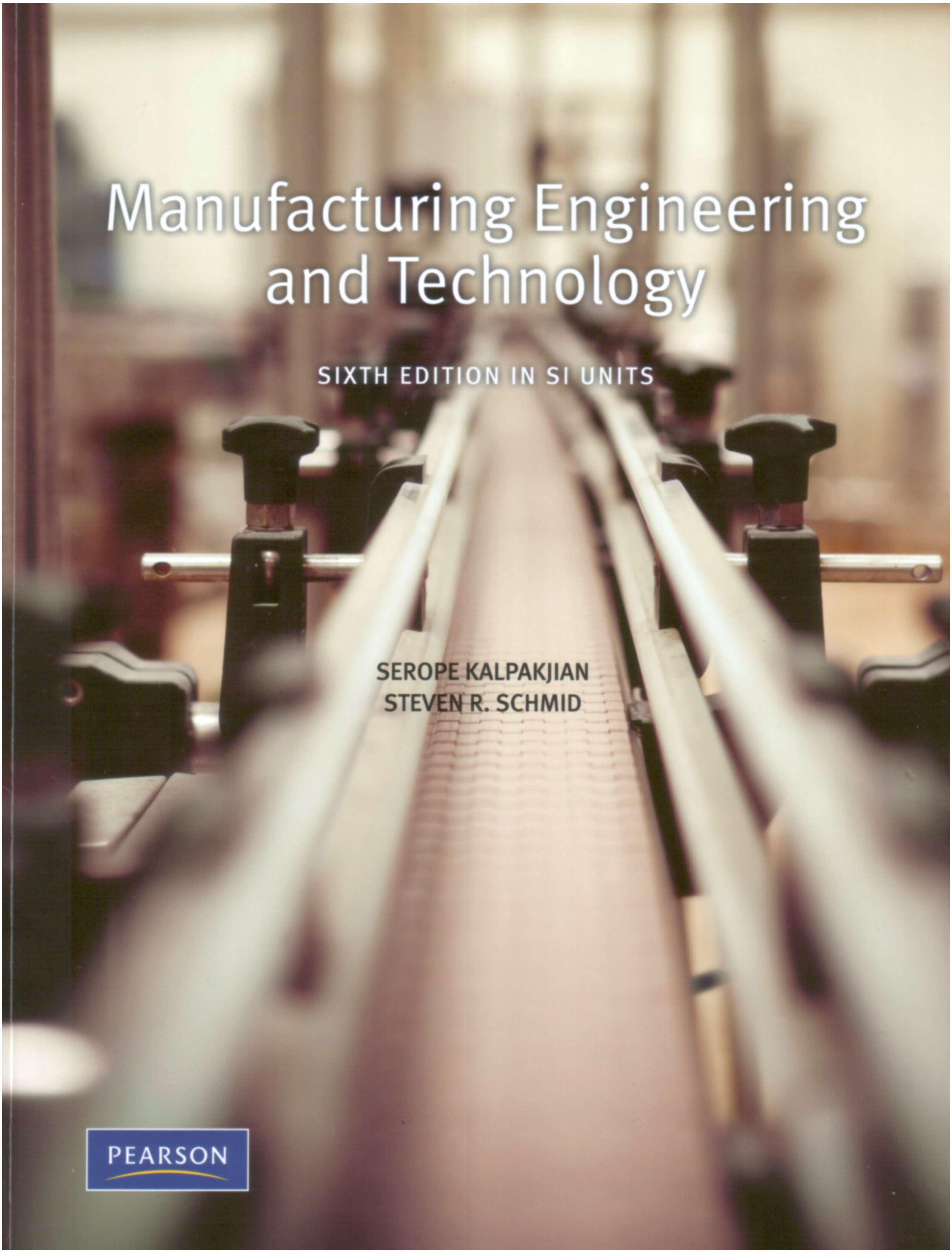


Manufacturing Engineering and Technology

SIXTH EDITION IN SI UNITS

SEROPE KALPAKJIAN
STEVEN R. SCHMID

PEARSON

MANUFACTURING ENGINEERING AND TECHNOLOGY

SIXTH EDITION IN SI UNITS

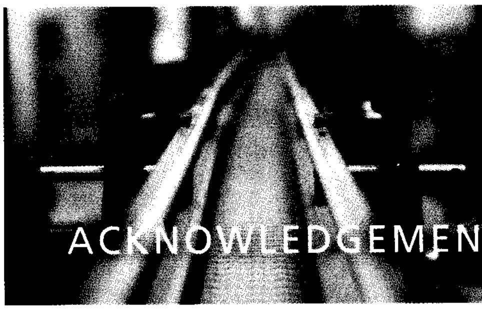
Serope Kalpakjian
Illinois Institute of Technology

Steven R. Schmid
The University of Notre Dame

SI Conversion by
Hamidon Musa
Universiti Teknologi Malaysia

Prentice Hall

Singapore London New York Toronto Sydney Tokyo Madrid
Mexico City Munich Paris Capetown Hong Kong Montreal



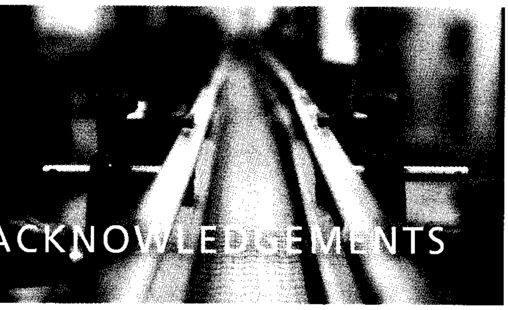
ACKNOWLEDGEMENTS

It gives us great pleasure to acknowledge the assistance of the following colleagues in the preparation and publication of this edition of the book: Kornel F. Ehmann, Northwestern University; Klaus J.Weinmann, University of California at Berkeley; Timotius Pasang, AUT University; Miguel Sellés Cantó, Escola Politècnica Superior d'Alcoi, Universitat Politècnica de València, Spain; and Kevin Wilson, Michael Giordano and Megan McGann, University of Notre Dame.We also acknowledge Kent M. Kalpakjian as the original author of the chapter on Fabrication of Microelectronic Devices.

We would like to thank our editor, Holly Stark, at Prentice Hall for her enthusiastic support and guidance; and Clare Romeo, Production Editor, for her meticulous editorial and production supervision. We would also like to acknowledge the help of the following in the production of this edition of the book: Greg Dulles, A/V Project Manager; Daniel Sandin, Media Editor; William Opaluch, Editorial Assistant; and Kenny Beck, Cover Design.

We are happy to present below a cumulative list of all those individuals who generously have made numerous contributions to various editions of our two books.

B.J. Aaronson
R. Abella
D. Adams
K. Anderson
S. Arellano
D.D. Arola
R.A. Arlt
V. Aronov
A. Bagchi
E.D. Baker
J. Barak
J. Ben-Ari
G.F. Benedict
S. Bhattacharyya
JT. Black
W. Blanchard
C. Blathras
G. Boothroyd
D. Bourell
B. Bozak
N.N. Breyer
C.A. Brown
R.G. Bruce
J. Cesarone
T.-C. Chang
R.L. Cheaney
A. Cheda
S. Chelikani
S. Chen
S.-W. Choi
A. Cinar
R.O. Colantonio
P. Cotnoir
P.J. Courtney
P. Demers
D. Descoteaux
M.F. DeVries
R.C. Dix
M. Dollar
D.A. Dornfeld
H.I. Douglas
M. Dugger
D.R. Durham
D. Duvall
S.A. Dynan
J. El Gomayel
M.G. Elliott
N.X. Fang
E.C. Feldy
J. Field
G.W. Fischer
D.A. Fowley
R.L. French
B.R. Fruchter
D. Furrer
R. Giese
E. Goode
K.L. Graham
P. Grigg
M. Grujicic
P.J. Guichelaar
B. Harriger
D. Harry
M. Hawkins
R.J. Hocken
E.M. Honig, Jr.
Y. Huang
S. Imam
R. Jaeger
C. Johnson
K. Jones
D. Kalisz
K.M. Kalpakjian
J. Kamman
S.G. Kapoor
R. Kassing
R.L. Kegg
W.J. Kennedy
R. Kerr
T. Kesavadas
B.D. King

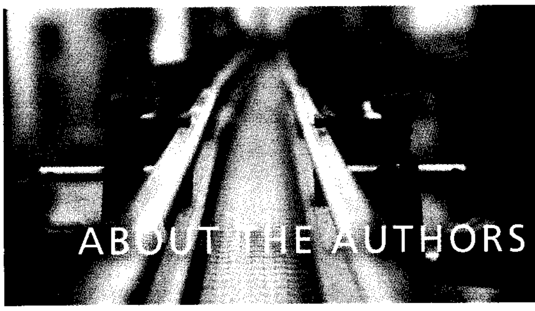


J.E. Kopf
R.J. Koronkowski
J. Kotowski
S. Krishnamachari
K.M. Kulkarni
T. Lach
L. Langseth
M. Laurent
M. Levine
B.S. Levy
J. Lewandowski
X.Z. Li
Z. Liang
B.W. Lilly
A. Lozier
D.A. Lucca
M. Madou
S. Mantell
L. Mapa
A. Marsan
R.J. Mattice
C. Maziar
T. McClelland
W. McClurg
L. McGuire
K.E. McKee
K.P. Meade
M.H. Miller
R. Miller
T.S. Milo
D.J. Morrison
S. Mostovoy
C. Nair
P.G. Nash
J. Nazemetz
E.M. Odom
N. Pacelli
U. Pal
S. Paolucci
S.J. Parelukar
T. Pasang
Z.J. Pei
J. Penaluna
C. Petronis
M. Philpott
M. Pradheeradhi
J.M. Prince
D.W. Radford
W.J. Riffe
R.J. Rogalla
Y. Rong
A.A. Runyan
P. Saha
G.S. Saletta
M. Salimian
M. Savic
W.J. Schoech
S.A. Schwartz
M. Selles
S. Shepel
R. Shivpuri
M.T. Siniawski
J.E. Smallwood
J.P. Sobczak
L. Soisson
P. Stewart
J. Stocker
L. Strom
A.B. Strong
K. Subramanian
T. Sweeney
W.G. Switalski
T. Taglialavore
M. Tarabishy
K.S. Taraman
R. Taylor
B.S. Thakkar
A. Trager
A. Tseng
C. Tszang
M. Tuttle
S. Vaze
J. Vigneau
G.A. Volk
G. Wallace
J.F. Wang
K.J. Weinmann
R. Wertheim
K. West
J. Widmoyer
K.R. Williams
G. Williamson
B. Wiltjer
J. Wingfield
P.K. Wright
N. Zabaras

We are also grateful to numerous organizations, companies, and individuals who supplied us with many of the illustrations and case studies. These contributions have been specifically acknowledged throughout the text.

It is with joy that we gratefully dedicate this book to Professor John A. Schey, our esteemed colleague and distinguished researcher and author, who has been a wise and great teacher, not only to generations of students but to his many colleagues as well. He and his wife, Gitta, have been treasured friends for many years.

SEROPE KALPAKJIAN
STEVEN R. SCHMID
JANUARY 2009



ABOUT THE AUTHORS

About the Author

Serope Kalpakjian is a professor emeritus of mechanical and materials engineering at the Illinois Institute of Technology, Chicago. He is the author of Mechanical Processing of Materials (Van Nostrand, 1967) and co-author of Lubricants and Lubrication in Metalworking Operations (with E.S. Nachtman, Dekker, 1985). Both of the first editions of his books Manufacturing Processes for Engineering Materials (Addison-Wesley, 1984) and Manufacturing Engineering and Technology (Addison Wesley, 1989) have received the M. Eugene Merchant Manufacturing Textbook Award of SME. He has conducted research in several areas in manufacturing processes, is the author of numerous technical papers and articles in professional journals, handbooks, and encyclopedias, and has edited several conference proceedings. He also has been editor and co-editor of various technical journals and has served on the editorial board of Encyclopedia Americana.

Professor Kalpakjian has received the Forging Industry Educational and Research Foundation Best Paper Award (1966); the Excellence in Teaching Award from the Illinois Institute of Technology (1970); the ASME Centennial Medallion (1980); the SME International Education Award (1989); a Person of the Millennium Award from IIT (1999); the Albert Easton White Outstanding Teacher Award from ASM International (2000); and the SME Outstanding Young Manufacturing Engineer Award for 2002 was named after him. Professor Kalpakjian is a Life Fellow of ASME; Fellow of SME; Fellow and Life Member of ASM International; Fellow Emeritus of The International Academy for Production Engineering (CIRP); and is a past president and founding member of the North American Manufacturing Research Institution/ SME. He is a high-honors graduate of Robert College (Istanbul), Harvard University, and the Massachusetts Institute of Technology.

ABOUT THE AUTHORS

About the Author

Steven R. Schmid is an Associate Professor in the Department of Aerospace and Mechanical Engineering at the University of Notre Dame, where he teaches and conducts research in the general areas of manufacturing, machine design, and tribology. He received his Bachelor's degree in Mechanical Engineering from the Illinois Institute of Technology (with Honors) and Master's and Ph.D. degrees, both in Mechanical Engineering, from Northwestern University. He has received numerous awards, including the John T. Parsons Award from SME (2000); the Newkirk Award from ASME (2000); the Kaneb Center Teaching Award (2000 and 2003); and the Ruth and Joel Spira Award for Excellence in Teaching (2005). He is also the recipient of a National Science Foundation CAREERS Award (1996) and an ALCOA Foundation Award (1994).

Professor Schmid is the author of almost 100 technical papers, has co-authored the texts Fundamentals of Machine Elements (McGraw-Hill), Fundamentals of Fluid Film Lubrication (Dekker), Manufacturing Processes for Engineering Materials (Prentice Hall), and has contributed two chapters to the CRC Handbook of Modern Tribology. He is an Associate Editor of the ASME Journal of Manufacturing Science and Engineering, and is a registered Professional Engineer and a Certified Manufacturing Engineer of SME.

About the Adaptor

Hamidon Musa is Associate Professor with the Manufacturing Division in the Faculty of Mechanical Engineering at Universiti Teknologi Malaysia. He graduated from the University of Birmingham in 1980 and did his post-graduate studies at Cranfield Institute of Technology, UK. His main area of interest is non-traditional machining.

Contents

• •

3 Physical Properties of Materials 88

4 Metal Alloys: Their Structure and Strengthening by Heat Treatment 100

5 Ferrous Metals and Alloys: Production, General Properties, and Applications 129

6 Nonferrous Metals and Alloys: Production, General Properties, and Applications 151

7 Polymers: Structure, General Properties, and Applications 171

26 Abrasive Machining and Finishing Operations 719

27 Advanced Machining Processes 759

Part V: Micromanufacturing and Fabrication of Microelectronic Devices 787

28 Fabrication of Microelectronic Devices 790

Case Studies

General Introduction

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 items that 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, the vast majority of objects around us consist of numerous individual pieces that are built and assembled by a combination of processes called **manufacturing.**

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 concerned with making products. A manufactured product may itself be used to make other products, such as (a) a large press, to shape flat sheet metal into automobile bodies, (b) a drill, for producing holes, (c) industrial sawing machines, for making clothing at high rates, and (d) numerous pieces of machinery, to produce an endless variety of individual items, ranging from thin wire for guitars and electric motors to crankshafts and connecting rods for automotive engines (Fig. I.1).

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

Because a manufactured item typically starts with raw materials, which are then subjected to a sequence of processes to make individual products, it has a certain *value*. For example, clay has some value as mined, but when it is made into a product such as cookware, pottery, an electrical insulator, or a cutting tool, value is *added* to the clay. Similarly, a nail has a value over and above the cost of the short piece of wire or rod from which it is made. Products such as computer chips, electric motors, and professional athletic shoes are known as *high-value-added* products.

A Brief History of Manufacturing. Manufacturing dates back to the period 5000–4000 B.C. (Table I.2), and thus, it is older than recorded history, the earliest forms of which were invented by the Sumerians around 3500 B.C. Primitive cave

EXAMPLES:

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

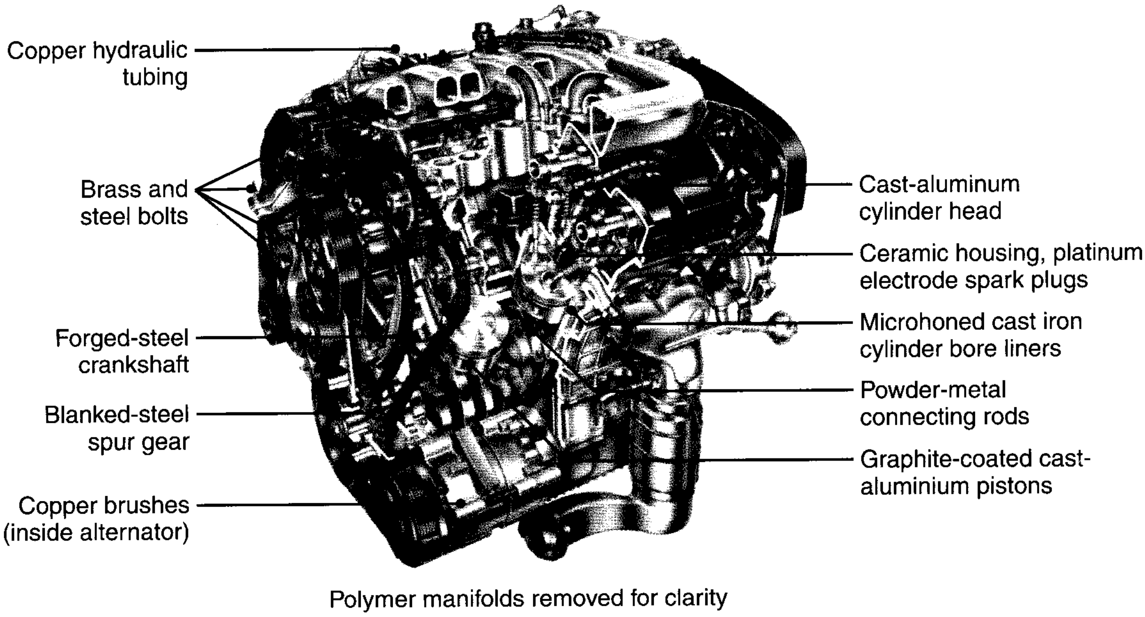


FIGURE I.1 Illustration of an automotive engine (the Duratec V-6), showing various components and the materials used in making them. *Source*: Courtesy of Ford Motor Company. Illustration by D. Kimball.

drawings, as well as markings on clay tablets and stone, needed (1) some form of a brush and some sort of "paint," as in the prehistoric *cave paintings* in Lascaux, France, estimated to be 16,000 years old; (2) some means of scratching the clay tablets and baking them, as in *cuneiform scripts* and *pictograms* of 3000 B.C.; and (3) simple tools for making incisions and carvings on the surfaces of stone, as in the *hieroglyphs* in ancient Egypt.

The manufacture of items for specific uses began with the production of various household artifacts, which were typically made of either 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 (an alloy of copper and tin), and brass (an alloy of copper and zinc). 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 higher levels of product quality. Note, for example, from Table I.2 that lathes for cutting screw threads already were available during the period from 1600 to 1700, but it was not until some three centuries later that automatic screw machines were developed.

