warm, or hot forming. One method is upsetting (see *heading*, Fig. 14.12) in which a 25-mm round blank is placed in a round die cavity to form the larger end. Another possibility is partial direct extrusion of a 38-mm diameter blank to reduce its diameter to 25 mm. Note that each of these processes produces little or no material waste, an important factor in green manufacturing.

This part also can be made by machining a 38-mm-diameter bar stock to reduce the lower section to 25 mm. Machining, however, will require much more time than forming it, and a considerable amount of material inevitably will be wasted as metal chips (see Table 40.4). However, unlike net-shape processes, which generally require special dies, machining involves no special tooling, and the operation can be carried out easily on a CNC lathe and at high rates. Note also that, alternatively, the part can be made in two separate pieces and then joined by welding, brazing, or adhesive bonding.

After these initial considerations, it appears that if only a few parts are needed, machining this part is the most economical method. However, for a high production quantity and rate as stated, producing this part by a heading operation or by cold extrusion (a variation of closed-die forging, see Section 15.4) would be an appropriate choice. Finally, note that if, for some technical reason, the top and bottom portions of the part must be made of different materials, then the part can be made in two pieces, and joining them would be the most appropriate choice.

40.10 Manufacturing Costs and Cost Reduction

The total cost of a product typically consists of material costs, tooling costs, fixed costs, variable costs, direct-labor costs, and indirect-labor costs. As a general guide to the costs involved, see the sections on the economics of each chapter concerning individual groups of manufacturing processes and operations. Depending on the particular company and the type of products made, different methods of cost accounting may be used, with methodologies of accounting procedures that can be complex and even controversial. Moreover, because of the numerous technical and operational factors involved, calculating individual cost factors can be challenging, time consuming, and not always reliable.

Costing systems, also called *cost justification*, typically include the following considerations: (a) intangible benefits of quality improvements and inventory reduction, (b) life-cycle costs, (c) machine usage, (d) cost of purchasing machinery compared with that of leasing it, (e) financial risks involved in implementing highly automated systems, and (f) implementation of new technologies and their impact on products.

Additionally, the costs to a manufacturer that are attributed directly to product liability continue to be a matter of major concern, and every product has a built-in cost to cover possible product liability claims. It has been estimated that liability suits against car manufacturers in the United States add about $500 to the indirect cost of an automobile, and because of the risks involved in its use, 20% of the price for a ladder is attributed to potential product liability costs.

Material Costs. Some cost data on materials are given in various tables throughout this book. Because of the different operations involved in producing raw materials (stock), their costs depend on the (a) type of material, (b) its processing history, and (c) its size, shape, and surface characteristics. Per unit weight, for example,

- Drawn round bars are less expensive than bars that are ground to close tolerances and fine surface finish
- Bars with square cross-sections are more expensive than round bars
- Cold-rolled plate is more expensive than hot-rolled plate of the same thickness
- Thin wire is more expensive than thick wire
- Solid stock is much less expensive than metal powders of the same type.

Tooling Costs. *Tooling costs* can be very high but they can be justified in high-volume production, such as automotive applications, where die costs can be $2 million or more. The expected life of tools and die, and their obsolescence because of product changes, also are major considerations.

Tooling costs are greatly influenced by the operation they perform. For example:

- Tooling cost for die casting is higher than that for sand casting
- Tooling costs for machining or grinding are much lower than those for powder metallurgy, forging, or extrusion
- Carbide tools are more expensive than high-speed steel tools but tool life is longer
- If a part is to be made by spinning, the tooling cost for conventional spinning is much lower than that for shear spinning
- Tooling for rubber-forming processes is less expensive than the die sets (male and female) used for the deep drawing and forming of sheet metals

Fixed Costs. These costs include electric power, fuel, taxes on real estate, rent, insurance, and capital (including depreciation and interest). A company has to meet fixed costs regardless of whether or not it has made a product; thus, fixed costs are not sensitive to production volume.

Table 40.6: Relative Costs for Machinery and Equipment.