TABLE 1.2

Historical Development of Materials, Tools, 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
	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)
Egypt: ~3100 B.C. to ~300 B.C.	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
Greece: ~1100 B.C. to 146 B.C.	2000–1000 B.C.	Wrought iron, brass				
Roman Empire: ~500 B.C. to 476 A.D.	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
Middle Ages: ~476 to 1492	1–1000 A.D.	Zinc, steel	Venetian glass	Armor, coining, forging, steel swords		Etching of armor
	1000–1500	Blast furnace, type metals, casting of bells, pewter	Crystal glass	Wire drawing, gold- and silversmith work		Sandpaper, windmill driven saw
Renaissance: 14th to 16th centuries	1500–1600	Cast-iron cannon, tinplate	Cast plate glass, flint glass	Waterpower 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

(continued)

TABLE I.2 (*continued*)

Historical Development of Materials, Tools, and Manufacturing Processes

Period	Dates	Metals and casting	Various materials and composites	Forming and shaping	Joining	Tools, machining, and manufacturing systems
First Industrial Revolution: ~1780 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 steels, 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
Second Industrial Revolution: 1947–	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, thermomechanical processing	Gas metal arc, gas tungsten arc, and electroslag welding; explosion welding	Electrical and chemical machining, automatic control

TABLE I.2

Historical Development of Materials, Tools, and Manufacturing Processes

Period	Dates	Metals and casting	Various materials and composites	Forming and shaping	Joining	Tools, machining, and manufacturing systems
Space age	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
	1970–1990	Compacted graphite, vacuum casting, organically bonded sand, automation of molding and pouring, rapid solidification, metal-matrix composites, semisolid metalworking, amorphous metals, shape-memory alloys (smart materials), computer simulation	Adhesives, composite materials, semiconductors, optical fibers, structural ceramics, ceramic-matrix composites, biodegradable plastics, electrically conducting polymers	Precision forging, isothermal forging, superplastic forming, dies made by computer-aided design and manufacturing, net-shape forging and forming, computer simulation	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, expert systems, artificial intelligence, computer simulation and optimization
Information age	1990–2000s	Rheocasting, computer-aided design of molds and dies, rapid tooling, TRIP and TWIP steels	Nanophase materials, metal foams, advanced coatings, high-temperature superconductors. machinable ceramics, diamond-like carbon, carbon nanotubes	Rapid prototyping, rapid tooling, environmentally friendly metalworking fluids	Friction stir welding, lead-free solders, laser butt-welded (tailored) sheet-metal blanks, electrically conducting adhesives	Micro- and nano fabrication, LIGA (a German acronym for a process involving lithography, electroplating, and molding), dry etching, linear motor drives, artificial neural networks, six sigma, three-dimensional computer chips

Source: J.A. Schey, C.S. Smith, R.F. Tylecote, T.K. Derry, T.I. Williams, S.R. Schmid, and S. Kalpakjian.

Although ironmaking began in the Middle East in about 1100 B.C., a major milestone was the production of steel in Asia during the period 600–800 A.D. A wide variety of materials continually began to be developed. Today, countless metallic and nonmetallic materials with unique properties are available, including *engineered materials* and various advanced materials. Among the available materials are industrial or high-tech ceramics, reinforced plastics, composite materials, and nanomaterials 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 and is also called the *First Industrial Revolution*, goods had been produced in *batches* and required much reliance on manual labor in all phases of their 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, basically with the development of textile machinery and machine tools for cutting metal. This technology soon moved to the United States, where it continued to be further developed.

A major advance in manufacturing occurred in the early 1800s with the design, production, and use of **interchangeable parts**, conceived by the American manufacturer and inventor Eli 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. Further developments soon followed, resulting in countless consumer and industrial products that 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 detailed review of Table I.2. Note particularly the progress that has been made during the 20th century, compared with that achieved during the 40-century period from 4000 B.C. to 1 B.C.

For example, in the Roman Empire (~500 B.C. to 476 A.D.), factories were available for the mass production of glassware; however, the methods used were generally very slow, and much manpower was involved 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 are punched at rates of 800 holes per minute, and (c) incandescent light bulbs are made at rates of more than 2000 bulbs per minute (see Example I.1), each costing less than one dollar.

EXAMPLE I.1 Incandescent Light Bulbs

The first incandescent lamp was made by T.A. Edison (1847–1931) in New Jersey and was first lit in 1879. A typical bulb then had a life of only about 13.5 hours. Numerous improvements have since been made in both materials and production methods for making light bulbs, with the main purposes being increasing their life and reducing production costs. This example briefly describes the typical sequence of methods used in manufacturing incandescent light bulbs.

The basic components of an incandescent (meaning "glowing with heat") light bulb are shown in Fig. I.2a. The light-emitting component is the filament, which, by the passage of current and due to its electrical resistance, is heated to incandescence to a temperature of 2200°–3000°C. Edison's first successful lamp had a carbon filament, although he and others also had experimented with carbonized paper and metals such as osmium, iridium, and

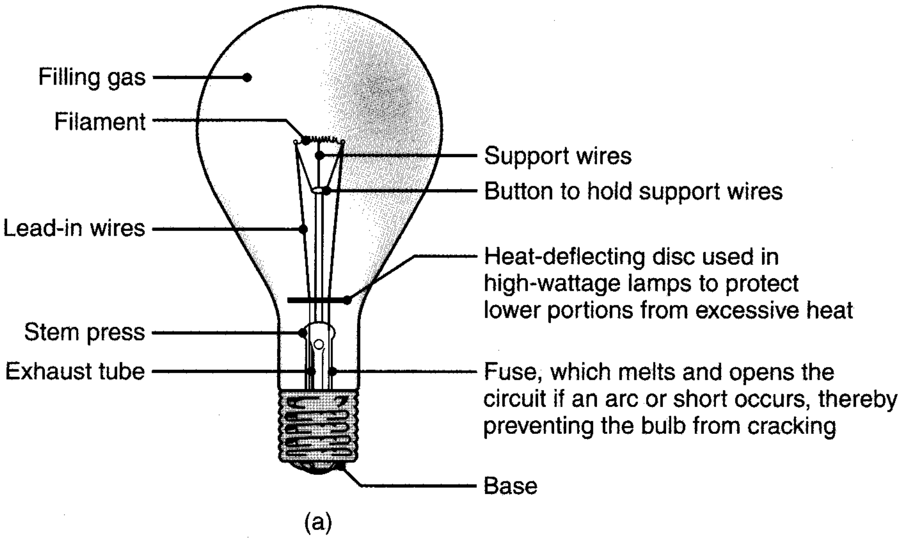


FIGURE I.2a Components of a common incandescent light bulb. *Source*: Courtesy of General Electric Company.

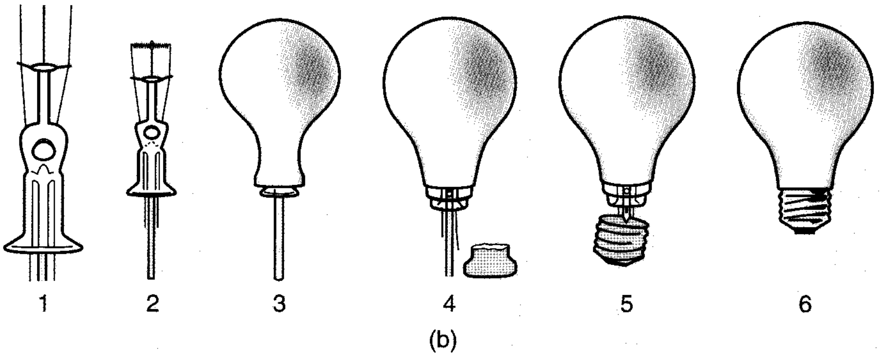


FIGURE I.2b Manufacturing steps in making an incandescent light bulb. *Source*: Courtesy of General Electric Company.

tantalum. However, none of these materials has the strength, resistance to high temperature, and long life as has tungsten (Section 6.8.3), which is now the most commonly used filament material.

The first step in manufacturing a light bulb is making the glass stem that supports the lead-in wires and the filament and connects them to the base of the bulb (Fig. I.2b). All these components are positioned, assembled, and sealed while the glass is heated by gas flames. The filament is then attached to the lead-in wires. The filament is made by powder metallurgy techniques (Chapter 17), which involves first pressing tungsten powder into ingots and sintering it (heating it without its melting). Next, the ingot is shaped into round rods by rotary swaging (Section 14.4) and then drawing it through a set of dies into thin wire (Sections 15.8 and 15.10). The wire diameter for a 60-W, 120-V bulb is 0.045 mm (0.0018 in.). The diameter must be controlled precisely, because if it is only 1% less than the diameter specified, the life of the bulb would be reduced by as much as 25% (because of the increased heat due to the higher electrical resistance of the wire).

Note from Fig. I.2a, as well as by direct obsevation of a clear light bulb, that the filament wire is coiled; this is done in order to increase the light-producing capacity of the filament. The spacing between the coils must be precise, so as to prevent a localized buildup of heat that might short the filament during its use.

The completed stem assembly (called the *mount*) is transferred to a machine that lowers a glass bulb over the mount. Gas flames are used to seal the rim of the mount to the neck of the bulb. The air in the bulb is then exhausted through the exhaust tube (which is an integral part of the glass stem), and the bulb is either evacuated or filled with inert gas. For 40-W bulbs and higher, the gas used is typically a mixture of nitrogen and argon. The exhaust tube is then sealed. The filling gas must be pure, as otherwise the inside surfaces of the bulb will blacken. It has been observed that just one drop of water in the gas that is used for half a million bulbs will cause blackening in all of them.

The next step involves attaching the metal base to the glass bulb with a special cement. The machine that performs this operation also solders or welds the lead-in wires to the base, to provide the electrical connection. The lead-in wires are usually made of nickel, copper, or molybdenum, and the support wires are made of molybdenum (Section 6.8). The portion of the lead-in wire that is embedded in the stem is made from an iron–nickel alloy, which has essentially the same coefficient of thermal expansion as that of the glass (Table 3.1), as otherwise the thermal stresses that develop may cause cracking of the glass stem. The bulb base is generally made from aluminum, replacing the more expensive brass base that was used many years ago. To reduce friction and thus allow easy insertion of the bulb into a socket, the metal base is coated with a special compound.

Several types of glasses (Section 8.4) are used, depending on the bulb type. The bulbs are made by blowing molten glass into a mold (Section 18.3.3). The inside of the bulb either is left clear or is frosted (thus making it translucent), to better diffuse the light and to reduce glare.

I.2 Product Design and Concurrent Engineering

Product design involves the creative and systematic prescription of the shape and characteristics of an artifact to achieve specified objectives while simultaneously satisfying several constraints. Design is a critical activity, because it has been estimated that as much as 80% of the cost of product development and manufacture 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 between 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, which may be new or a modified version of an existing product. The market for the product and its anticipated use(s) also must be clearly defined; this aspect involves the assistance of market analysts and sales personnel who will bring valuable and timely input to the manufacturer, especially regarding market trends.

The Design Process. Traditionally, design and manufacturing activities have taken place *sequentially*, as shown in Fig. I.3a. 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 (a) to use a different material, such as a polymer or a ceramic, instead of a metal or (b) to use the same material, but in a different condition, such as a softer instead of a harder material or a

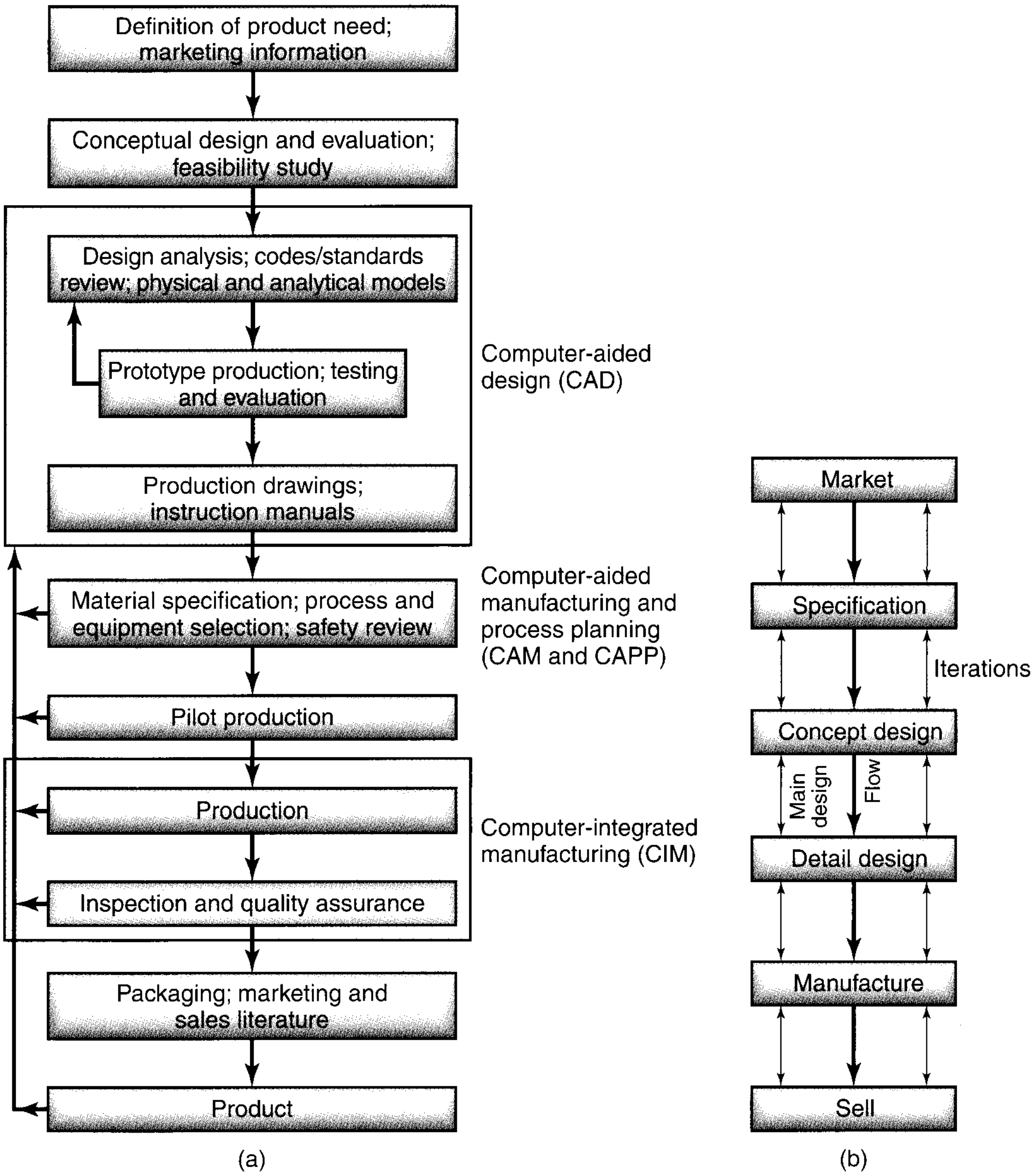


FIGURE I.3 (a) Chart showing various steps involved in *traditional* design and manufacture of a product. Depending on the complexity of the product and the type of materials used, the time span between the original concept and the marketing of the product may range from a few months to several years. (b) Chart showing general product flow in *concurrent engineering*, from market analysis to marketing the product. *Source*: After S. Pugh.

material with a smoother surface finish, or (c) to modify the design of a component in order to make it easier, faster, and less expensive to manufacture. Note that these decisions must take place at the material-specification stage (the sixth box from the top in Fig. I.3a).

Each of the modifications just described will necessitate a repeat of the design analysis stage (the third box from the top in Fig. I.3a) and the subsequent stages, to ensure that the product will still meet all specified requirements and will function satisfactorily. A later change from, say, a forged to a cast 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 is taking place 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.3b. Note that, although this concept, also called **simultaneous engineering**, still has the same general product-flow sequence as in the traditional approach (Fig. I.3a), it now contains several deliberate modifications. From the earliest stages of product design and engineering, *all* relevant disciplines are now *simultaneously* involved. As a result, any iterations that may have to be made will require a smaller effort and thus result 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. Such companies are generally referred to as *small businesses.* 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 the following 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, product 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 have become highly simplified through the use of **computer-aided design** (CAD) and **computer-aided engineering** (CAE) techniques.

CAD systems are capable of rapid and complete analyses of designs, whether it be a simple shelf bracket or a shaft in large and complex structures. The Boeing 777 passenger airplane, for example, was designed completely by computers in a process known as **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 constructed 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 developed is stored and can be retrieved, displayed, printed, and transferred anytime and anywhere within a company's organization. Design modifications can be made and optimized (as is often the practice in engineering, especially in the production of large structures) directly, easily, and at any time.

Computer-aided manufacturing involves all phases of manufacturing, by utilizing and processing the large amount of information on materials and processes gathered and stored in the organization's database. Computers greatly assist in organizing the information developed and performing 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 (Chapter 36).

On the basis of the models developed and analyzed in detail, product designers then finalize the geometric features of each of the product's components, including specifying their dimensional tolerances and surface-finish 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). Indeed, dimensional tolerances are equally important for small products as well as for car bodies or airplanes. 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 **rapid prototyping** (RP, see Chapter 20). Using CAD/CAM and various specialized technologies, designers can now make prototypes rapidly and at low cost, from metallic or nonmetallic materials such as plastics and ceramics.

Prototyping new components by means of traditional methods (such as casting, forming, and machining) could cost an automotive company hundreds of millions of dollars a year, with some components requiring a year or more to complete. Rapid prototyping can significantly reduce costs and the associated product-development times. Rapid-prototyping techniques are now advanced to such a level that they also can be used for low-volume (in batches typically of fewer than 100 parts) economical production of a variety of actual and functional parts to be assembled into products.

Virtual Prototyping. *Virtual prototyping* 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 drawn and developed. Virtual prototyping has been gaining importance, especially because of the availability of low-cost computers and simulation and analysis tools.

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 (1) the characteristics, capabilities, and limitations of materials, manufacturing processes, machinery, equipment, and tooling and (2) variability in machine performance, dimensional accuracy and surface finish of the workpiece, processing time, and the effect of processing methods 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 minimum product cost.

The concepts of **design for assembly** (DFA), **design for manufacture and assembly** (DFMA), and **design for disassembly** (DFD) are all important aspects of all manufacturing. Methodologies and computer software are now available for design for assembly, utilizing 3-D conceptual designs and solid models. Subassembly, assembly, and disassembly times and costs can now be minimized, while product integrity and performance are maintained. Experience has indicated that a product which is easy to assemble is usually also easy to disassemble.

Assembly is an important phase of manufacturing and requires a consideration of the ease, speed, and cost of putting together the numerous individual components of a product (Fig. I.4). Assembly costs in manufacturing operations can be substantial, typically ranging from 20 to 60% of the total product cost. *Disassembly* of a product is an equally important consideration, for maintenance, servicing and recycling of individual components.

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 operations has its own specific characteristics, assembly times, advantages and limitations, associated costs, and special design considerations. Individual parts may be assembled by hand or by a variety of automatic equipment and industrial robots. The choice depends on factors such as product complexity, the number of components to be assembled, the care and protection required to prevent damage to the surfaces of the parts, and the 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 important in product design. Products often have to be disassembled to varying degrees in order to service 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.

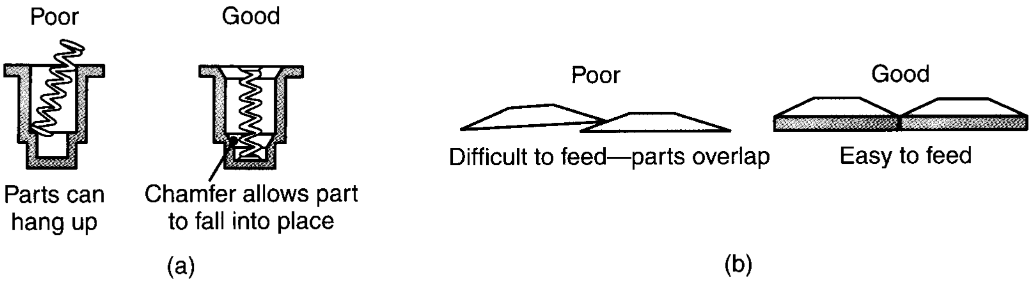


FIGURE I.4 Redesign of parts to facilitate assembly. *Source*: After G. Boothroyd and P. Dewhurst.

I.4 Green Design and Manufacturing

In the United States alone, 9 million passenger cars, 300 million tires, 670 million compact fluorescent lamps, and more than 5 billion kilograms of plastic products are discarded each year. Every three months, industries and consumers discard enough aluminum to rebuild the U.S. commercial air fleet. Note that, as indicated subsequently, the term *discarding* implies that the products have reached the end of their useful life; it does not necessarily mean that they are wasted and 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:

a. Chips from machining and trimmed materials from sheet forming, casting, and molding operations.
b. Slag from foundries and welding operations.
c. Additives in sand used in sand-casting operations.
d. Hazardous waste and toxic materials used in various products.
e. Lubricants and coolants in metalworking and machining operations.
f. Liquids from processes such as heat treating and plating.
g. Solvents from cleaning operations.
h. Smoke and pollutants from furnaces and gases from burning fossil fuels.

The adverse effects of these activities, their damage to our environment and to the Earth's ecosystem, and, ultimately, their effect on the quality of human life are now widely recognized and appreciated. 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 **green design and manufacturing** is now in common usage in all industrial activities, with a major emphasis on **design for the environment** (DFE). Also called **environmentally conscious design and manufacturing**, 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 design and production.

These goals, which increasingly have become global, 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, they lead to new soil that can sustain life. Thus, product design involves the use of (usually) organic materials. The products function well for their intended life and can then be safely discarded.
- *Industrial cycle*: The materials in the product are recycled and reused continuously. For example, aluminum beverage cans are recycled and reused after they have served their intended purpose. To demonstrate the economic benefits of this approach, it has been determined 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%.

One of the basic principles 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 100 million of the 300 million discarded automobile tires are reused in various ways.

Cradle-to-cradle Production. A term coined in the 1970s and also called C2C, *cradle-to-cradle* production 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. Certification procedures for companies are now being developed for cradle-to-cradle production, as they have been for quality control (Section 36.6). *Cradle-to-grave* production, also called *womb-to-tomb* production, has a similar approach, but does not necessarily consider or take on the responsibility of recycling.

Cradle-to-cradle production especially 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 and solar energy.
6. Maintaining ecosystems by minimizing the environmental impact of all activities.
7. Using materials and energy sources that are locally available, 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, thus 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 Green Design and Manufacturing. In reviewing the various activities described thus far, note that there are overarching relationships among the basic concepts of DFMA, DFD, DFE, and DFR. These relationships can be summarized as guidelines, now rapidly being accepted worldwide:

1. Reduce waste of materials, by refining product design, reducing the amount of materials used in products, and selecting 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 produce 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, and whenever possible, encourage the use of renewable sources of energy.
6. Encourage recycling by using materials that are a part of either industrial or biological cycling, but not both in the same product assembly. Ensure proper handling and disposal of all waste in the case of materials used that are not part of an industrial or biological cycle.

I.5 Selection of Materials

An increasingly wide variety of materials are now available, each type having its own (a) material properties and manufacturing characteristics, (b) advantages and limitations, (c) material and production costs, and (d) consumer and industrial applications (Part I). The selection of materials for products and their components is typically made in consultation with materials engineers, although design engineers may also be sufficiently experienced and qualified to do so. At the forefront of new materials usage are industries such as the aerospace and aircraft, automotive, military equipment, and sporting goods industries.

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** (also called *smart materials*), **amorphous alloys, semiconductors,** and **superconductors** (Chapters 6, 18 and 28).

As new developments continue, the selection of an appropriate material for a particular application becomes even more challenging. Also, there are continously 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.

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, machined, joined, and heat treated with relative ease. As Table I.3 illustrates, no one material has the same manufacturing characteristics. Another consideration is *appearance*, which includes such characteristics as color, surface texture, and feel, all of which can play a significant role in a product's acceptance by the public.

Availability. As will be emphasized throughout this book, the economic aspect of material selection is as important as technological considerations (Chapter 40). Thus, the availability of materials is a major concern in manufacturing. Furthermore, if materials are not available in the desired quantities, shapes, dimensions, and surface

TABLE I.3

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.

texture, substitute materials 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. Consequently, a country's *self-reliance* on resources, especially energy, is an often-expressed political goal, but is challenging to achieve. Geopolitics (defined briefly as the study of the infuence of a nation's physical geography on its foreign policy) must thus be a consideration, particularly during periods of global hostility.

Service Life. We all have had the experience of a shortened service life of a product, which often can be traced to (a) improper selection of materials, (b) improper selection of production methods, (c) insufficient control of processing variables, (d) defective parts or manufacturing-induced defects, (e) poor maintenance, 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, bolt, or turbine blade or a burned-out electric motor.
- Does not function properly or perform within 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.

Throughout various chapters, this text describes the types of failure of a component or a product resulting, for example, from (a) design deficiencies, (b) improper material selection, (c) incompatibility of materials in contact, which produces friction, wear, and galvanic corrosion, (d) defects in raw materials, (e) defects induced during manufacturing, (f) improper component assembly, and (g) improper product use.

Material Substitution in Products. For a variety of reasons, numerous substitutions are often made in materials, as evidenced by a simple inspection and comparison of common products such as home appliances, sports equipment, or 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) aluminum vs. copper wire, (d) plastic vs. steel car bumper, (e) plastic vs. metal toy, and (f) alloy steel vs. titanium submarine hull.

The two examples that follow give typical details of the major factors involved in material substitution in common products.

EXAMPLE I.2 Baseball Bats

Baseball bats for the major and minor leagues are generally made of wood from the northern white ash tree, a wood that has high dimensional stability, a high elastic modulus and strength-to-weight ratio, and high shock resistance. Wooden bats can, however, break during their use and may cause serious injury. The wooden bats are made on semiautomatic lathes and then subjected to finishing operations for appearance and labeling. The straight uniform grain required for such bats has become increasingly difficult to find, particularly when the best wood comes from ash trees that are at least 45 years old.

For the amateur market and for high school and college players, aluminum bats (top portion of Fig. I.5) have been made since the 1970s as a cost-saving alternative to wood. The bats are made by various metalworking operations, described throughout Part III. Although, at first, their performance was not as good as that of wooden bats, the technology has advanced to a great extent. Metal bats are now made mostly from high-strength aluminum tubing, as well as other metal alloys. The bats are designed to have the same center of percussion (known as the sweet spot, as in tennis racquets) as wooden bats, and are usually filled with polyurethane or cork for improved sound damping and for controlling the balance of the bat.

Metal bats possess such desirable performance characteristics as lower weight than wooden bats, optimum weight distribution along the bat's length, and superior impact dynamics. Also, as documented by scientific studies, there is a general consensus that metal bats outperform wooden bats.

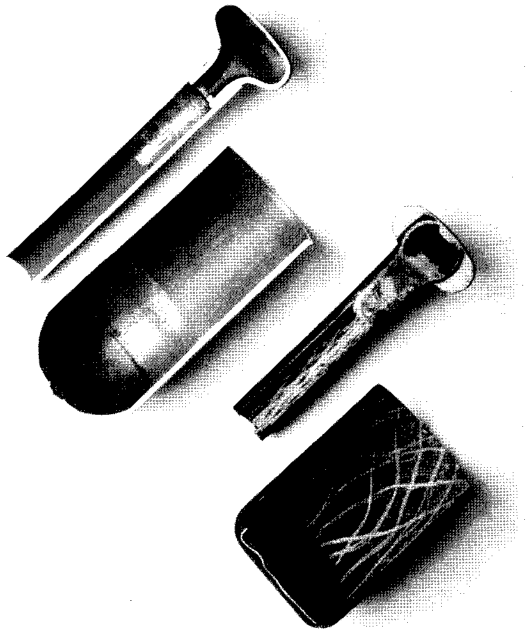


FIGURE I.5 Cross sections of baseball bats made of aluminum (top two) and composite material (bottom two).

Developments in bat materials now include composite materials (Chapter 9) consisting of high-strength graphite and glass fibers embedded in an epoxy resin matrix. The inner woven sleeve (lower portion of Fig. I.5) is made of Kevlar fibers (an aramid), which add strength to the bat and dampen its vibrations. These bats perform and sound much like wooden bats.

Source: Mizuno Sports, Inc.

EXAMPLE I.3 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 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

As will be described throughout this text, there is often more than one method that can be employed to produce a component for a product from a given material. The following broad categories of manufacturing methods are all applicable to metallic as well as nonmetallic materials:

a. **Casting** (Fig. I.6a): Expendable mold and permanent mold (Part II).

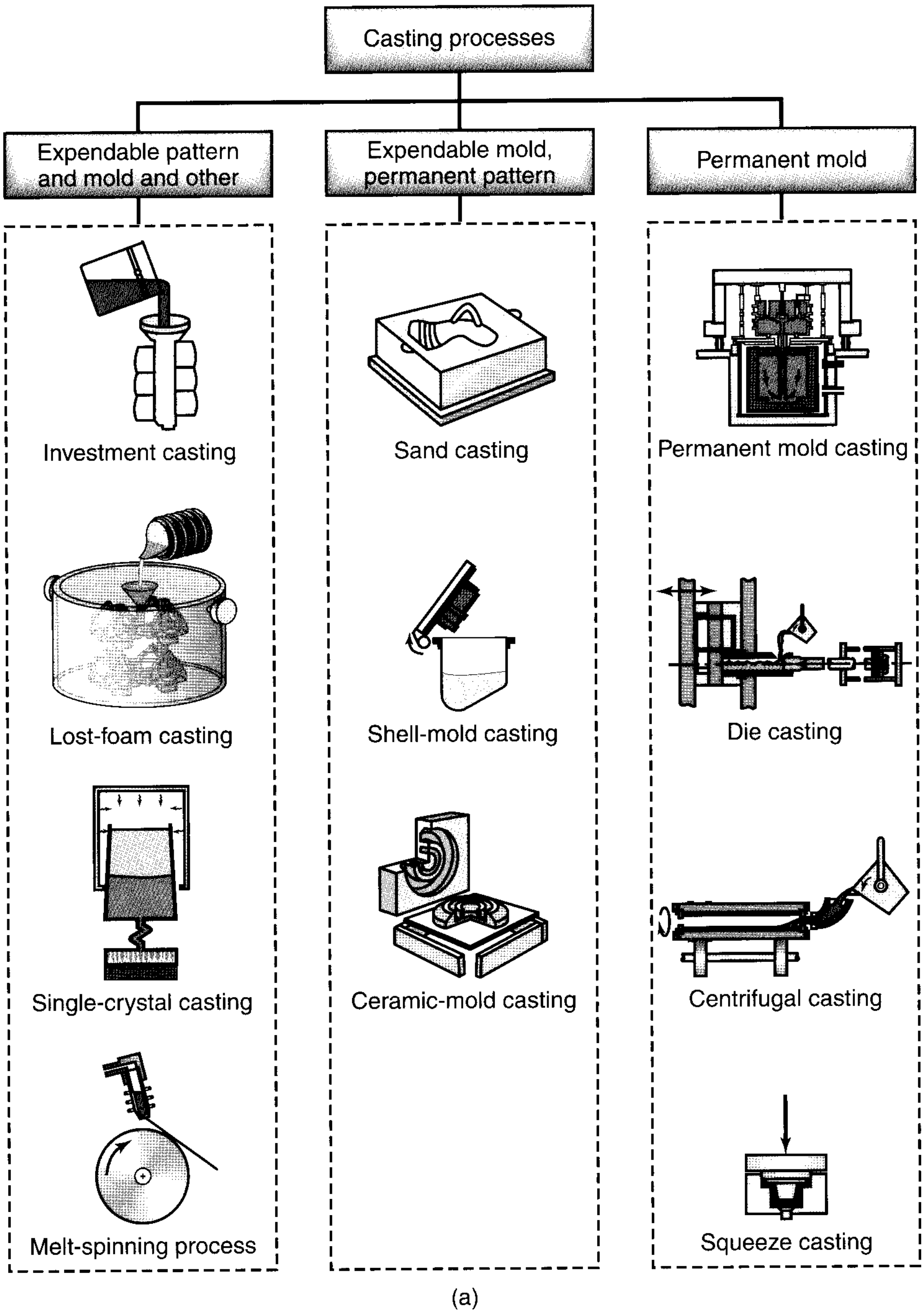


FIGURE I.6a Schematic illustrations of various casting processes.

b. **Forming and shaping** (Figs. I.6b through I.6d): Rolling, forging, extrusion, drawing, sheet forming, powder metallurgy, and molding (Part III).

c. **Machining** (Fig. I.6e): 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).

d. **Joining** (Fig. I.6f): Welding, brazing, soldering, diffusion bonding, adhesive bonding, and mechanical joining (Part VI).

e. **Finishing:** Honing, lapping, polishing, burnishing, deburring, surface treating, coating, and plating (Chapters 26 and 34).

f. **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

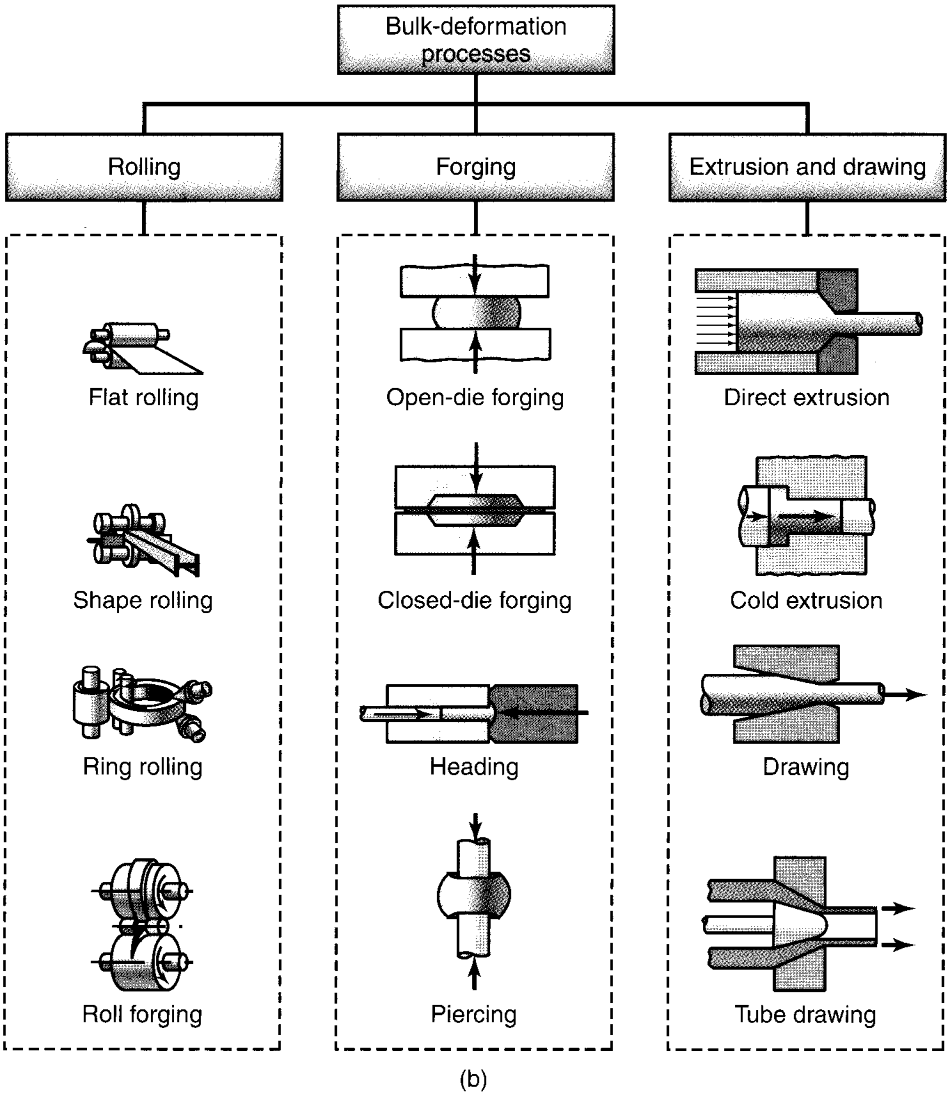


FIGURE I.6b Schematic illustrations of various bulk-deformation processes.

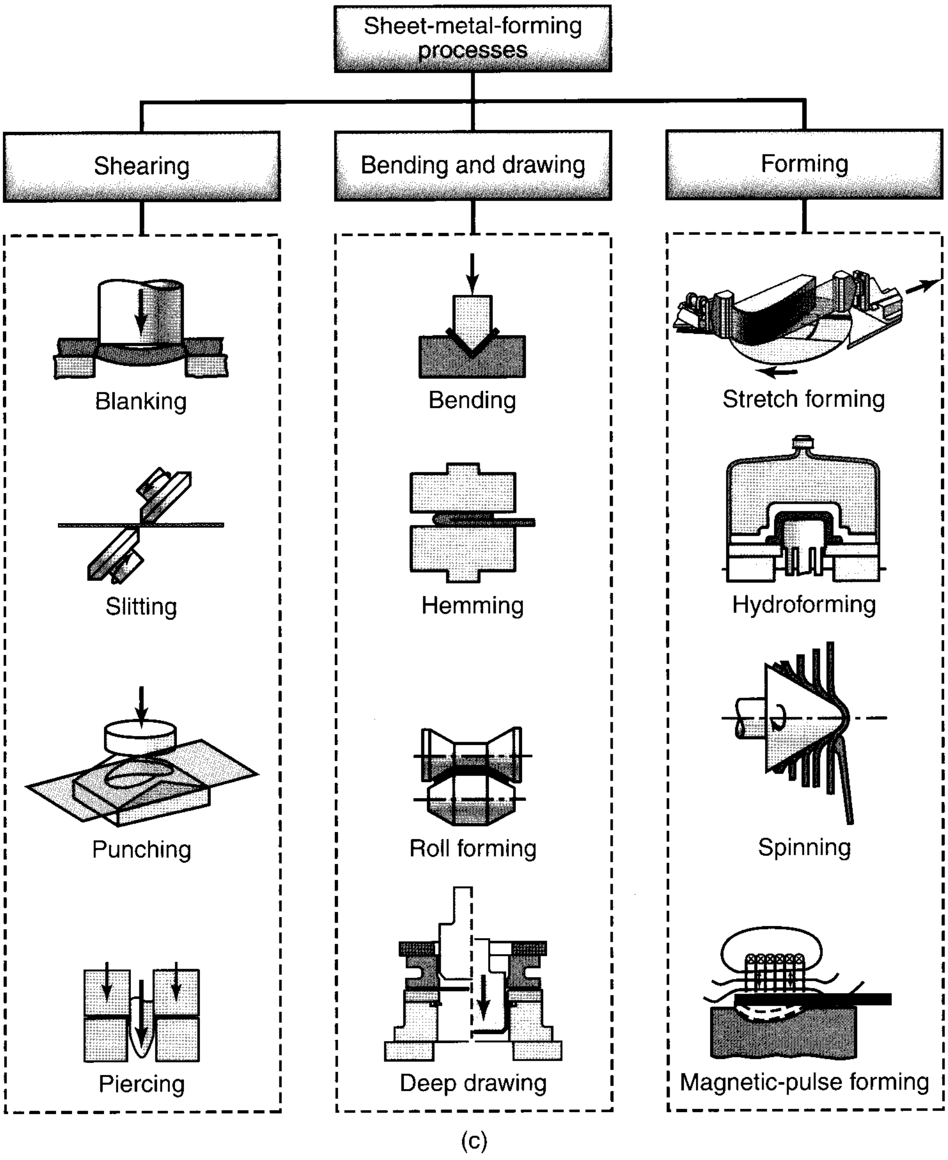


FIGURE I.6c Schematic illustrations of various sheet-metal-forming processes.

systems (MEMS) and nanoelectromechanical systems (NEMS), typically involving processes such as lithography, surface and bulk 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 produced, including the dimensional tolerances 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:

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

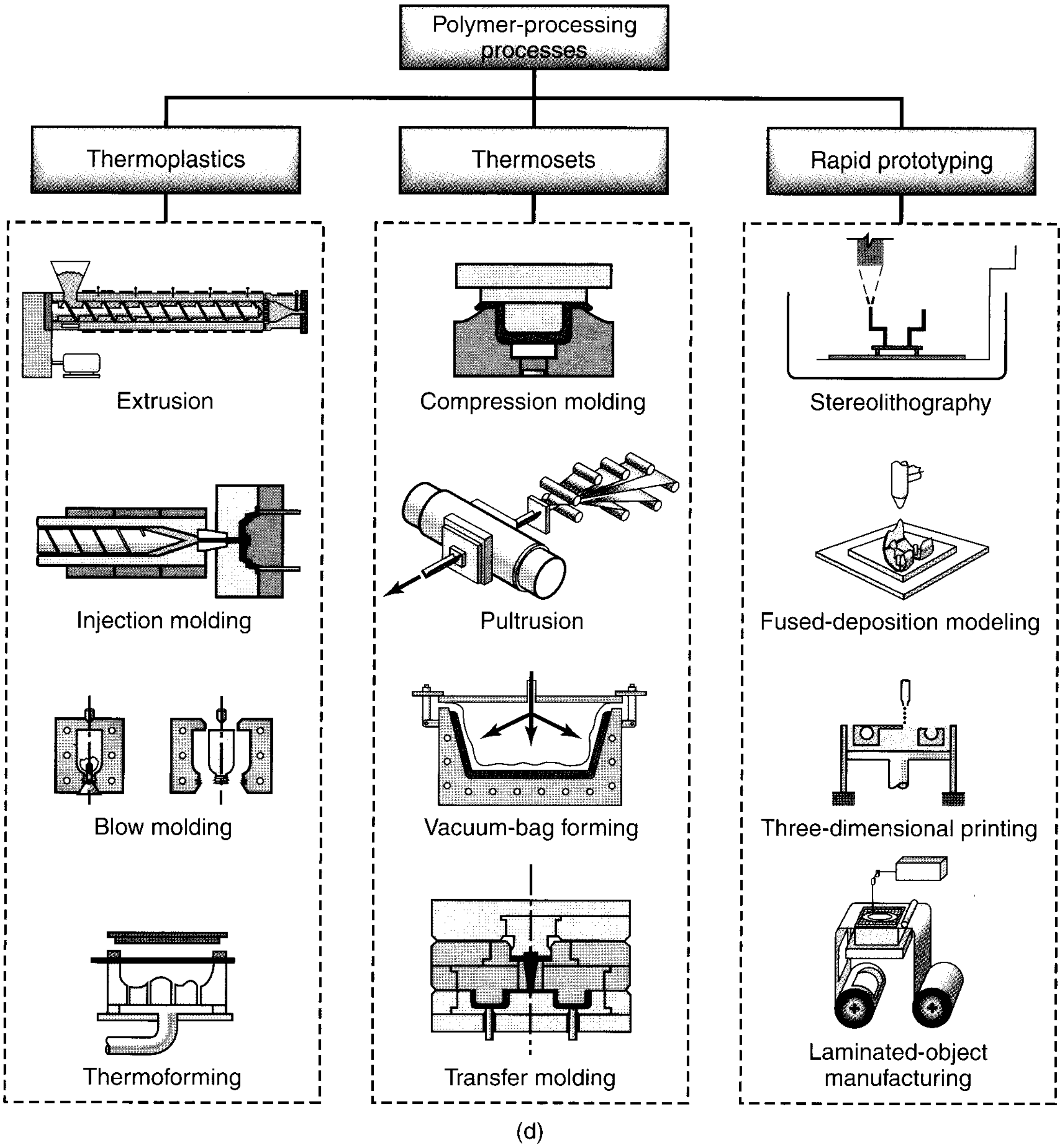


FIGURE I.6d Schematic illustrations of various polymer-processing methods.

b. Metals that have been preshaped at room temperature become less formable during subsequent processing, which, in practice, is often required 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 problems 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 are now being replaced by laser cutting, as shown in Fig. I.7 on p. 24, thus eliminating the need for hard tools, which have only fixed shapes and can be expensive and time consuming to make.