Automatic screw machine	M-H
Boring mill, horizontal	M-H
Broaching	M-H
Deep drawing	M-H
Die casting	M-H
Drilling	L-M
Electrical-discharge machining	L-M
Electron-beam welding	M-H
Extruder, polymer	L-M
Extrusion press	M-H
Flexible manufacturing cell and system	H-VH
Forging	M-H
Fused deposition modeling	L
Gas tungsten-arc welding	L
Gear shaping	L-H
Grinding	L-H
Headers	L-M
Honing, lapping	L-M
Injection molding	M-H
Laser-beam welding	M-H
Lathes	L-M
Machining center	L-M
Mechanical press	L-M
Milling	L-M
Powder-injection molding	M-H
Powder metallurgy	L-M
Powder metallurgy, HIP	M-H
Resistance spot welding	L-M
Ring rolling	M-H
Robots	L-M
Roll forming	L-M
Rubber forming	L-M
Sand casting	L-M
Spinning	L-M
Stereolithography	L-M
Stamping	L-M
Stretch forming	M-H
Transfer lines	H-VH
Ultrasonic welding	L-M

Note: L = low; M = medium; H = high; VH = very high. Costs vary greatly, depending on size, capacity, options, and level of automation and computer controls. See also the sections on economics in various chapters.

Capital Costs. These costs represent machinery, tooling, equipment, and investment in buildings and land. As can be seen in Table 40.6, the cost of machines and systems can vary widely, depending on numerous factors. In view of the generally high equipment costs (particularly those involving transfer lines and flexible-manufacturing cells and systems), high production quantities and rates are essential to justify such large expenditures, as well as to keep product costs at or below the all-important competitive level. Lower unit costs (cost per piece) can be achieved by continuous production, involving around-the-clock operation, as long as demand warrants it. Equipment maintenance also is essential to ensure high productivity; as any breakdown of machinery, leading to *downtime*, can be very costly, by as much as thousands of dollars per hour.

Direct-labor Costs. *Direct-labor costs* involve labor that is directly involved in manufacturing products, also known as *productive labor*. These costs include the costs of all labor from the time raw materials are first handled to the time when the product is manufactured, a period generally referred to as *floor-to-floor time*. Direct-labor costs are calculated by multiplying the labor rate (the hourly wage, including benefits) by the amount of time that the worker spends producing the particular part.

The time required for producing a part depends not only on its specified size, shape, dimensional accuracy, and surface finish, but also on the workpiece material itself. For example, the cutting speeds for machining high-temperature alloys are lower than those for machining aluminum or plain-carbon steels. Labor costs in manufacturing and assembly vary greatly from country to country (see Table I.4 in the General Introduction). It is thus understandable that most of the products one purchases today are either made or assembled in countries where labor costs are low. Note, for example, that the global market for bicycles in 2011 was $61 billion, 66% of which were made in China. Firms located in countries with high labor rates tend, on the other hand, to emphasize *high value added* manufacturing tasks or high automation levels and the associated increases in productivity, such that the labor component of the cost is significantly reduced.

For labor-intensive industries, such as clothing and textiles, steelmaking, petrochemicals, and chemical processing, manufacturers generally consider moving production to countries with a lower labor rate, a practice known as **outsourcing**. While this approach can be financially attractive, the cost savings anticipated may not always be realized because of the following hidden costs associated with outsourcing:

- International shipping is far more involved and time consuming than domestic shipping. For example, it takes roughly four to six weeks for a container ship to bring a product from China to the United States or Europe, an interval that has continued to increase because of homeland security issues. Moreover, shipping costs can fluctuate significantly and in an unpredictable manner; it is therefore not possible to reliably predict or budget shipping costs.

- Lengthy shipping schedules indicate that the benefits of just-in-time manufacturing approach and its associated cost savings may not be realized. Moreover, because of the long shipping times, schedules are rigid, design modifications cannot be made easily, and companies cannot readily address changes in the market or in demand. Thus, companies that outsource can lose agility and may also have difficulties in following lean-manufacturing approaches.

- Legal systems are not as well established in countries with lower labor rates as they are in other countries. Procedures that are common in the United States and the European Union, such as accounting audits, protection of patented designs and intellectual property, and conflict resolution, are more difficult to enforce or obtain in other countries.