The laser path in this cutting operation is computer controlled, thereby increasing the operation's flexibility and its capability of producing an infinite variety of

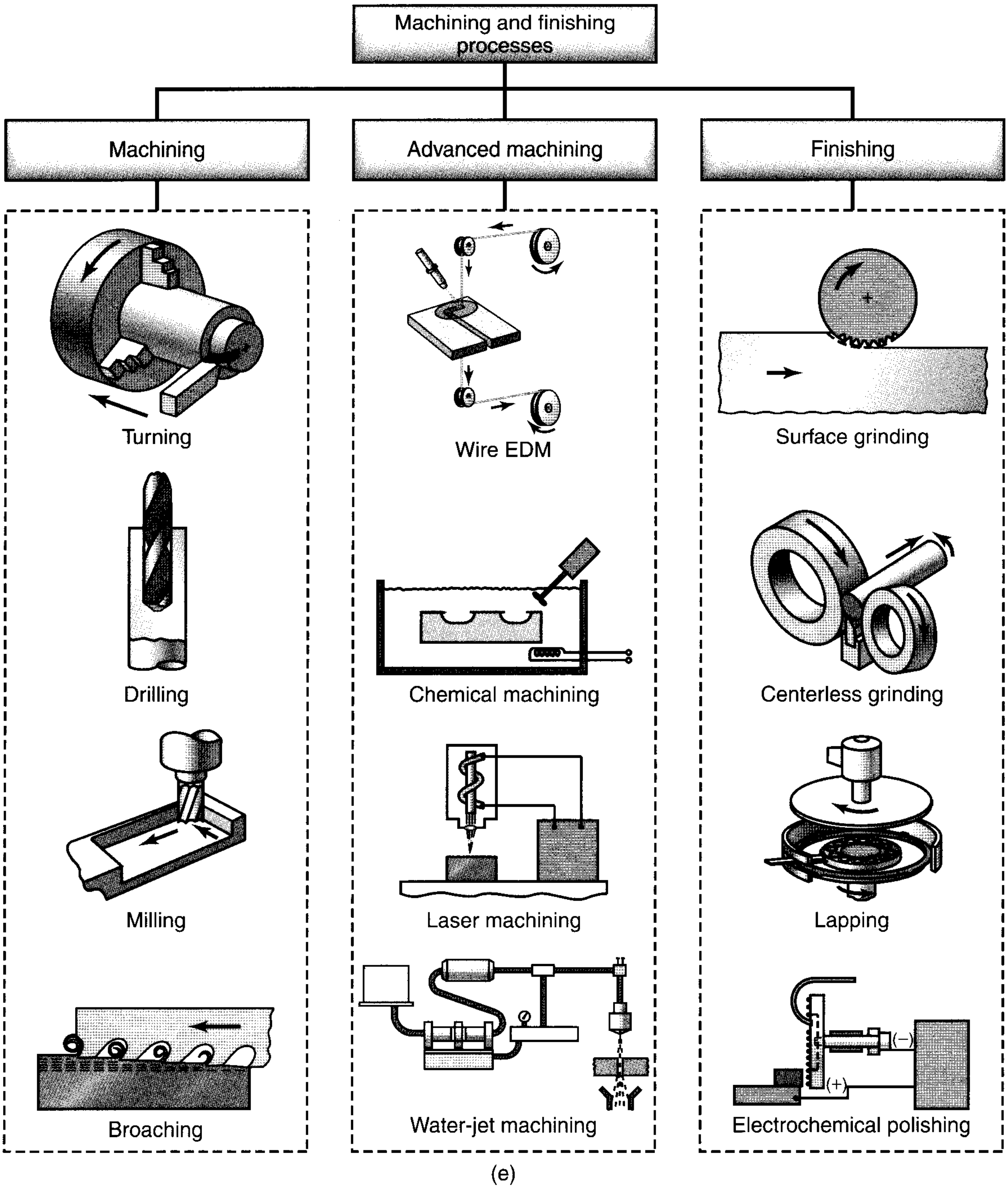
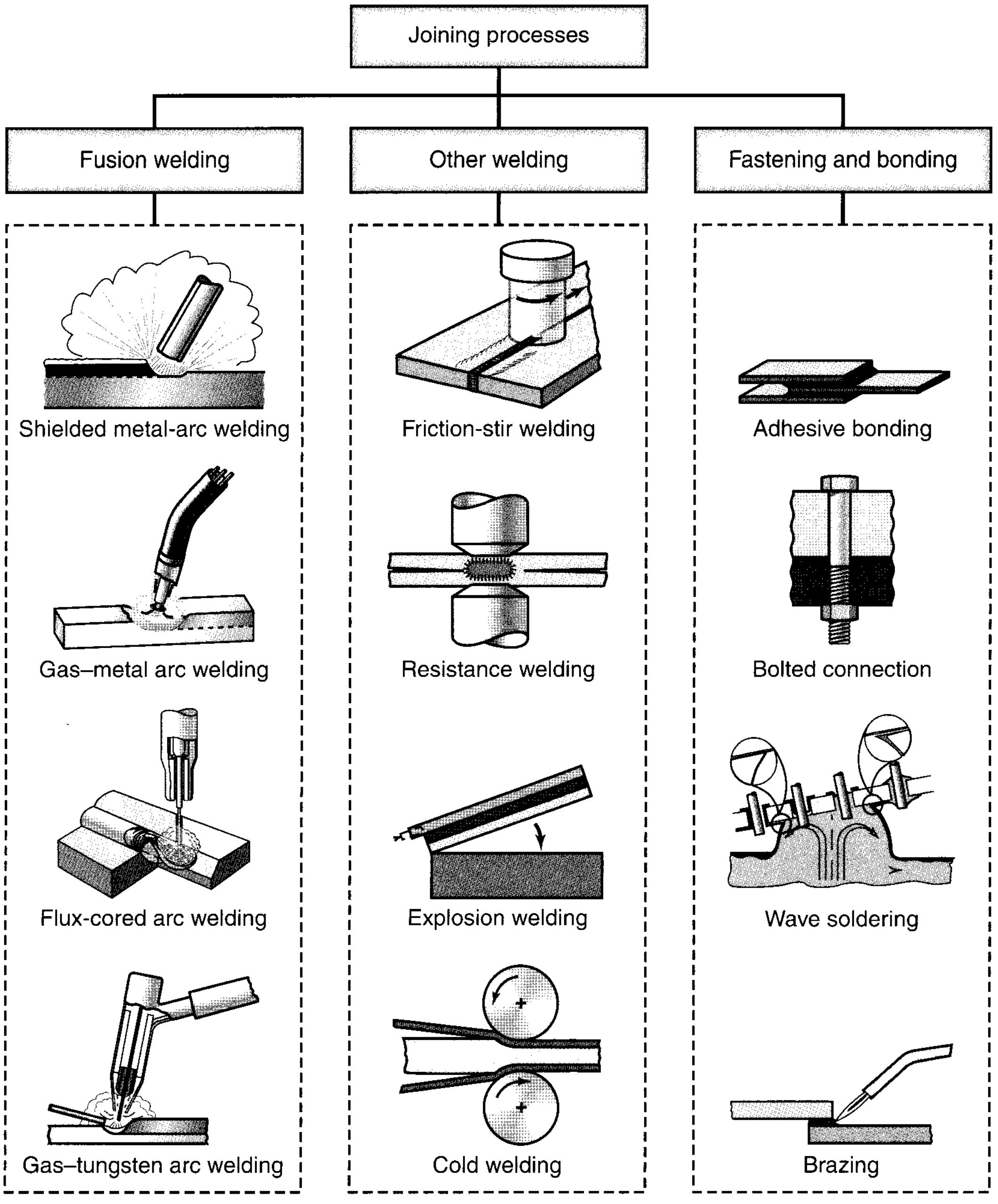


FIGURE I.6e Schematic illustrations of various machining and finishing processes.

shapes accurately, repeatedly, and economically. However, because of the high heat involved in using lasers, the surfaces produced after cutting have very different characteristics (such as discoloration and a different surface texture) than those produced by traditional methods. This difference can have significant effects not only on the appearance of the material, but especially on its subsequent processing and in the



(f)

FIGURE I.6f Schematic illustrations of various joining processes.

service life of the product. Moreover, the inherent flexibility of laser cutting is countered by the fact that it is a much slower operation than traditional punching.

In process selection, several factors can have a major role, such as the 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 the part size and the

level of complexity, the parts may be precision cast, fabricated from individual pieces, or produced by powder-metallurgy techniques.

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

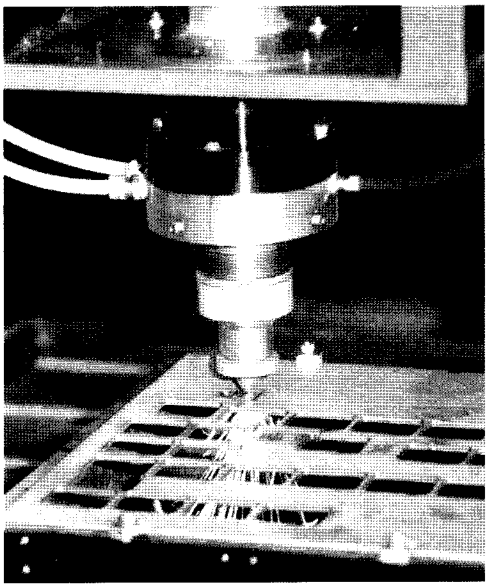
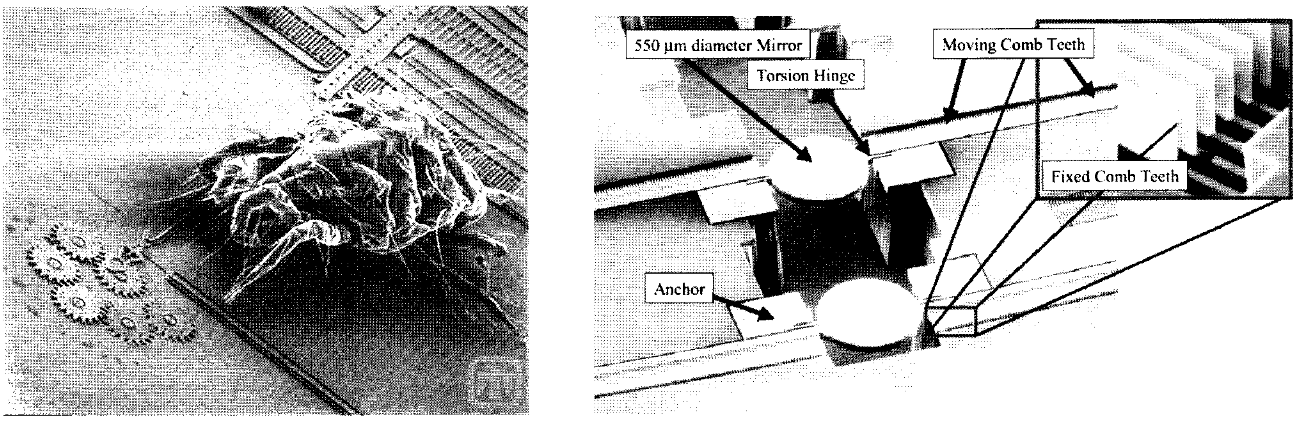


FIGURE I.7 Cutting sheet metal with a laser beam. *Source*: Courtesy of Rofin-Sinar, Inc., and Society of Manufacturing Engineers.

The size of manufactured products, and the machinery and equipment involved in processing them, vary widely, ranging from microscopic gears and machanisms of micrometer size, as illustrated in Fig. I.8, to (a) the main landing gear for the Boeing 777 aircraft, which is 4.3 m high and includes three axles and six wheels; (b) the runner for the turbine for a hydroelectric power plant, which is 4.6 m in diameter and weighs 50,000 kg; and (c) a large steam turbine rotor weighing 300,000 kg.

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 (that may have been employed in the past) may well have to be substituted with another. Consider, for example, the following products that can be produced by any of the sets of processes indicated: (a) cast vs. forged crankshaft, (b) stamped sheet-metal vs. forged or cast automobile wheels, (c) cast vs. stamped sheet-metal frying pan, (d) injection molded vs. extruded or cast polymer bracket, and (e) welded vs. riveted sheet-metal safety hood for a machine.

Many varieties of such products are widely available in the marketplace. However, a customer's preference will depend on his or her particular needs, which include factors such as the product's appeal to the customer, its cost, whether maintenance is required,



(a) (b)

FIGURE I.8 (a) Microscopic gears with dust mite. *Source*: Courtesy Sandia National Laboratory; (b) A movable micromirror component of a light sensor. Note the 100-μm scale at the bottom of the figure. *Source*: Courtesy of R. Mueller, University of California at Berkeley.

whether the product is for industrial or consumer use, the parameters to which the product will be subjected (such as temperatures and chemicals), and any environmental concerns that have to be addressed.

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 desired 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 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 operation is often necessary to produce the part. For example, a cast or forged gear or crankshaft generally will not have the necessary dimensional characteristics, thus requiring 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 are now reaching the magnitude of the atomic lattice. Various techniques, including the use of highly sophisticated technologies (see micromechanical and microelectromechanical device fabrication in Chapter 29), are rapidly being developed to attain such extreme accuracy. Also, mirror-like surfaces on metals can now be produced by machining with a very sharp diamond with a nose radius of 250 micrometers as the cutting tool. The machine is highly specialized, with very high stiffness (to minimize deflections, as well as vibration and chatter, during machining) and 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 (e.g., the annual quantity) and the rate (number of pieces made per unit time) are important economic considerations in determining the appropriate processes and the types of machinery required. Note, for example, that light bulbs, beverage cans, fuel-injection nozzles, and hubcaps are produced in numbers and at rates that are much higher than those for jet engines and tractors.

Following is a brief outline of the general types of production, in increasing order of annual quantities produced:

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

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

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

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

EXAMPLE I.4 Saltshaker and Pepper Mill

The saltshaker and pepper mill set shown in Fig. I.9 consists of metallic as well as 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 desirable characteristics for this application and 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 (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 a rolling operation (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).

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*: Reproduced with permission from *Success Stories on P/M Parts*, Metal Powder Industries Federation, Princeton, NJ, 1998.

I.7 Computer-integrated Manufacturing

Computer-integrated manufacturing (CIM), as the name suggests, integrates the software and 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.

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 the 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

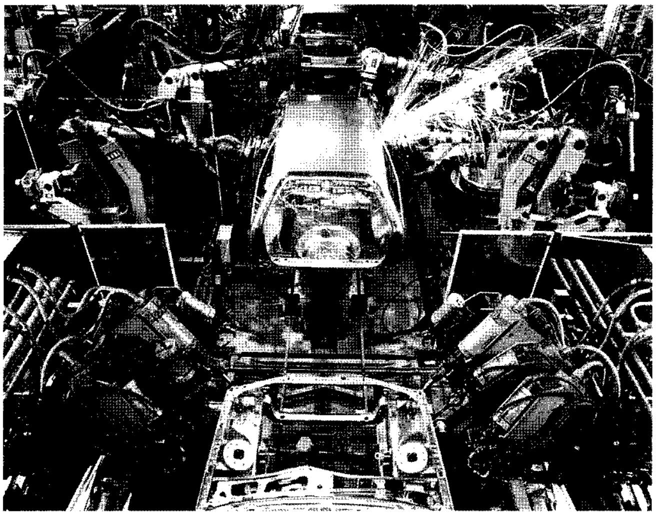


FIGURE I.10 Automated spot welding of automobile bodies in a mass-production line. *Source*: Courtesy of Ford Motor Company.

and to minimize manufacturing cost. For example, machining forces, temperature, surface finish, and the dimensions of the part can be constantly monitored; if they move outside the specified range, the system adjusts the appropriate variables until the parameters are within the specified range.

3. **Industrial robots.** Introduced in the early 1960s, industrial robots (Fig. I.10) have rapidly been replacing humans, especially in operations that are repetitive, dangerous, and boring. As a result, variability in product quality is decreased and productivity improved. Robots are particularly effective in assembly operations, and some (*intelligent robots*) have been developed with sensory-perception capabilities and movements that simulate those of humans.
4. **Automated materials handling.** Computers have made possible highly efficient handling of materials and components 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 continue to be developed to replace assembly by human operators, although humans still have to perform some operations. Assembly costs can be high, depending on the type of product; consequently, products are now being designed so that they can be assembled more easily, and faster by automated machinery, thus reducing the total manufacturing cost.
6. **Computer-aided process planning** (CAPP). By optimizing process planning, this system is capable of improving productivity, product quality, and consistency and hence 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 parts can be grouped and produced by classifying them into families according to similarities in design and the manufacturing processes employed to produce them. In this way, part designs and process plans can be standardized and new

parts (based on similar parts made previously) can be produced efficiently and economically.

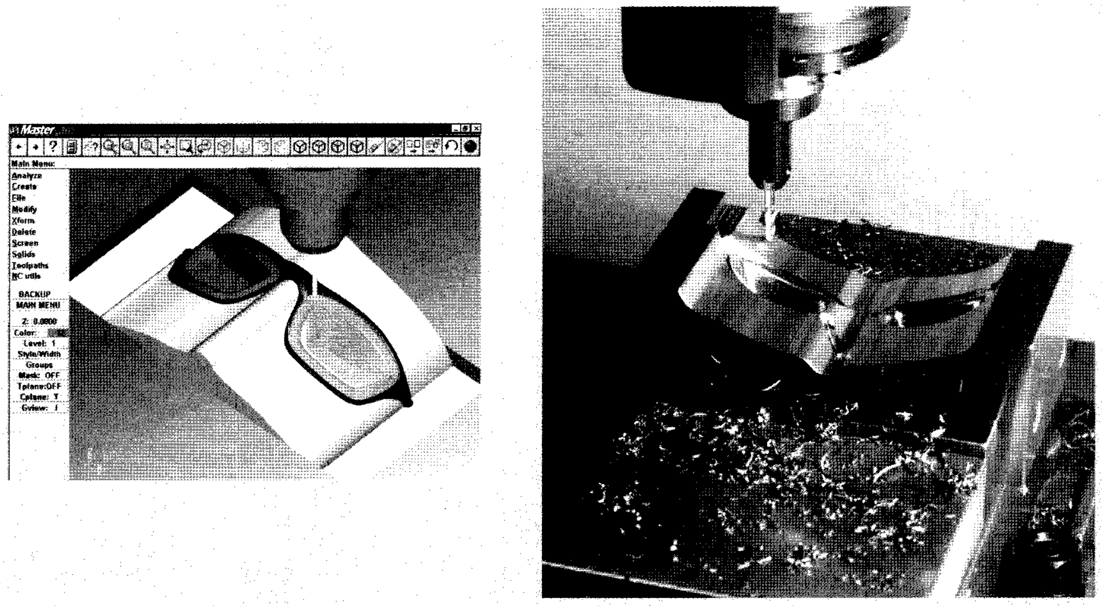
8. **Just-in-time production** (JIT). The principle behind JIT is that (1) supplies of raw materials and parts are delivered to the manufacturer just in time to be used, (2) parts and components are produced just in time to be made into subassemblies, and (3) products are assembled and finished just in time to be delivered to the customer. As a result, inventory carrying costs are low, defects in components are detected right away, productivity is increased, and high-quality products are made at low cost.
9. **Cellular manufacturing** (CM). This system utilizes workstations that consist of a number of *manufacturing cells*, each containing various production machines controlled by a central robot, with each machine performing a different operation on the part, including inspection.
10. **Flexible manufacturing systems** (FMS). These systems integrate manufacturing cells into a large production facility, with all of the cells interfaced with a central computer. Although very costly, flexible manufacturing systems are capable of producing parts efficiently, but in relatively small quantities, and of quickly changing manufacturing sequences required for different 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 various tasks and solving difficult real-life problems, much as human experts would, including expediting the traditional iterative process in design optimization.
12. **Artificial intelligence** (AI). Computer-controlled systems are now capable of learning from experience and of making decisions that optimize operations and minimize costs, ultimately replacing human intelligence.
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 processes, diagnosing problems in machine performance, and conducting financial planning and managing a company's manufacturing strategy.

EXAMPLE I.5 Mold for Making Sunglasses Frames

The metal mold used for injection molding of plastic sunglasses is made on a computer numerical-control milling machine, by using a cutter (called a ball-nosed end mill), as illustrated in Fig. I.11. First, a model of the sunglasses is made using a computer-aided design software package, from which a model of the mold is automatically generated. The geometric information is sent to the milling machine, and the machining steps are planned.

Next, an offset is added to each surface to account for the nose radius of the end mill during machining, thus determining the cutter path (i.e., the path followed by the center of rotation of the machine spindle). The numerical-control programming software executes this machining program on the milling machine, producing the die cavity with appropriate dimensions and tolerances. Electrical-discharge machining (Section 27.5) can also be used to make this mold; however, it was determined that the procedure was about twice as expensive as machining the mold by computer numerical control, and it produced molds with lower dimensional accuracy.

Source: Courtesy of Mold Threads, Inc.



(a) (b)

FIGURE I.11 Machining a mold cavity for making sunglasses. (a) Computer model of the sunglasses as designed and viewed on the monitor. (b) Machining of the die cavity, using a computer numerical-control milling machine. *Source*: Courtesy of Mastercam/CNC Software, Inc.

I.8 Quality Assurance and Total Quality Management

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

Because products are typically made through several manufacturing steps and operations, each step can involve its own significant variations in performance, which can occur even within a relatively short time. A production machine, for example, may perform differently when it is first turned on than when it warms up during its use or when the ambient temperature in the plant fluctuates. Consequently, *continuous control of processes* (known as *online monitoring*) is a critical factor in maintaining product quality, and the objective must be 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 manufacture 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, as shown in Table I.4,
- Is suitable for its intended purposes, and
- Can be maintained with relative ease.

TABLE I.4

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 observed, depending on the quality of the product and how well it has been maintained.

TABLE I.5

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

As Table I.5 indicates, producing defective products can be very costly to the manufacturer, with costs varying by orders of magnitude.

Pioneers in quality control, particularly W.E. Deming (1900–1993), J.M. Juran (1904–2008), and G. Taguchi (1924–), 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 indication of strict quality control, computer chips are now made 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 *sigma*. Three sigma in manufacturing would result in 2700 defective parts per million, which is much too high 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 in which the factors involved in a manufacturing process and their interactions are studied simultaneously. For example, the variables affecting dimensional accuracy or surface finish in a machining operation can be identified readily, thus making it possible for appropriate preventive on-time actions 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). 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, and they are now the world standard for quality.

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

Some examples of the need for proper ergonomic considerations are represented by (a) a mechanism that is difficult to operate manually, causing injury to the worker, (b) a poorly designed keyboard that causes pain to the user's hands and arms during its normal use (known as *repetitive stress syndrome*), and (c) a control panel on a machine that is difficult to reach or use safely and comfortably.

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

As is widely known, a product's malfunction or failure can cause bodily injury or even death, as well as financial loss to an individual, to bystanders, or to an organization. This important topic is referred to as *product liability*, and the laws governing it generally vary from state to state and from country to country. Among the numerous examples of products that could involve liability are the following:

- A grinding wheel shatters and blinds a worker.
- A cable supporting a platform snaps, allowing the platform to drop and cause bodily harm or death.
- Automotive brakes suddenly become inoperative because of the failure of a particular component of the brake system.
- Production machinery lacks appropriate safety guards.
- Electric and pneumatic tools lack appropriate warnings and instructions for their safe use.
- Aircraft landing gears fail to descend and lock properly.

I.9 Lean Production and Agile Manufacturing

Lean production is a methodology that involves a thorough assessment of each activity of a company, with the basic purpose of minimizing 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 from the viewpoint of the customer 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 operation, (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 those for productive and for nonproductive labor.

The lean production strategy requires a fundamental change in corporate culture, as well as an understanding of the importance of *cooperation and teamwork* among the company's workforce and management. Lean production does not necessarily require cutting back on a company's physical or human resources; rather, 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.

Agile Manufacturing. The principle behind *agile manufacturing* is ensuring *agility*—and hence *flexibility*—in the manufacturing enterprise, so that it can respond rapidly and effectively to changes in product demand and the needs of the customer. Flexibility can be achieved through people, equipment, computer hardware and software, and advanced communications systems. As an example of this approach, it has been predicted that the automotive industry could configure and build a car in three days and that, eventually, the traditional assembly line will be replaced by a system in which a nearly custom made car will be produced by combining several individual modules.

The methodologies of both lean and agile production require that a manufacturer **benchmark** its operations. Benchmarking involves assessing the competitive position of other manufacturers with respect to one's own position (including product quality, production time, and manufacturing cost) and setting realistic goals for the future. Benchmarking thus becomes a *reference point* from which various measurements can be made and to which they can be compared.

I.10 Manufacturing Costs and Global Competition

Always critically important, the 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*, 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.6. The percentages indicated can, however, vary significantly depending on product type.

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

1. **Materials.** Raw-material costs depend on the material itself, as well as on supply and demand. Low cost may not be the deciding factor if the cost of processing a

TABLE I.6

Typical Cost Breakdown in Manufacturing	
Design	5%
Materials	50%
Manufacturing	
Direct labor	15%
Indirect labor	30%

particular material is higher than that for a more expensive material. For example, a low-cost piece of material may require more time to machine or form than one of higher cost, thus increasing production costs.

2. **Tooling.** Tooling costs include those for cutting tools, dies, molds, workholding devices, and fixtures. Some cutting tools cost as little as \$2 and as much as about \$100 for materials such as 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 sheet metal to make automobile fenders.
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 thousand to over a million dollars. Although the cost of computer-controlled machinery can be very high, such an expenditure may well be warranted in view of its long-range benefit of reducing 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* pertains to servicing of the total manufacturing operation; it is also called *nonproductive labor* or *overhead*. Direct-labor costs may be only 10 to 15% of the total cost (Table I.6), but it can be as much as 60% for labor-intensive products. Reductions in the direct-labor share of manufacturing costs can be achieved by such means as extensive automation, computer control of all aspects of manufacturing, the implementation of modern technologies, and increased efficiency of operations.

As shown in Table I.7, there is a worldwide disparity in labor costs, by an order of magnitude. It is not surprising that today numerous consumer products are manufactured mostly in Pacific Rim countries, especially China, or they are assembled in Mexico. Likewise, software and information technologies are often much less costly to develop in countries such 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 such countries.

Outsourcing. A more recent trend has been *outsourcing*, defined as the purchase by a company of parts or labor from an outside source, from either another company or another country, in order to reduce design and manufacturing costs. There is increasing evidence, however, that, depending on the type of product, manufacturing abroad can have significant challenges, including the rising cost of shipping. Another problem is the social impact and political implications of any ensuing lowered employment, especially in the European Union countries and the United States.

TABLE I.7

Approximate Relative Hourly Compensation for Workers in Manufacturing in 2006 (United States = 100)

Norway	154	Italy	96
Germany	137	Japan	81
Denmark	127	Spain	73
Austria	122	Korea (South)	56
Belgium	121	New Zealand	54
Switzerland	119	Israel	48
Netherlands	118	Singapore	45
Finland	117	Portugal	32
Sweden	114	Czech Republic	27
France, United Kingdom	112	Argentina, Slovakia	22
Ireland	103	Poland	21
United States, Australia	**100**	Mexico	12
Canada	97	China, Philippines	5

Note: Compensation can vary significantly with benefits. *Source*: U.S. Department of Labor.

I.11 General Trends in Manufacturing

Following are some general trends that have been observed regarding various aspects of manufacturing today:

Global manufacturing trends

1. Product variety and complexity continue to increase.
2. Product life cycles are becoming shorter.
3. Markets have become multinational and global competition has been increasing rapidly.
4. Market conditions fluctuate widely.
5. Customers are consistently demanding high-quality, low-cost products and on-time delivery.

Materials

6. Material composition, purity, and defects (impurities, inclusions, and flaws) are coming under more control in order to further enhance overall properties, manufacturing characteristics, reliability, and service life.
7. Developments have occurred in the selection of materials for improved recyclability.
8. Developments continue in nanomaterials, nanopowders, composites, superconductors, semiconductors, amorphous alloys, shape-memory alloys (smart materials), tool and die materials, and coatings.
9. Testing methods and equipment, including the use of advanced computers and software, particularly for ceramics, carbides, and composite materials, are continually being improved.
10. Increasing control over the thermal treatment of materials is resulting in more predictable and reliable properties.

11. Weight savings are being achieved with the use of materials with higher strength-to-weight and stiffness-to-weight ratios, particularly in the automotive and aerospace industries.

Manufacturing operations

12. Improvements are being made in predictive models of the effects of material-processing parameters on product integrity, applied during a product's design stage.
13. Developments continue in ultraprecision manufacturing, micromanufacturing, and nanomanufacturing, approaching the level of atomic dimensions.
14. Computer simulation, modeling, and control strategies are being applied to all areas of manufacturing.
15. Rapid-prototyping technologies are increasingly being applied to the production of tooling and direct digital manufacturing.
16. Optimization of manufacturing processes and production systems are making them more agile.

Manufacturing systems

17. Advances in computer software and hardware are being applied to all aspects of production.
18. Developments have occurred in control systems, industrial robots, automated inspection, handling and assembly, and sensor technology.
19. Lean production and information technology are being implemented as tools to help meet major global challenges.

Goals in manufacturing

20. View manufacturing activities not as individual, separate tasks, but as making up a *large system*, with all its parts interrelated.
21. Meet all *design requirements, product specifications*, and relevant *national and international standards* for products.
22. Build *quality* into the product at each stage of its production.
23. Implement the most *economical and environmentally friendly (green)* manufacturing methods.
24. Continually evaluate advances in *materials, production methods*, and *computer integration*, with a view toward realizing their appropriate, timely, and economical implementation.
25. Adopt production methods that are sufficiently *flexible* in order to rapidly respond to changing global market demands and provide on-time delivery to the customer.
26. Continue efforts aimed at achieving *higher levels of productivity* and eliminating or minimizing waste with optimum use of an organization's resources.
27. Cooperate with customers for timely feedback for *continuous improvement* of a company's products.

Fundamentals of Materials: Their Behavior and Manufacturing Properties

PART I

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

In order to emphasize the importance of the topics to be described, let's look at a typical automobile as an example of a common product that utilizes a wide variety of materials, as shown in Fig. I.1. These materials were selected primarily because, not only did they possess the desired properties and characteristics for the intended functions of specific parts, but also they were the ones that could be manufactured at the lowest cost.

For example, steel was chosen for much of the body because it is strong, easy to shape, and inexpensive. Plastics were used in many components because of characteristics such as light weight, resistance to corrosion, a wide selection of colors, and ease of manufacturing into complex shapes and at low cost. Glass was chosen for all the windows not only because it is transparent, but also because it is

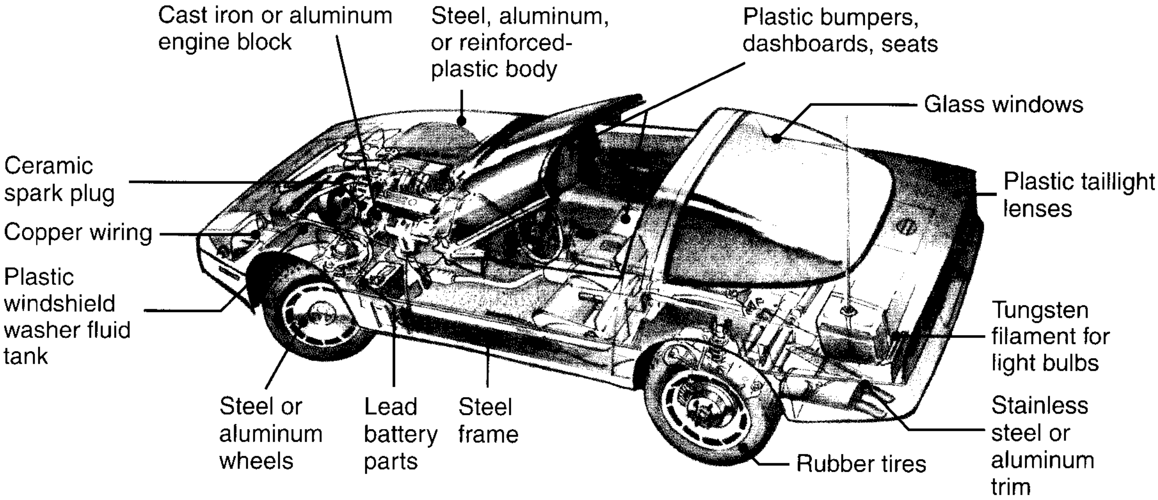


FIGURE I.1 Some of the metallic and nonmetallic materials used in a typical automobile.

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 small screws to wheels. In recent years, fuel economy and the need for improved performance have driven the substitution of materials, 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, the selection of materials for individual components in a product requires a thorough understanding of their properties, functions, and manufacturing costs. A typical automobile is an assemblage of some 15,000 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. The task of engineers thus becomes very challenging, especially with the ever-increasing variety of materials that are now available, as outlined in Fig. I.2.

A general outline of the topics described in Part I is given in Fig. I.3. The fundamental knowledge presented on the behavior, properties, and characteristics of materials will help the reader understand their relevance to the manufacturing processes described in Parts II through IV. This knowledge will also make it possible for us to analyze the often complex interrelationships among materials, manufacturing processes, machinery and tooling, and the economics of manufacturing operations.

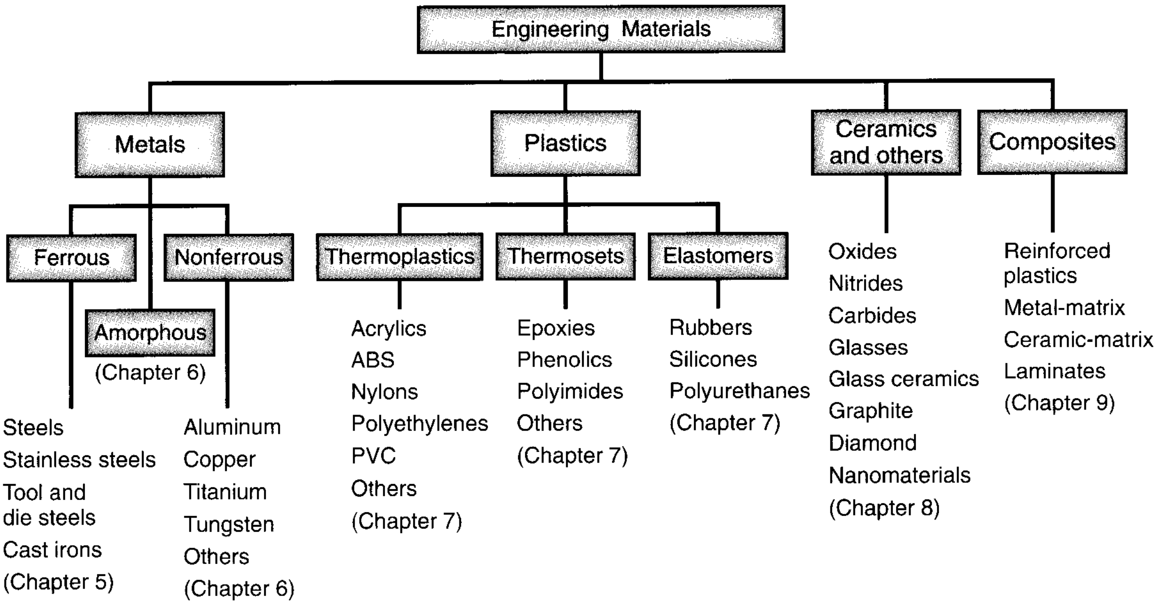


FIGURE I.2 An outline of the engineering materials described in Part I.

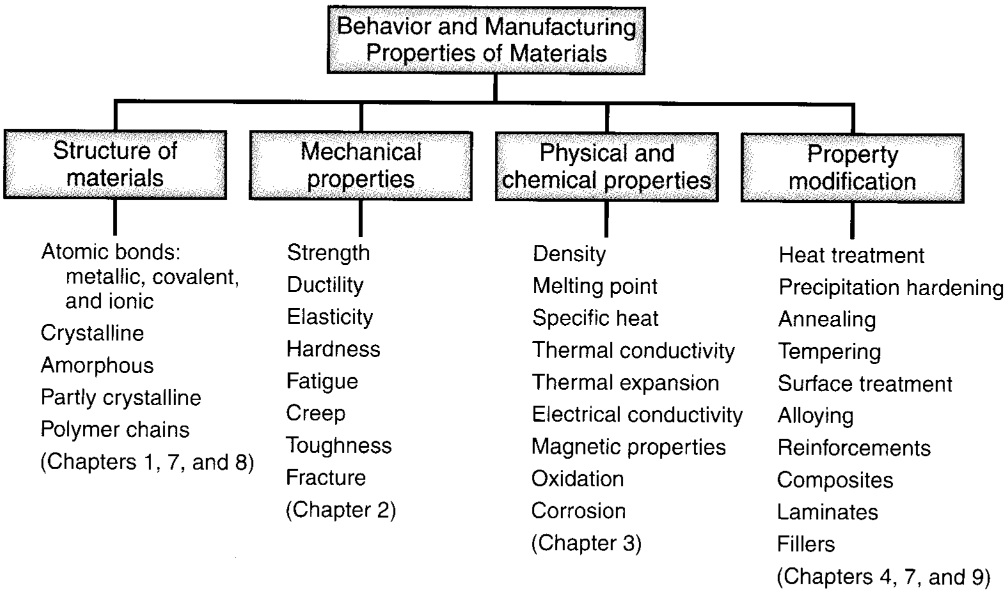


FIGURE I.3 An outline of the behavior and the manufacturing properties of materials decribed in Part I.

CHAPTER

1

The Structure of Metals

- This chapter introduces the subject of the crystalline structure of metals and explains why it is important in determining their properties and behavior.
- It begins with a review of the types of atomic bonds and explains the importance of ionic, covalent, and metallic bonds.
- Metal structures and the arrangement of atoms within the structure are then examined, and the types of imperfections in the crystal structure and their effects are presented.
- Recognizing that even a small metal pin contains thousands of crystals, or grains, the effects of grains and grain boundaries are examined, followed by a discussion of strain hardening and anisotropy of metals.
- All of these topics are examined within the framework of their relevance to manufacturing processes.