- Because payments typically are expected on the basis of units completed and sold, the consequences of defective products and their rates can become significant.

- There are various other hidden costs in outsourcing, such as increased paperwork and documentation, lower productivity from existing employees because of lower morale, and difficulties in communication.

For all these reasons, manufacturers that have outsourced manufacturing activities have not been able to realize the hoped for benefits and, consequently, a trend toward **reshoring** manufacturing efforts has recently begun.

Indirect-labor Costs. These costs are generated in the servicing of the total manufacturing operation, consisting of such activities as supervision, maintenance, quality control, repair, engineering, research, and sales, as well as the cost of office staff. Because they do not contribute directly to the production of specific products, these costs are referred to as the *overhead* or *burden rate* and are charged proportionally to all products. The personnel involved in these activities are categorized as **nonproductive labor**.

Manufacturing Costs and Production Quantity. One of the most significant factors in manufacturing costs is the *production quantity* (see Tables 37.2 and 37.3). A large production quantity obviously requires high production rates which, in turn, require the use of **mass production** techniques that involve special machinery (**dedicated machinery**) and employ proportionally less direct labor. At the other extreme, a smaller production quantity usually means a larger direct-labor involvement.

Small-batch production usually involves general-purpose machines, such as lathes, milling machines, and hydraulic presses. The equipment is versatile, and parts with different shapes and sizes can be produced by appropriate changes in the tooling. However, direct-labor costs are high because these machines usually are operated by skilled labor.

In **medium-batch production**, the quantities are larger and general-purpose machines are equipped with various jigs and fixtures, or they can be computer controlled. To further reduce labor costs, machining centers and flexible-manufacturing systems are important alternatives. Generally, for quantities of 100,000 or more, the machines are designed for specific purposes, and they perform a variety of specific operations with very little direct labor involved.

Cost Reduction. Cost reduction requires a study of how the costs described previously are interrelated, using *relative costs* as an important parameter. The unit cost of a product can vary widely. For example, some parts may be made from expensive materials but require very little processing, as in the case of minted gold coins; consequently, the cost of materials relative to that of direct labor is high.

By contrast, some products may require several complex and expensive production steps to process relatively inexpensive materials, such as carbon steels. An electric motor, for example, is made of relatively inexpensive materials, yet several different manufacturing operations are involved in making the housing, rotor, bearings, brushes, and various other components. Unless highly automated, assembly operations for such products can become a significant portion of the overall cost (Section 37.9).

In the 1960s, labor accounted for as much as 40% of the production cost; today, it can be as low as 5%, depending on the type of product and level of automation; see Table I.6. Note also that the contribution of the design phase is only 5%, yet it is the design phase that generally has the largest influence on the quality and success of a product in the marketplace.

There are various opportunities for cost reduction, as has been discussed in various chapters throughout this book. Introducing more automated systems and the adoption of up-to-date technologies are an obvious means of reducing some costs. However, this approach must be undertaken after a thorough **cost–benefit analysis**, which requires reliable input data and consideration of all the technical as well as the human factors involved. Advanced technologies, some of which can be very costly to implement, should be considered only after a complete analysis of the more obvious cost factors, known as **return on investment** (ROI).

40.10.1 Value Analysis

Manufacturing adds value to materials as they become discrete parts, components, and products and are then marketed. Because the value is added in individual stages during the creation of the product, the utilization of **value** analysis, also called *value engineering*, *value control*, and *value management*, is important. *Value analysis* is thus a system that evaluates each step in design, material and process selection, and operations in order to manufacture a product that performs all of its intended functions and does so at the lowest possible cost.