1.1 Introduction

Why are some metals hard and others soft? Why are some metals brittle, while others are ductile and thus can be shaped easily without fracture? Why is it that some metals can withstand high temperatures, while others cannot? Why is it that a sheet metal may behave differently when stretched in one direction versus another?

These and numerous other questions can be answered by studying the **atomic structure** of metals—that is, the arrangement of the atoms within the metals—because atomic structure greatly influences the properties and behavior of those metals. This knowledge then guides us in controlling and predicting the behavior and performance of metals in various manufacturing processes. Understanding the structure of metals also allows us to predict and evaluate their **properties**, thus allowing us to make appropriate selections for specific applications under various conditions. In addition to atomic structure, several other factors influence the properties and behavior of metals. Among these are the composition of the metal, impurities and vacancies in the atomic structure, grain size, grain boundaries, environment, size and surface condition of the metal, and the methods by which metals are made into products.

The topics described in this chapter and their sequence are outlined in Fig. 1.1. The structure and general properties of materials other than metals are described in

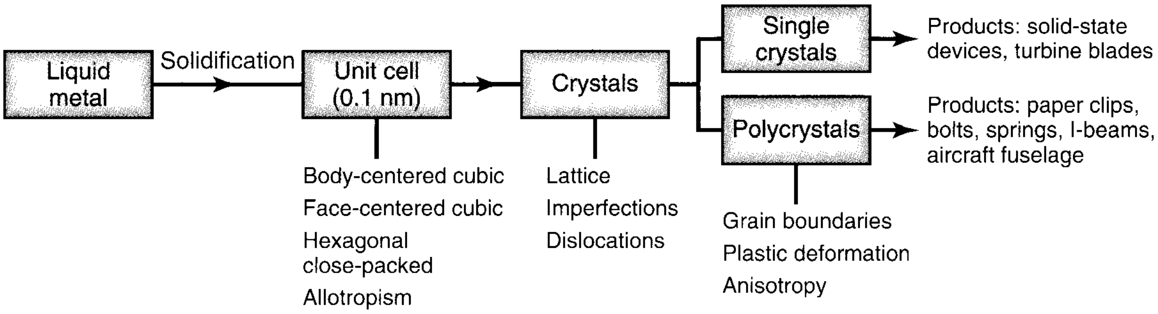


FIGURE 1.1 An outline of the topics described in Chapter 1.

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.

1.2 Types of Atomic Bonds

All matter is made up of **atoms** containing a **nucleus** of *protons* and *neutrons* and surrounding clouds, or orbits, of *electrons*. In recent decades, a large number of subatomic particles have been identified, with additional complexities in the nucleus. For the purposes of this textbook, however, the planetary model of a nucleus orbited by electrons is sufficient.

The number of protons in the nucleus determines whether an atom will be metallic, nonmetallic, or semimetallic. 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 combine to form **molecules**. Molecules are held together by attractive forces called **bonds** through electron interaction. The basic types of atomic attraction associated with electron transfer, called **primary** or **strong bonds**, are as follows:

- **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; such salt consists of Na^+ and Cl^- ions (hence the term *ionic bond*), which are strongly attracted to each other. Also, the attraction is between all adjacent ions, allowing crystalline structures to be formed, as discussed in Section 1.3. Molecules with ionic bonds generally have poor ductility and low thermal and electrical conductivity.
- **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," etc. Polymers consist of large molecules that are covelantly bonded together; water (H_2O) and nitrogen gas (N_2) are additional common examples of molecules formed from 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; thus, they cannot complete the outer shell of other self-mated atoms. Instead, metals and alloys form *metallic bonds*, whereby the available electrons are shared by all atoms in contact. The resultant electron cloud provides attractive forces to hold the atoms together and results in generally high thermal and electrical conductivity.

In addition to the strong attractive forces associated with electrons, weak or secondary attractions occur between molecules. Also referred to as **van der Waals forces**, these forces arise from the attraction of opposite charges without electron transfer. As an example, water molecules consist of one oxygen 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.

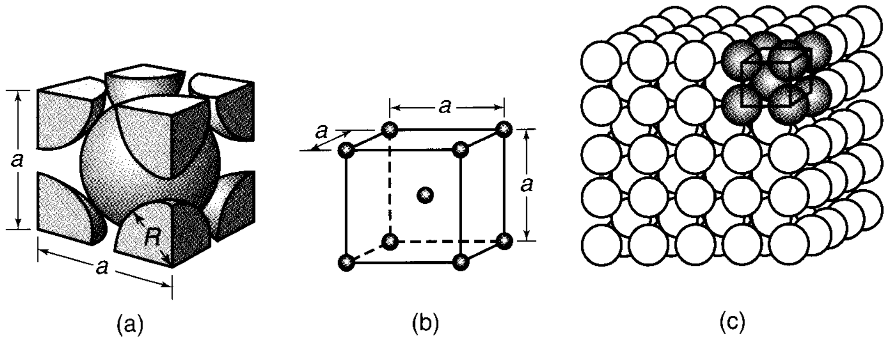
1.3 The Crystal Structure of Metals

When metals solidify from a molten state, the atoms arrange themselves into various orderly configurations, called **crystals**; this atomic arrangement is called **crystal structure** or **crystalline structure**. The smallest group of atoms showing the characteristic **lattice structure** of a particular metal is known as a **unit cell**. It is the building block of a crystal, and a single crystal can have many unit cells.

The following are the three basic atomic arrangements in metals:

1. **Body-centered cubic** (bcc); examples: alpha iron, chromium, molybdenum, tantalum, tungsten, and vanadium.
2. **Face-centered cubic** (fcc); examples: gamma iron, aluminum, copper, nickel, lead, silver, gold, and platinum.
3. **Hexagonal close-packed** (hcp); examples: beryllium, cadmium, cobalt, magnesium, alpha titanium, zinc, and zirconium.

These structures are represented by the illustrations given in Figs. 1.2 through 1.4; 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 in the figure are known as **hard-ball** or **hard-sphere** models; they can be likened to tennis balls arranged in various configurations in a box.



(a) (b) (c)

FIGURE 1.2 The body-centered cubic (bcc) crystal structure: (a) hard-ball model; (b) unit cell; and (c) single crystal with many unit cells.

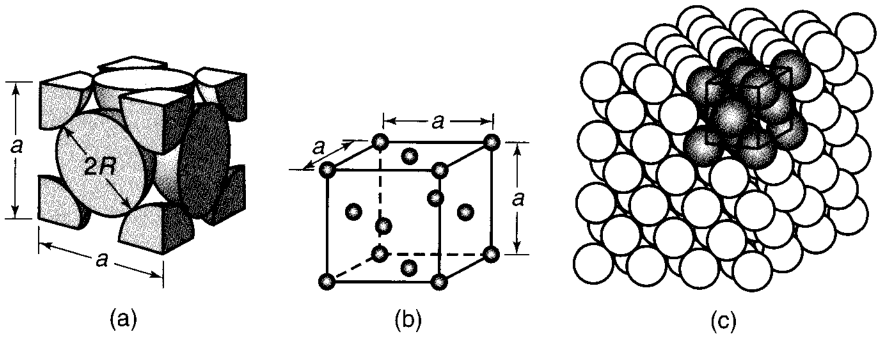


FIGURE 1.3 The face-centered cubic (fcc) crystal structure: (a) hard-ball model; (b) unit cell; and (c) single crystal with many unit cells.

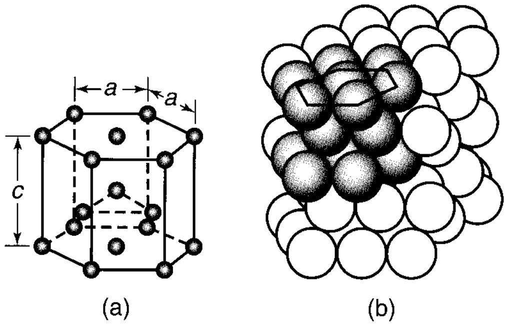


FIGURE 1.4 The hexagonal close-packed (hcp) crystal structure: (a) unit cell; and (b) single crystal with many unit cells.

In the three structures illustrated, 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**, often to improve various properties of the metal. Crystal structure plays a major role in determining the properties of a particular metal.

The reason that metals form different crystal structures is to *minimize* the energy required to fill space. Tungsten, for example, forms a bcc structure because that structure involves less energy than other structures do; likewise, aluminum forms an fcc structure. At different temperatures, however, the same metal may form different structures, because of a lower energy requirement. For example, as described in Chapter 4, iron forms a bcc structure (alpha iron) below 912°C and above 1394°C, but it forms an fcc structure (gamma iron) between 912°C and 1394°C.

The appearance of more than one type of crystal structure is known as **allotropism** or **polymorphism** (meaning "many shapes"). Because the properties and behavior of a metal depend greatly on its crystal structure, allotropism is an important factor in heat treatment of metals, as well as 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 m long and up to 300 mm in diameter, with applications such as turbine blades and semiconductors (see Sections 11.15 and 34.3). However, most metals used in manufacturing 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**; 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 on the crystal structure 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.

There are two basic mechanisms by which plastic deformation takes place in crystal structures. One is the slipping of one plane of atoms over an adjacent plane (called the **slip plane**) under a **shear stress** (Fig. 1.5a). Note that this behavior is much like the 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.

Just as it takes a certain magnitude of force to slide playing cards against each other, a single crystal requires a certain amount of shear stress (called **critical shear stress**) to undergo permanent deformation. Thus, there must be a shear stress of sufficient magnitude within a crystal 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. 1.5a, where *a* is the spacing of the atomic planes and *b* is inversely proportional to the atomic density in the atomic plane. As *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; this property is called **anisotropy**. A simple example of anisotropy is the behavior of woven cloth, which stretches differently when pulled in different directions. Another example is

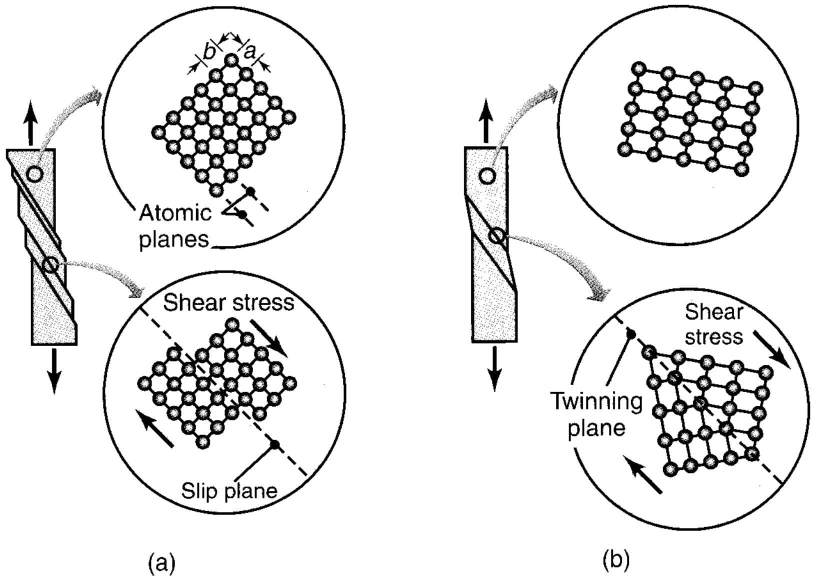


FIGURE 1.5 Permanent deformation of a single crystal under a tensile load. The highlighted grid of atoms emphasizes the motion that occurs within the lattice. (a) Deformation by slip. The *b*/*a* ratio influences the magnitude of the shear stress required to cause slip. (b) Deformation by twinning, involving 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.

the behavior of plywood, which is much stronger in the planar direction than along its thickness direction. Note, for example, how plywood splits easily when a thick nail is driven through its thickness.

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. 1.5b). Twins form abruptly and are the cause of the creaking sound ("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 its direction of slip is known as a *slip system*. In general, metals with 5 or more slip systems are ductile, whereas those with fewer than 5 slip systems are not.

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 crystal, however, the required shear stress is high. Metals with bcc structures 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 is low because of the relatively low *b/a* ratio. These metals generally have moderate strength and good ductility.
3. The **hexagonal close-packed** crystal has 3 slip systems and therefore has a low probability of slip; however, more slip systems become active at elevated temperatures. Metals with hcp structures are generally brittle at room temperature.

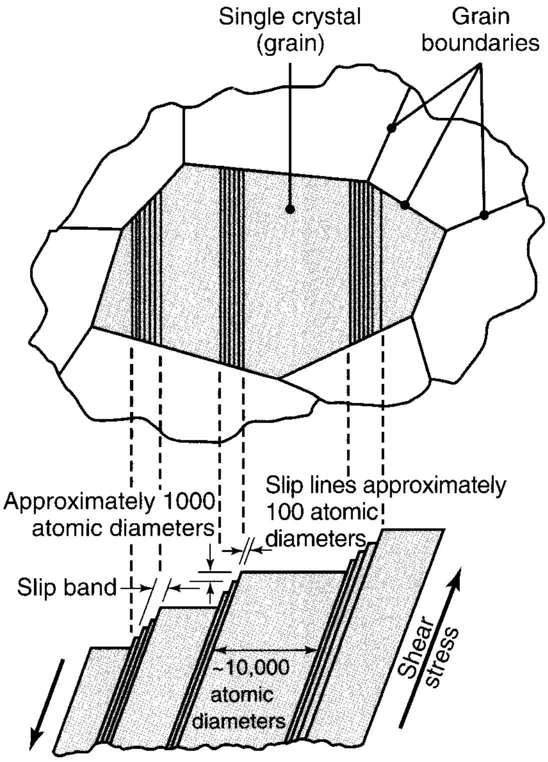


FIGURE 1.6 Schematic illustration of slip lines and slip bands in a single crystal (grain) subjected to a shear stress. A slip band consists of a number of slip planes. The crystal at the center of the upper illustration is an individual grain surrounded by several other grains.

Note in Fig. 1.5a that the portions of the single crystal that have slipped 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 seen from electron microscopy that what appears to be a single slip plane is actually a **slip band** consisting of a number of slip planes (Fig. 1.6).

1.4.1 Imperfections in the Crystal Structure of Metals

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

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. 1.7);
2. *Linear,* or *one-dimensional, defects*, called **dislocations** (Fig. 1.8);
3. *Planar,* or *two-dimensional, imperfections*, such as **grain boundaries** and **phase boundaries** (see Section 1.5);

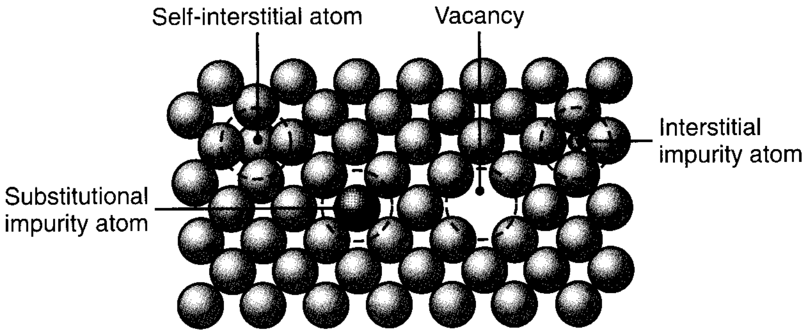


FIGURE 1.7 Schematic illustration of types of defects in a single-crystal lattice: self-interstitial, vacancy, interstitial, and substitutional.

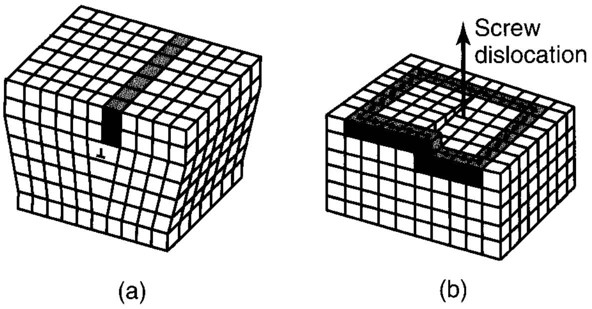


FIGURE 1.8 Types of dislocations in a single crystal: (a) edge dislocation; and (b) screw dislocation.

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

Mechanical and electrical properties of metals, such as yield stress, fracture strength, and electrical conductivity, are adversely affected by 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 (e.g., modulus of elasticity and modulus of rigidity), are not sensitive to these defects; these properties are known as **structure insensitive**.

Dislocations. First observed in the 1930s, *dislocations* are defects in the orderly arrangement of a metal's atomic structure. Because a slip plane containing a dislocation (Fig. 1.9) requires less 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 theoretical strengths of metals.

There are two types of dislocations: **edge** and **screw** (Fig. 1.8). 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 the tail and moves toward the head. Another analogy is the movement of a large carpet on the floor by first forming a hump at one end and moving the hump gradually to the other end. 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.

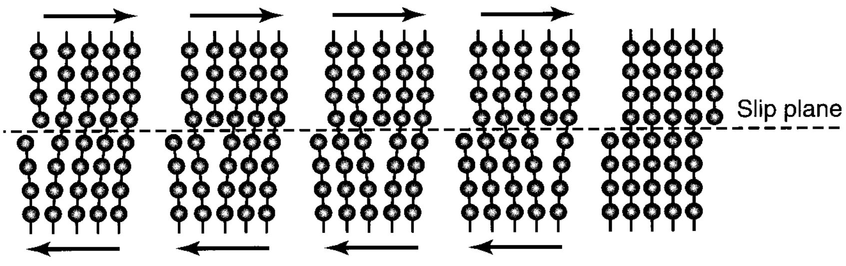


FIGURE 1.9 Movement of an edge dislocation across the crystal lattice under a shear stress. Dislocations help explain why the actual strength of metals is much lower than that predicted by theory.

1.4.2 Work Hardening (Strain Hardening)

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

1. Become entangled and interfere with each other, and
2. Be impeded by barriers, such as grain boundaries, impurities, and inclusions in the material.

The increased shear stress required to overcome entanglements and impediments 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 the metal's strength. Work hardening is used extensively for strengthening in metalworking processes at ambient temperatures. Typical examples are producing sheet metal for automobile bodies and aircraft fuselages by cold rolling (Chapter 13), producing the head of a bolt by forging (Chapter 14), and strengthening wire by reducing its cross section by drawing it through a die (Chapter 15).

1.5 Grains and Grain Boundaries

When a mass of molten metal begins to solidify, crystals begin to form independently of each other at various locations within the liquid mass; they have random and unrelated orientations (Fig. 1.10). Each of these crystals then 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 developed depends on the number of different sites at which individual crystals begin to form (note that there are seven in Fig. 1.10a) and 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 rate of growth of the crystals is high (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.

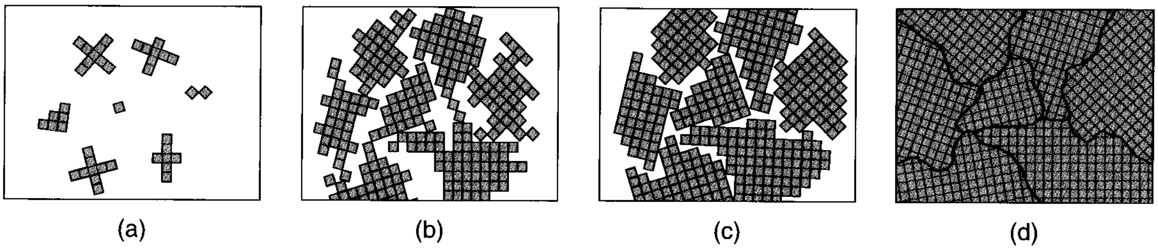


FIGURE 1.10 Schematic illustration of the stages during the solidification of molten metal; each small square represents a unit cell. (a) Nucleation of crystals at random sites in the molten metal; note that the crystallographic orientation of each site is different. (b) and (c) Growth of crystals as solidification continues. (d) Solidified metal, showing individual grains and grain boundaries; note the different angles at which neighboring grains meet each other.

Note in Fig. 1.10d that the growing grains eventually interfere with and impinge upon one another. The surfaces that separate these 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.) Thus, because its many grains have random crystallographic orientations, the behavior of a piece of polycrystalline metal is essentially **isotropic**; that is, its properties do not vary with the direction of testing.

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).