In this analysis, developed at the General Electric Co. during World War II, a monetary value is established for each of two product attributes: (a) **use value**, reflecting the functions of the product, and (b) **esteem** or **prestige value**, reflecting the attractiveness of the product that makes its ownership desirable. The *value of a product* is then defined as

$$\text{Value} = \frac{\text{Product function and performance}}{\text{Product cost}} \tag{40.1}$$

Although there are different versions, value analysis basically consists of the following six phases:

1. *Information phase*: gathering data and determining costs
2. *Analysis phase*: defining functions and identifying problems as well as opportunities
3. *Creativity phase*: seeking ideas in order to respond to problems and opportunities without judging the value of each idea
4. *Evaluation phase*: selecting the ideas to be developed and identifying the costs involved
5. *Implementation phase*: presenting facts, costs, and values to the company management; developing a plan and to motivate positive action, all in order to obtain a commitment of the resources necessary to accomplish the task
6. *Review phase*: reexamining the overall value-analysis process in order to make necessary adjustments.

Value analysis is usually coordinated by a value engineer and conducted jointly by designers; manufacturing engineers; and quality-control, purchasing, and marketing personnel, and it must have the full support of a company's top management. The implementation of value analysis in manufacturing can result in such benefits as (a) significant cost reduction, (b) reduced lead times, (c) better product quality and performance, (d) a reduced time for manufacturing the product, and (e) reduced product weight and size.

Summary

- Regardless of how well a product meets design specifications and quality standards, it also must meet economic criteria in order to be competitive in the domestic and global marketplace.
- Important considerations in product design and manufacturing include manufacturing characteristics of materials, product life expectancy, life-cycle engineering, and an awareness of minimizing any potential harm to the environment and the Earth's ecosystem.
- Substitution of materials, modification, and simplifying product design, and relaxing of dimensional tolerance and surface finish requirements are among important methods of cost reduction.
- The total cost of a product includes several elements, such as the costs of materials, tooling, capital, labor, and overhead. Material costs can be reduced through careful selection without compromising design and service requirements, functions, specifications, or standards for good product quality.
- Labor costs generally are becoming an increasingly smaller percentage of production costs in highly industrialized countries, and to counteract lower wages in some developing countries, labor costs can be further reduced through highly automated and computer-controlled equipment and operations.

Key Terms

Burden rate

Capital costs

Cost–benefit analysis

Cost justification

Cost reduction

Cradle-to-cradle

Cradle-to-grave

Dedicated machines

Direct labor

Downtime

Economic order quantity

Fixed costs

Floor-to-floor time
Indirect labor
Lead time
Life-cycle assessment
Nonproductive labor
Outsourcing
Overhead
Process capabilities
Production quantity
Production rate
Recycling
Relative costs
Reshoring
Return on investment
Scrap
Smart databases
Sustainable manufacturing
Value
Value analysis
World class

Bibliography

Agustiady, T., and Badiru, A.B., **Sustainability Utilizing Lean Six Sigma Techniques**, CRC Press, 2012.

Anderson, D.M., **Design for Manufacturability & Concurrent Engineering**, CIM Press, 2010.

Andrae, A.S.G., **Global Life Cycle Assessments of Material Shifts**, Springer, 2009.

Ashby, M.F., **Materials Selection in Mechanical Design**, 4th ed., Butterworth-Heinemann, 2010.

ASM Handbook, Vol. 20: **Materials Selection and Design**, ASM International, 1997.

Boothroyd, G., Dewhurst, P., and Knight, W., **Product Design for Manufacture and Assembly**, 3rd ed., CRC Press, 2010.

Bralla, J.G., **Design for Manufacturability Handbook**, 2nd ed., McGraw-Hill, 1999.

Cook, H.F., and Wissmann, L.A., **Value Driven Product Planning and Systems Engineering**, Springer, 2010.

Curran, M.A., **Goal and Scope Definition of Life Cycle Assessment**, Springer, 2016.

Design for Manufacturability, Vol. 6, Society of Manufacturing Engineers, 1992.

Fiksel, J., **Design for the Environment**, 2nd ed., McGraw-Hill, 2011.

Giudice, F., La Rosa, G., and Risitano, A., **Product Design for the Environment**, CRC Press, 2006.

Madu, C.N. (ed.), **Handbook of Environmentally Conscious Manufacturing**, Springer, 2011.

Muthu, S.S., **The Carbon Footprint Handbook**, CRC Press, 2015.

Poli, C., **Design for Manufacturing: A Structured Approach**, Butterworth-Heinemann, 2001.

Ribbens, J., **Simultaneous Engineering for New Product Development: Manufacturing Applications**, Wiley, 2000.

Seliger, G., (ed.), **Sustainability in Manufacturing**, Springer, 2010.

Swift, K.G., and Booker, J.D., **Process Selection: From Design to Manufacture**, 2nd ed., Butterworth-Heinemann, 2003.

Ulrich, K., and Eppinger, S., **Product Design and Development**, 5th ed., McGraw-Hill, 2011

Wenzel, H., Hauschild, M., and Alting, L., **Environmental Assessment of Products**, Springer, 2003.

Review Questions

40.1. Explain what is meant by "manufacturing properties" of materials.

40.2. Why is material substitution an important aspect of manufacturing engineering?

40.3. What factors are involved in the selection of manufacturing processes? Explain why they are important.

40.4. How is production quantity significant in process selection? Explain.

40.5. List and describe the major costs involved in manufacturing.

40.6. Describe life-cycle assessment and life-cycle engineering.

40.7. Define what is meant by economic order quantity.

40.8. Explain the difference between direct-labor cost and indirect-labor cost.

40.9. Describe your understanding of the following terms: (a) life expectancy, (b) life-cycle engineering, (c) sustainable manufacturing, and (d) green manufacturing.

40.10. What is the difference between production quantity and production rate?

40.11. Is there a significant difference between cradle-to-grave and cradle-to-cradle production? Explain.

40.12. How would you define value? Explain.

40.13. Define sustainable manufacturing.

40.14. What is the meaning and significance of the term "return on investment"? Explain.

Qualitative Problems

40.15. Describe the major considerations involved in selecting materials for products.

40.16. What is meant by manufacturing process capabilities? Select four different manufacturing processes and describe their capabilities.

40.17. Comment on the magnitude and range of scrap shown in Table 40.4 and the reasons for the variations.

40.18. Explain why the value of the scrap produced in a manufacturing process depends on the type of material and processes involved.

40.19. Describe your observations concerning the information given in Table 6.1 and the reasons for those observations.

Material	Relative cost
Gold	70,000
Silver	680
Molybdenum alloys	200–250
Nickel	40
Titanium alloys	25–40
Copper alloys	8–10
Zinc alloys	1.5–3.5
Stainless steels	2–9
Magnesium alloys	2–4
Aluminum alloys	1.5–3
High-strength low-alloy steels	1.4
Gray cast iron	1.2
Carbon steel	1
Nylons, acetals, and silicon rubber*	1.1–2
Rubber*	0.2–1
Other plastics and elastomers*	0.2–2

* As molding compounds.

Note: Costs vary significantly with quantity of purchase, supply and demand, size and shape, and various other factors.

40.20. Other than the size of the machine, what factors are involved in the range of prices in each machine category shown in Table 40.6? Explain.

40.21. Explain why it takes different amounts of energy to produce different materials. Consider both the material and the processing history.

40.22. Refer to Table 40.2 and explain why it is essential to recycle aluminum and magnesium. From a lifecycle standpoint, explain why aluminum and magnesium should or should not be used in automobiles.

40.23. Other than the size of the machine, what factors are involved in the range of prices in each machine category shown in Table 40.6?

40.24. Explain how the high cost of some of the machinery listed in Table 40.6 can be justified.

40.25. On the basis of the topics covered in this book, explain the reasons for the relative positions of the curves shown in Fig. 40.3.

40.26. What factors are involved in the shape of the curve shown in Fig. 40.2? Explain.

40.27. Is it always desirable to purchase stock that is close to the final dimensions of a part to be manufactured? Explain your answer, and give some examples.

40.28. Describe the problems that may have to be faced in reducing the quantity of materials in products. Give some examples.

40.29. Explain the reasons that there is a strong desire in industry to practice near-net-shape manufacturing.

40.30. State and explain your thoughts concerning cradle-to-cradle manufacturing.

40.31. Why is the amount of scrap produced in a manufacturing process important?

40.32. Discuss the advantages to long lead times, if any, in production.

40.33. Review Table 40.2 and estimate the carbon footprint of materials (mass of carbon produced per mass or volume of material) if the energy used to produce the material is obtained from (a) hydroelectric power, wind, or nuclear energy; (b) coal.