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

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

$$N = 2^{n-1}. \tag{1.1}$$

Because grains are typically extremely small, many grains can occupy a piece of metal (Table 1.1). Grain sizes between 5 and 8 are generally considered fine grained.

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	1,020
5	256	2,900
6	512	8,200
7	1,024	23,000
8	2,048	65,000
9	4,096	185,000
10	8,200	520,000
11	16,400	1,500,000
12	32,800	4,200,000

A grain size of 7 is generally acceptable for sheet metals for making car bodies, appliances, and kitchen utensils (Chapter 16).

EXAMPLE 1.1 Number of Grains in the Ball of a Ballpoint Pen

Assume that the ball of a ballpoint pen is 1 mm in diameter and has an ASTM grain size of 10. Calculate the number of grains in the ball.

Solution A metal with an ASTM grain size of 10 has 520,000 grains per mm^3. (See Table 1.1.) The volume of the 1-mm-diameter ball is

$$V = \frac{4\pi r^3}{3} = \frac{4\pi(0.5 \text{ mm})^3}{3} = 0.5236 \text{ mm}^3.$$

The total number of grains is calculated by multiplying the volume by the grains per mm^3, or

$$\text{No. grains} = (0.5236 \text{ mm}^3)(520{,}000 \text{ grains/mm}^3) = 272{,}300$$

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 the movement of dislocations, grain boundaries also influence strain hardening. These effects depend on temperature, deformation rate, and the type and amount of impurities present *along* the grain boundaries.

Because the atoms along the grain boundaries are packed less efficiently and are more disordered, 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, and thus they can be more easily removed or chemically bonded to another atom. For example, a metal surface becomes rougher when etched or subjected to corrosive environments. (See also *end grains in forging*, in Section 14.5).

At elevated temperatures, 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) results from grain-boundary sliding (see Section 2.8).

Grain-boundary embrittlement. When exposed to certain low-melting-point metals, a normally ductile and strong metal can crack when subjected to very low external stresses. Examples of such behavior are (a) aluminum wetted with a mercury–zinc amalgam or liquid gallium, and (b) copper at elevated temperature wetted with lead or bismuth. These added 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, a phenomenon known as **solid-metal embrittlement.**

Hot shortness is caused by local melting of a constituent or of an impurity in the grain boundary at a temperature below the melting point of the metal itself. When subjected to plastic deformation at elevated temperatures (*hot working*), a piece of metal crumbles along its grain boundaries; examples are antimony in copper, leaded steels (Section 21.7.1), and leaded brass. To avoid hot shortness, the metal is usually worked at a lower temperature in order to prevent softening and melting along the grain boundaries. Another form of embrittlement is **temper embrittlement** in alloy steels, which is 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 (a process known as *cold working*), the grains become deformed and elongated, as shown schematically in Fig. 1.11. Deformation may be carried out, for example, by compressing the metal piece, 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 to make an automobile body (Chapter 16). 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, because of the 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 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.

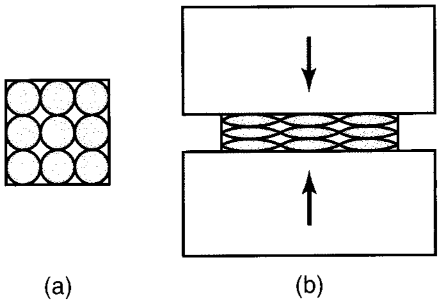


FIGURE 1.11 Plastic deformation of idealized (equiaxed) grains in a specimen subjected to compression (such as occurs in the forging or rolling of metals): (a) before deformation; and (b) after deformation. Note the alignment of grain boundaries along a horizontal direction; this effect is known as *preferred orientation*.

Anisotropy (Texture). Note in Fig. 1.11b that, as a result of plastic deformation, the grains have elongated in one direction and contracted in the other. Consequently, this piece of 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 takes places and on how uniformly the metal is deformed. Note from the direction of the crack in Fig. 1.12, for example, that the ductility of the cold-rolled sheet in the transverse direction is lower than that 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 reduces magnetic-hysteresis losses and thus improves the efficiency of transformers. (See also *amorphous alloys*, Section 6.14.) There are two general types of anisotropy in metals: *preferred orientation* and *mechanical fibering*.

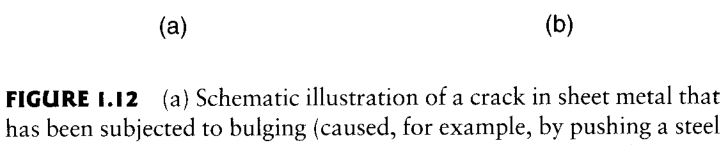


FIGURE 1.12 (a) Schematic illustration of a crack in sheet metal that has been subjected to bulging (caused, for example, by pushing a steel ball against the sheet). Note the orientation of the crack with respect to the rolling direction of the sheet; this sheet is anisotropic. (b) Aluminum sheet with a crack (vertical dark line at the center) developed in a bulge test; the rolling direction of the sheet was vertical. *Courtesy*: J.S. Kallend, Illinois Institute of Technology.

Preferred Orientation. Also called crystallographic anisotropy, *preferred orientation* can be best described by referring to Fig. 1.5a. When a single-crystal piece of 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. By contrast, slip planes under compression tend to align themselves in a direction perpendicular to the direction of the compressive force.

Mechanical Fibering. This is a type of anistropy that results from the alignment of inclusions (*stringers*), impurities, and voids in the metal during deformation. Note that if the spherical grains in Fig. 1.11a were coated with impurities, these impurities would align themselves in a generally horizontal direction after deformation. Because impurities weaken the grain boundaries, this piece of metal will now be weaker and less ductile when tested in the vertical direction. As an analogy, consider plywood, which is strong in tension along its planar direction, but peels off (splits) easily when pulled in tension in its thickness direction.

1.7 Recovery, Recrystallization, and Grain Growth

We have seen that plastic deformation at room temperature causes distortion of the grains and grain boundaries (leading to anisotropic behavior), a general increase in strength, and a decrease in ductility. These effects can be reversed, and the properties of the metal can be 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 the heating 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 piece are relieved. Subgrain boundaries begin to form (a process called **polygonization**), with no significant change in mechanical properties such as hardness and strength (Fig. 1.13).
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. 1.13). Lead, tin, cadmium, and zinc recrystallize at about room temperature; consequently, they do not work harden when cold worked.

   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 (**stored energy**) also increase. This energy supplies some of the work required for recrystallization.

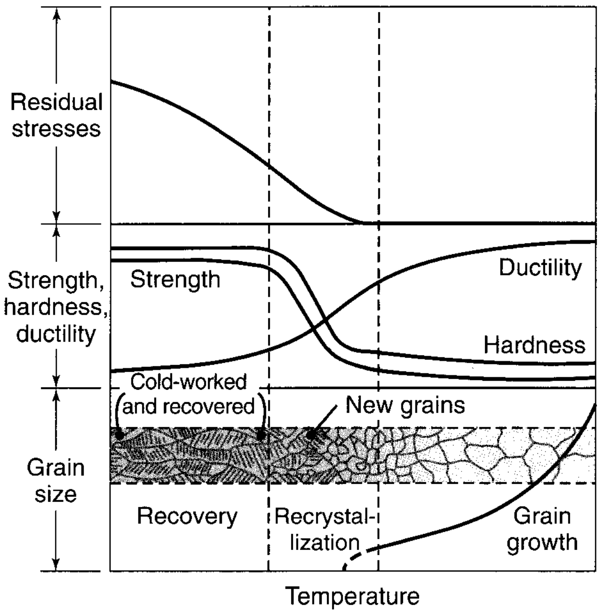


FIGURE 1.13 Schematic illustration of the effects of recovery, recrystallization, and grain growth on mechanical properties and on the shape and size of grains. Note the formation of small new grains during recrystallization. *Source*: After G. Sachs.

Recrystallization is also a function of time, because it involves **diffusion**—the movement and exchange of atoms across grain boundaries.

The effects on recrystallization of temperature, time, and plastic deformation by cold working are as follows:

- **a.** For a constant amount of deformation by cold working, the time required for recrystallization decreases with increasing temperature;
- **b.** The more 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 of converting a coarse-grained structure to one having a finer grain, and thus one 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.

3. **Grain growth.** If the temperature of the metal is raised further, the grains begin to grow, and their size may eventually exceed the original grain size; called *grain growth*, this phenomenon adversely affects mechanical properties (Fig. 1.13). 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 that is usually, but not necessarily, 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 a dimensionless quantity, 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 to 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 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.
- Grain boundaries have a major influence on the behavior of metals. Boundaries can undergo embrittlement, severely reducing ductility at elevated temperatures (hot shortness); they are also responsible for the creep phenomenon, which is due to grain boundary sliding.
- Metals may be plastically deformed at room, warm, or high temperatures, their behavior and workability depending 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 (preferred orientation or mechanical fibering), whereby the properties are different in different directions.
- The effects of cold working can be reversed by annealing the metal: heating it in a certain temperature range for a given period of time, thereby allowing the successive processes 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
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., 2005; Vol. 2, *An Introduction to Microstructures, Processing and Design.* Butterworth-Heinemann, 2005.

Ashby, M., Shercliff, H., and Cebon, D., **Materials: Engineering, Science, Processing and Design,** Butterworth-Heinemann, 2007.

Askeland, D.R., and Fulay, P.P., **The Science and Engineering of Materials,** 5th ed, Cengage-Engineering, 2005.

Callister, W.D., Jr., **Materials Science and Engineering: An Introduction,** 7th ed., Wiley, 2006.

Callister, W.D., Jr., and Rethwisch, D.G., **Fundamentals of Materials Science and Engineering: An Integrated Approach,** 3rd ed., Wiley, 2007.

Shackelford, J.F., **Introduction to Materials Science for Engineers,** 6th ed., Prentice Hall, 2004.

REVIEW QUESTIONS

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

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

1.3. Define anisotropy. What is its significance?

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

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

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

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

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

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

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

1.11. Explain the difference between recovery and recrystallization.

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

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

1.14. Describe what the orange peel effect is.

1.15. Some metals, such as lead, do not become stronger when cold worked. Explain the reason.

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

1.17. Differentiate between stress relaxation and stress relieving.

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

QUALITATIVE PROBLEMS

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

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

1.21. 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.

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

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

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

1.25. 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.

1.26. 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.

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

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

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

1.30. It was stated in this chapter that twinning usually occurs in hcp materials, but Fig. 1.5b shows twinning in a rectangular array of atoms. Can you explain the discrepancy?

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

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

1.33. Comment on your observations regarding Fig. 1.13.

QUANTITATIVE PROBLEMS

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

1.35. The atomic weight of copper is 63.55, meaning that 6.023×10^{23} atoms weigh 63.55 grams. The density of copper is 8970 kg/m^3, and pure copper forms fcc crystals. Estimate the diameter of a copper atom.

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

1.37. A strip of metal is reduced from 40 mm in thickness to 20 mm by cold working; a similar strip is reduced from 40 mm to 30 mm. Which one of these cold-worked strips will recrystallize at a lower temperature? Why?

▶**1.38.** A paper clip is made of wire that is 120 mm long and 0.75 mm in diameter. If the ASTM grain size is 9, how many grains are there in the paper clip?

▶**1.39.** How many grains are on the surface of the head of a pin? Assume that the head of a pin is spherical with a 1-mm diameter and has an ASTM grain size of 2.

▶**1.40.** The unit cells shown in Figs. 1.2 through 1.4 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.

▶**1.41.** Show that the lattice constant a in Fig. 1.3a 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.

▶**1.42.** 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.

1.43. 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 175×, instead of the 100× that is required by the ASTM standards. Determine the correct grain size.

1.44. If the diameter of the aluminum atom is 0.5 nm, how many atoms are there in a grain of ASTM grain size 8?

SYNTHESIS, DESIGN AND PROJECTS

1.45. 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.

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

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

1.48. 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. 1.2–1.4. Comment on your observations.

1.49. Take a deck of playing cards, place a rubber band around it, and then slip the cards against each other to represent Figs. 1.5a and 1.6. 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?

1.50. 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.

1.51. 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.

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 various mechanical properties.
- The tension test is described first. This test is commonly used for quantifying a number of material parameters, including elastic modulus, yield stress, ultimate strength, ductility, and toughness.
- Compression tests are then described. These tests are useful because they more closely simulate manufacturing processes; nonetheless, they have the unavoidable drawback of contributing friction to the testing.
- 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 applicability are then explored.
- Fatigue involves the failure of material due to cyclic or repeating loads, whereas creep is deformation due to the application of a constant load over an extended period.
- The chapter ends with descriptions of the types of, and factors involved in, failure and fracture of materials.

2.1 Introduction

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

As indicated in Table I.2, a wide variety of metallic and nonmetallic materials is now available, with an equally wide range of properties, as shown qualitatively in Table 2.1. 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 properties.

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	

2.2 Tension

The **tension test** is the most common method for determining the *mechanical properties* of materials, such as strength, ductility, toughness, elastic modulus, and strain-hardening capability. The tension test first requires the preparation of a **test specimen**, as shown in Fig. 2.1a. Although most tension-test specimens are solid and round, they can also be flat or tubular. The specimen is prepared generally according

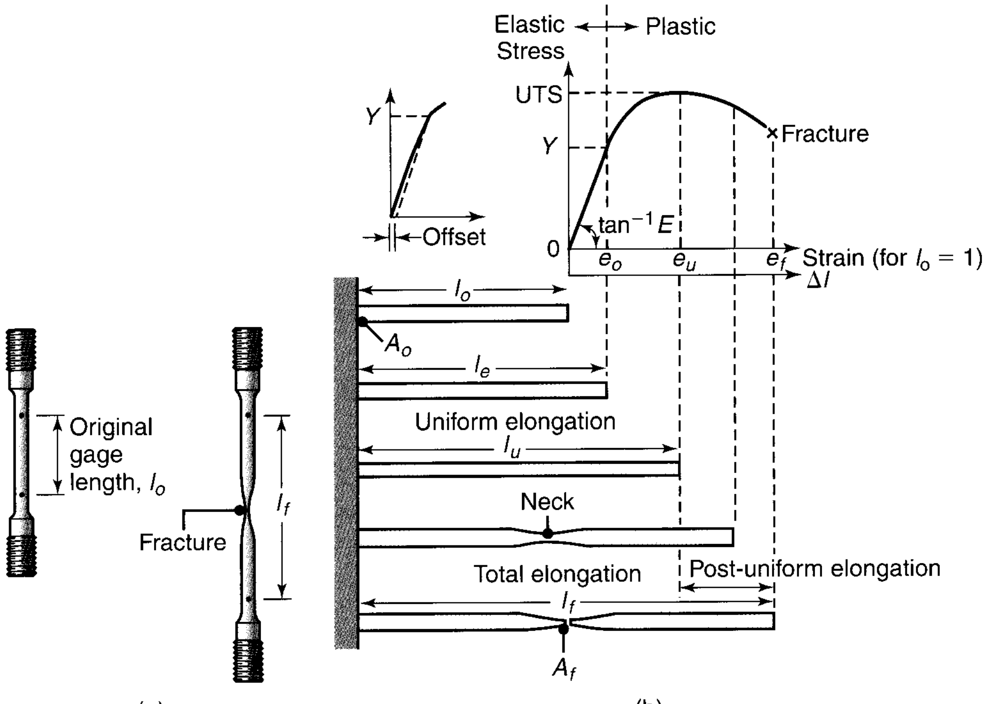


FIGURE 2.1 (a) A standard tensile-test specimen before and after pulling, showing original and final gage lengths. (b) Stages in specimen behavior in a tension test.

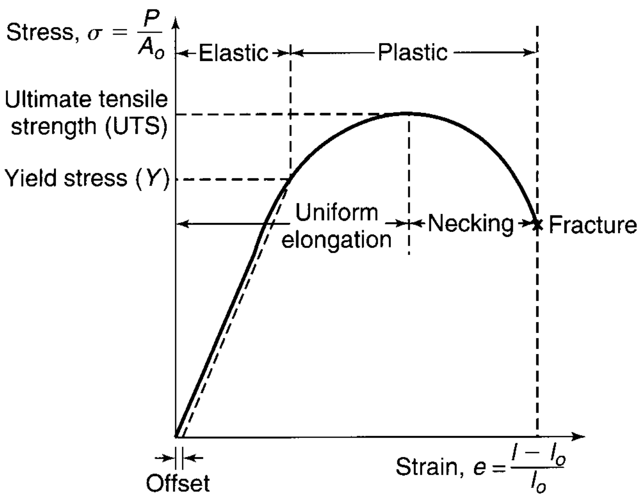


FIGURE 2.2 A typical stress–strain curve obtained from a tension test, showing various features.

to ASTM specifications; 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. It is mounted in the jaws of a tension-testing machine equipped with various accessories and controls so that the specimen can be tested at different temperatures and rates of deformation.

2.2.1 Stress–Strain Curves

A typical deformation sequence in a tension test is shown in Fig. 2.1b. When the load is first applied, the specimen elongates in proportion to the load, called **linear elastic** behavior (Fig. 2.2). If 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** (**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

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

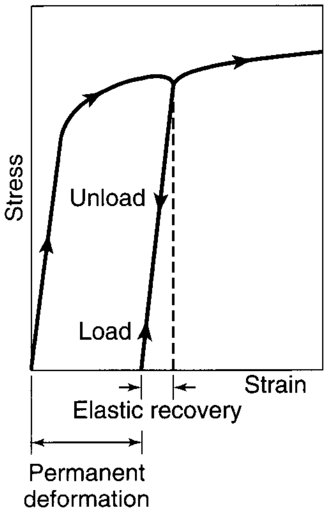


FIGURE 2.3 Schematic illustration of the loading and the unloading of a tensile-test specimen. Note that, during unloading, the curve follows a path parallel to the original elastic slope.

where l is the instantaneous length of the specimen.

As the load is increased, the specimen begins to undergo *nonlinear* elastic deformation at a stress 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,** Y, of the material is reached. The yield stress and other properties of various metallic and nonmetallic materials are given in Table 2.2.

For soft and ductile materials, it may not be easy 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. Therefore, Y is usually defined by drawing a line with the same slope as the linear elastic curve, but that is **offset** by a strain of 0.002, or 0.2% elongation. The yield stress is then defined as the stress where this 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** throughout 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 (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** (**UTS**), of the material. Values for UTS for various materials are given in Table 2.2.

TABLE 2.2

Mechanical Properties of Various Materials at Room Temperature

Metals (wrought)	E (GPa)	Y (MPa)	UTS (MPa)	Elongation in 50 mm (%)	Poisson's ratio, ν
Aluminum and its alloys	69–79	35–550	90–600	45–4	0.31–0.34
Copper and its alloys	105–150	76–1100	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, Y, and UTS and the highest values for elongation are for pure metals. Multiply gigapascals (GPa) by 145,000 to obtain pounds per square in. (psi), megapascals (MPa) by 145 to obtain psi.

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 finally fractures at the necked region (Fig. 2.1a); the engineering stress at fracture is known as the **breaking** or **fracture stress.**

The ratio of stress to strain in the elastic region is 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).

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 hence the **stiffness** of the material. The higher the E value, 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 sheet when they are loaded.

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.

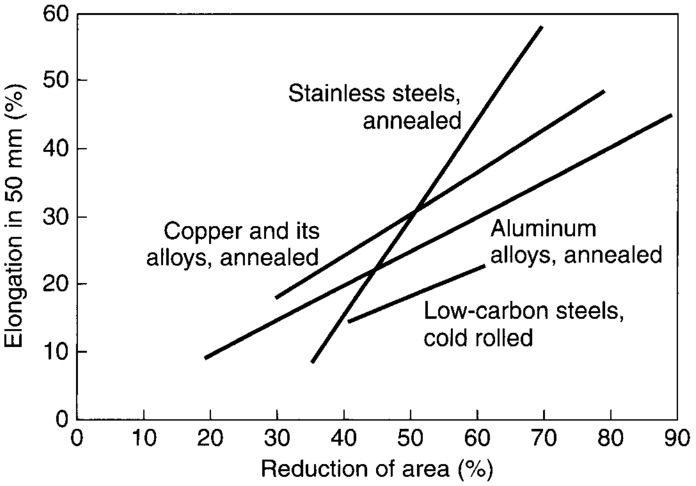


FIGURE 2.4 Approximate relationship between elongation and tensile reduction of area for various groups of metals.