40.34. Explain why the larger the quantity per package of food products, the lower is the cost per unit weight.

40.35. List and explain the advantages and disadvantages of outsourcing manufacturing activities to countries with low labor costs.

Synthesis, Design, and Projects

40.36. As you can see, Table 40.5 lists only metals and their alloys. On the basis of the information given in various chapters in this book and in other sources, prepare a similar table for nonmetallic materials, including ceramics, plastics, reinforced plastics, and both metal-matrix and ceramic-matrix composite materials.

40.37. Is it always desirable to purchase stock that is close to the final dimensions of a part to be manufactured? Explain why or why not and give some examples.

40.38. What course of action would you take if the supply of a raw material selected for a product line becomes unreliable? Explain.

40.39. Estimate the position of the curves for the following processes in Fig. 40.3: (a) centerless grinding, (b) electrochemical machining, (c) chemical milling, and (d) extrusion.

40.40. Review Fig. I.3 in the General Introduction and present your own thoughts concerning the two flowcharts. Would you want to make any modifications, and if so, what would they be?

40.41. Over the years, numerous consumer products (such as rotary-dial telephones, analog radio tuners, turntables, and vacuum tubes) have become obsolete or nearly so, while many new products have entered the market. Make two lists: a comprehensive list of obsolete products that you can think of and a list of new products. Comment on the reasons for the changes you observe.

40.42. List and discuss the different manufacturing methods and systems that have enabled the manufacture of new products. (These products and systems are known as *enabling technologies*.)

40.43. Select three different products, and make a survey of the changes in their prices over the past 10 years. Discuss the possible reasons for the changes.

40.44. Describe your own thoughts concerning the replacement of aluminum beverage cans with cans made of steel.

40.45. Select three different products commonly found in homes. State your opinions on (a) what materials were used in each product, (b) why those particular materials were chosen, (c) how the products were manufactured, and (d) why those particular processes were used.

40.46. Comment on the differences, if any, among the designs, materials, and processing and assembly methods used for making products such as hand tools and ladders for professional use and those for consumer use.

40.47. The cross-section of a jet engine is shown in Fig. 6.1. On the basis of the topics covered in this book, select any three individual components of such an engine and describe the materials and processes that you would use in making them in quantities of, say, 1000.

40.48. Inspect some products around your home, and describe how you would go about taking them completely apart quickly and recycling their components. Comment on their design regarding the ease with which they can be disassembled.

40.49. What products do you know of that would be very difficult to disassemble for recycling purposes? Explain.

40.50. Conduct a literature search and perform a lifecycle assessment of a typical automobile. Estimate the amount of energy needed to produce the car from its raw materials, and compare this to the energy that is consumed by the car during its intended life of 160,000 km. What recommendations would you make regarding the use of aluminum and magnesium instead of steel in cars?

40.51. Discuss the trade-offs involved in selecting between the two materials for each of the applications listed:

1. Sheet metal versus reinforced plastic chairs;
2. Forged versus cast crankshafts;
3. Forged versus powder-metallurgy connecting rods;
4. Plastic versus sheet-metal light-switch plates;
5. Glass versus metal water pitchers;
6. Sheet-metal versus cast hubcaps;
7. Steel versus copper nails;
8. Wood versus metal handles for hammers.

Also, discuss the typical conditions to which these products are subjected in their normal use.

40.52. Discuss the factors that influence the choice between the following pairs of processes to make the products indicated:

1. Sand casting versus die casting of a fractional electric-motor housing;

2. Machining versus forming of a large-diameter bevel gear;
3. Forging versus powder-metallurgy production of a cam;
4. Casting versus stamping a sheet-metal frying pan;
5. Making outdoor summer furniture from aluminum tubing versus cast iron;
6. Welding versus casting of machine-tool structures;
7. Thread rolling versus machining of a bolt for high-strength application;
8. Thermoforming a plastic versus molding a thermoset to make the blade for an inexpensive household fan.

Index

E

G

H

I

J

U