2.2.2 Ductility

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

$$\text{Elongation} = \frac{(l_f - l_o)}{l_o} \times 100, \qquad (2.4)$$

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

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

where A_o and A_f are, respectively, the original and final (fracture) cross-sectional areas of the test specimen. Reduction of area and elongation are generally interrelated, as shown in Fig. 2.4 for some typical metals. Thus, the ductility of a piece of chalk is zero, because it does not stretch at all or reduce in cross section; by contrast, a ductile specimen, such as putty or chewing gum, stretches and necks considerably before it fails.

2.2.3 True Stress and True Strain

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; thus, engineering stress does not represent the *actual* stress to which the specimen is subjected.

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

$$\sigma = \frac{P}{A}. \qquad (2.6)$$

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

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

Note from Eqs. (2.2) and (2.7) that, for small values of strain, the engineering and true strains are approximately 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 are consistent with actual physical phenomena in the deformation of materials. Let's assume, for example, a hypothetical situation: A specimen 50 mm in height is compressed between flat platens to a final height of zero; in other words, we have deformed the specimen infinitely. 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. 2.5a; also, Fig. 2.2), and then to divide the load (vertical axis) by the original cross-sectional area, A_o, and 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. 2.5b) has the same shape as the load–elongation curve shown in Fig. 2.5a. (In this example, $A_o = 36.1\ \text{mm}^2$ and $A_f = 10.3\ \text{mm}^2$.)

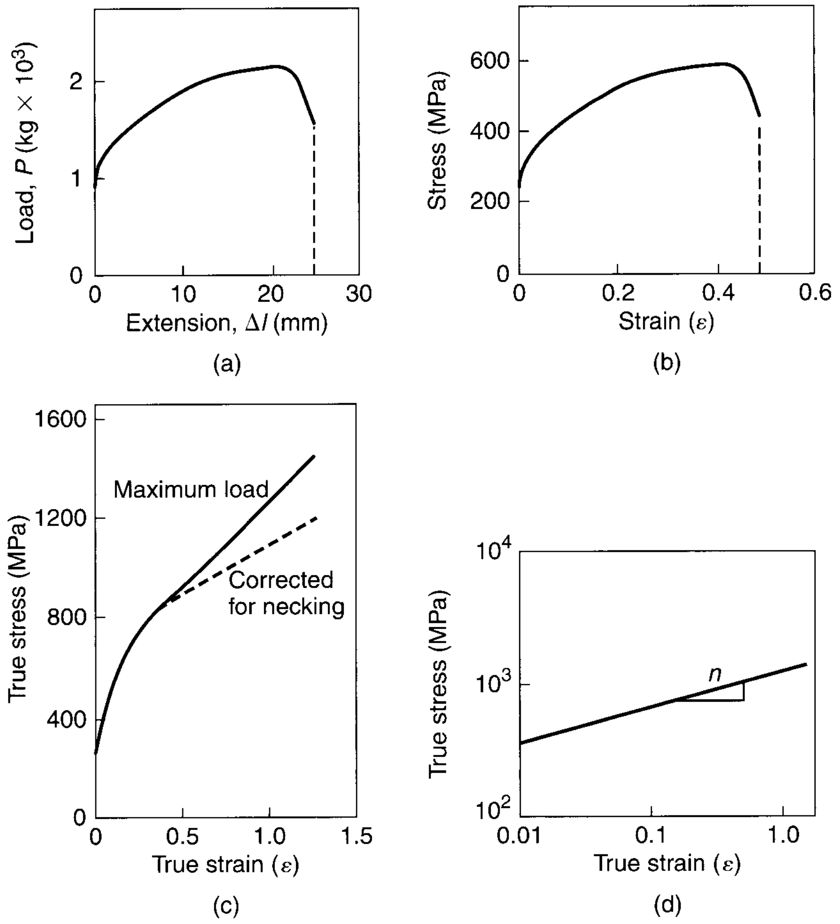


FIGURE 2.5 (a) Load–elongation curve in tension testing of a stainless steel specimen. (b) Engineering stress–engineering strain curve, drawn from the data in Fig. 2.5a. (c) True stress–true strain curve, drawn from the data in Fig. 2.5b. Note that this curve has a positive slope, indicating that the 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. 2.5c. The correction is due to the triaxial state of stress that exists in the necked region of the specimen.

TABLE 2.3

Typical Values for *K* and *n* for Metals at Room Temperature

	K (MPa)	n
Aluminum		
1100–O	180	0.20
2024–T4	690	0.16
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	2,070	0.50
Copper, annealed	315	0.54
Steel		
Low-C, annealed	530	0.26
4135, annealed	1,015	0.17
4135, cold-rolled	1,100	0.14
4340, annealed	640	0.15
304 stainless, annealed	1,275	0.45
410 stainless, annealed	960	0.10

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. 2.5c. 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 gives higher stress values than the actual true stress; hence, to compensate, the curve must be corrected downward.

The true stress–true strain curve in Fig. 2.5c 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. 2.5c is plotted on a log–log graph, it is found that the curve is approximately a straight line (Fig. 2.5d). The slope of the curve is equal to the exponent n. Thus, the higher the slope, the greater is the strain-hardening capacity of the material—that is, the stronger and harder it becomes as it is strained.

True stress–true strain curves for a variety of metals are given in Fig. 2.6. When they are reviewed in detail, some differences between Table 2.3 and Fig. 2.6 will be noted; these discrepancies result from the fact that different sources of data and different specimens are involved. Note that the elastic regions have been deleted, because the slope in this region is very high. As a result, the point of intersection of each curve with the vertical axis in this figure can be considered to be the yield stress, 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 material deformed and indicates the work required to plastically deform a 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.

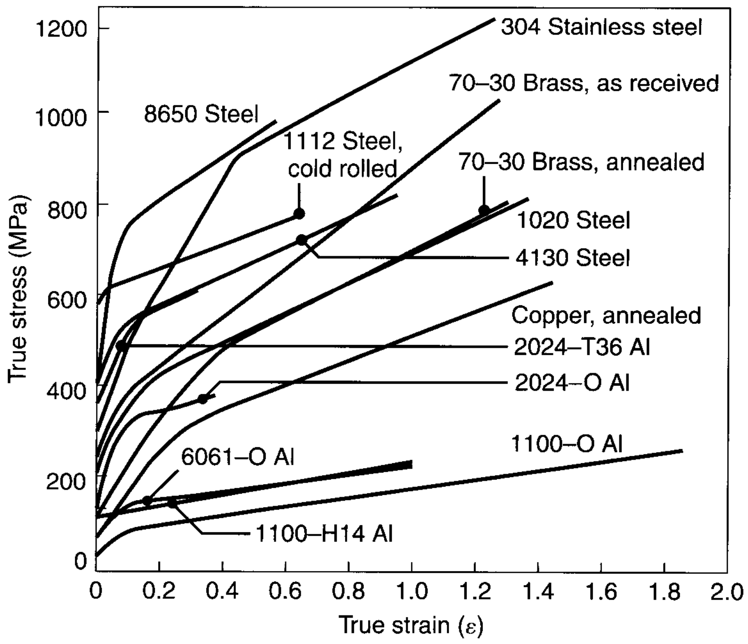


FIGURE 2.6 True stress–true strain curves in tension at room temperature for various metals. The curves start at a finite level of stress: The elastic regions have too steep a slope to be shown in this figure; thus, each curve starts at the yield stress, Y, of the material.

2.2.5 Strain at Necking in a Tension Test

As noted 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 because the cross-sectional 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 piece of material can experience before it begins to neck. This observation is important, particularly in regard to sheet-metal-forming operations that involve the stretching of the workpiece material (Chapter 16). It can be seen in Table 2.3 that annealed copper, brass, and stainless steel have high n values; this means that they can be stretched uniformly to a greater extent than can the other metals listed.

EXAMPLE 2.1 Calculation of Ultimate Tensile Strength

This example shows that the UTS of a material can be calculated from its K and n values. Assume that a material has a true stress–true strain curve given by

$$\sigma = 690^{0.5}\ \text{psi}.$$

Calculate the true ultimate tensile strength and the engineering UTS of this material.

Solution Because the necking strain corresponds to the maximum load, the necking strain for this material is

$$\epsilon = n = 0.5,$$

the *true* ultimate tensile strength is

$$\sigma = Kn^n = 690(0.5)^{0.5} = 488\ \text{MPa}.$$

The true area at the onset of necking is obtained from

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

Thus,

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

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,

$$P = (488)(0.606)(A_o) = 2900A_o \text{ kg}.$$

Since UTS $= P/A_o$.

$$UTS = 296 \text{ MPa}.$$

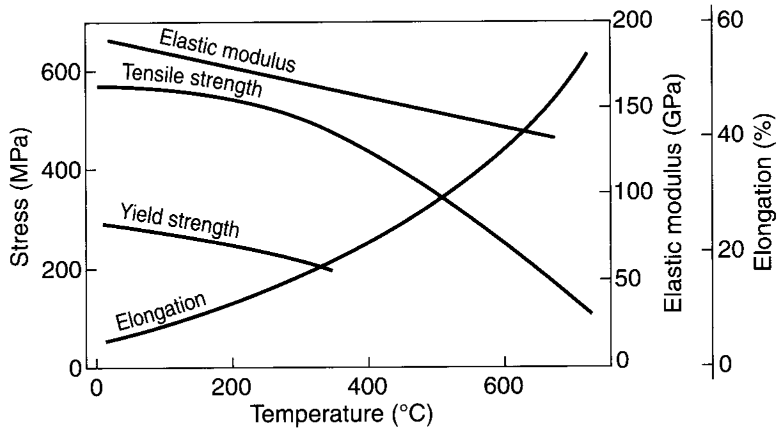


FIGURE 2.7 Effect of temperature on mechanical properties of a carbon steel. Most materials display similar temperature sensitivity for elastic modulus, yield strength, ultimate strength, and ductility.

2.2.6 Temperature Effects

Increasing the temperature generally has the following effects on stress–strain curves (Fig. 2.7):

a. The ductility and toughness incease, and

b. The yield stress and the modulus of elasticity decrease.

Temperature also affects the strain-hardening exponent of most metals, in that n decreases with increasing temperature. The influence of temperature is, however, best described in conjunction with the rate of deformation.

2.2.7 Rate-of-deformation Effects

Just as we can blow up a balloon or stretch a rubber band at different rates, we can shape a piece of material in a manufacturing process at different speeds. Some machines, such as hydraulic presses, form materials at low speeds; others, such as mechanical presses, form materials at high speeds. To incorporate such effects, it is common practice to strain a specimen at a rate corresponding to that which will be experienced in the actual manufacturing process.

The **deformation rate** is defined as the speed at which a tension test is being carried out, in units of, say, m/s. The **strain rate**, on the other hand, is a function of

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}

the specimen's length. A short specimen elongates proportionately more during the same period than does a long specimen. For example, let's take two rubber bands, one 20 mm and the other 100 mm long, respectively, and elongate them both by 10 mm within a period of 1 second. The engineering strain in the shorter specimen is $\frac{10}{20} = 0.5$; that in the longer is $\frac{10}{100} = 0.1$. Thus, the strain rates are $0.5\ \text{s}^{-1}$ and $0.1\ \text{s}^{-1}$, respectively, with the short band being subjected to a strain rate five times higher than that for the long band, although they are both being stretched at the same deformation rate.

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 range encountered in practice, strain rates are usually stated in terms of orders of magnitude, such as $10^2\ \text{s}^{-1}$, $10^4\ \text{s}^{-1}$, and so on.

The typical effects that temperature and strain rate jointly have on the strength of metals are shown in Fig. 2.8. Note that increasing the strain rate increases the strength of the material (**strain-rate hardening**). The slope of these curves is called the **strain-rate sensitivity exponent**, m. The value of m is obtained from log–log plots, provided that the vertical and horizontal scales are the same (unlike those in Fig. 2.8). A slope of 45° would therefore indicate a value of $m = 1$. The relationship is given by the equation

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

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. 2.8, it can be seen that the sensitivity of strength to strain rate increases with temperature; in other words, m increases with increasing temperature. Also, the slope is relatively flat at room temperature; that is, m is very low. This condition is true for most metals, but not for those that recrystallize at room

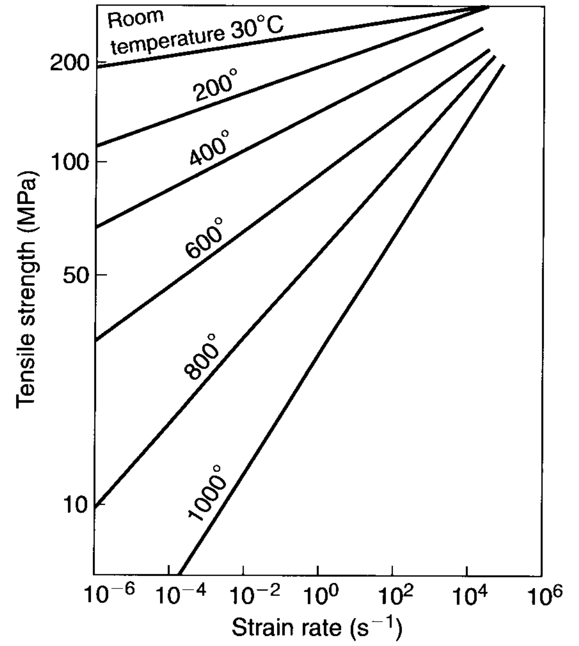


FIGURE 2.8 The effect of strain rate on the ultimate tensile strength for aluminum. Note that, as the temperature increases, the slopes of the curves increase; thus, strength becomes more and more sensitive to strain rate as temperature increases. *Source*: J.H. Hollomon.

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 superplastic materials (see below).

The magnitude of the strain-rate sensitivity exponent significantly influences necking in a tension test. With increasing m, the material stretches farther 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 (Section 16.10).

Superplasticity. The term *superplasticity* refers to the capability of some materials to undergo large uniform elongation prior to necking and fracture in tension. The elongation ranges from a few hundred percent to as much as 2000%. Common nonmetallic materials exhibiting superplastic behavior are bubble gum and glass (at elevated temperatures) and thermoplastics. As a result, glass and thermoplastics can successfully be formed into complex shapes such as beverage bottles and lighted advertising signs. 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 many times their original length.

2.2.8 Hydrostatic Pressure Effects

Various tests have been performed to determine the effect of hydrostatic pressure on mechanical properties of materials. Test results 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. This beneficial effect of hydrostatic pressure has been exploited in metalworking processes, particularly 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 many metals and alloys in nuclear applications, extensive studies have been conducted on the effects of radiation on mechanical properties. Typical changes in the properties of steels and other metals exposed to high-energy radiation are increased yield stress, tensile strength, and hardness, and decreased ductility and toughness. Radiation has similar detrimental effects on the behavior of plastics.

2.3 Compression

Many operations in manufacturing, particularly processes such as forging, rolling, and extrusion (Part III), are performed with the workpiece subjected to compressive stresses. The **compression test**, in which the specimen is subjected to a compressive load, gives information that is useful for estimating forces and power requirements in these processes. This test is usually carried out by compressing a solid cylindrical specimen between two well-lubricated flat dies (platens). Because of friction between the specimen and the platens, the specimen's cylindrical surface bulges, an effect is called **barreling** (see Fig. 14.3). Note that slender specimens can *buckle* during this test; thus, the height-to-diameter ratio of the solid cylindrical specimen is typically less than 3:1. (See also Section 14.4 on *heading*.)

Because of barreling, the cross-sectional area of the specimen changes along its height, and obtaining the stress–strain curves in compression can be difficult.

Furthermore, friction dissipates energy, so the compressive force is higher than it otherwise would be in order to supply the work required to overcome friction. With effective lubrication, friction can be minimized, and 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, it can be seen that the true stress–true strain curves coincide. This behavior does not hold true for *brittle* materials, which are generally stronger and more ductile in compression than in tension. (See Table 8.1.)

When a metal is subjected to tension into the plastic range, and then the load is released and a compressive load is applied, the yield stress in compression is found to be lower than that in tension. This phenomenon is known as the **Bauschinger effect** (after J. Bauschinger, who reported it in 1881), and it is exhibited in varying degrees by all metals and alloys. Because of the lowered yield stress in the direction opposite that of the original load application, the phenomenon is also called **strain softening** or **work softening.**

Disk Test. For brittle materials such as ceramics and glasses (Chapter 8), a **disk test** has been developed, in which the disk is subjected to compression between two hardened flat platens (Fig. 2.9). When the material is loaded as shown, tensile stresses develop perpendicular to the vertical centerline along the disk; fracture begins and the disk splits in half vertically. The *tensile stress*, σ, in the disk is uniform along the centerline and can be calculated from the formula

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

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 platens. These strips also protect the platens from being damaged during the test. The fracture at the center of the specimen has been utilized in the manufacture of *seamless tubing* (Section 13.6).

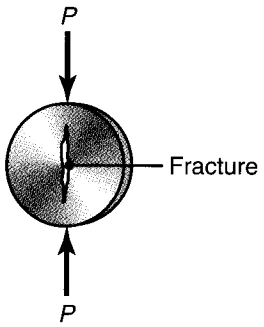


FIGURE 2.9 Disk test on a brittle material, showing the direction of loading and the fracture path.

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 metal cutting (Section 21.2). The test method generally used to determine properties of materials in shear is the **torsion test.** In order to obtain an approximately uniform stress and strain distribution along the cross section, this test is usually performed on a thin tubular specimen.

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

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

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

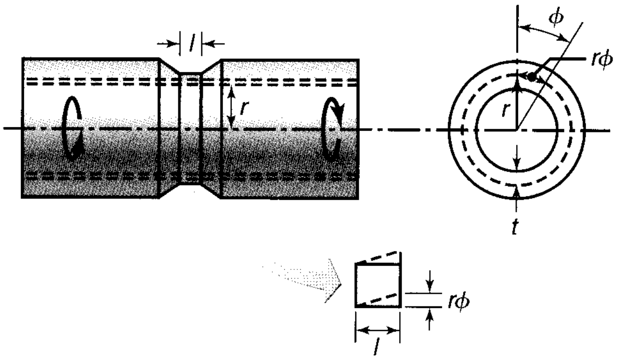


FIGURE 2.10 A typical torsion-test specimen; it is mounted between the two heads of a testing machine and twisted. Note the shear deformation of an element in the reduced section of the specimen.

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

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

where l is the length of tube subjected to torsion 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, a quantity related to the modulus of elasticity, E. The angle of twist, ϕ, to fracture in the torsion of solid round bars at elevated temperatures is also useful in estimating the forgeability of metals. The greater the number of twists prior to failure, the better is the forgeability (Section 14.5).

2.5 Bending (Flexure)

Preparing specimens from brittle materials, such as ceramics and carbides, is difficult because of the problems involved in shaping and machining them to proper dimensions. Furthermore, such specimens are sensitive to surface defects and notches, and clamping brittle test specimens for testing is difficult. Also, improper alignment of the test specimen may result in a nonuniform stress distribution along the cross section.

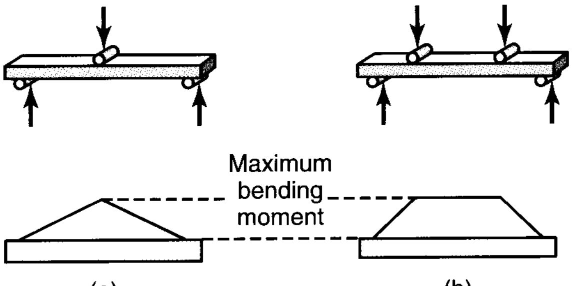


FIGURE 2.11 Two bend-test methods for brittle materials: (a) three-point bending; (b) four-point bending. The areas on the beams represent the bending-moment diagrams, described in texts on the mechanics of solids. Note the region of constant maximum bending moment in (b); by contrast, the maximum bending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

A commonly used test method for brittle materials is the **bend** or **flexure test**, which usually involves a specimen that has a rectangular cross section and is supported, as shown in Fig. 2.11. The load is applied vertically, at either one point or two points; as a result, 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.

These stresses 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 in this volume than in that in Fig. 2.11a. Consequently, the four-point test gives a lower modulus of rupture than the three-point test.

2.6 Hardness

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. **Hardness** is usually defined as *resistance to permanent indentation*; thus, steel is harder than aluminum, and aluminum is harder than lead. Hardness, however, is not a fundamental property, because the resistance to indentation depends on the shape of the indenter and on the load applied.

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. Commonly used hardness tests are described next.

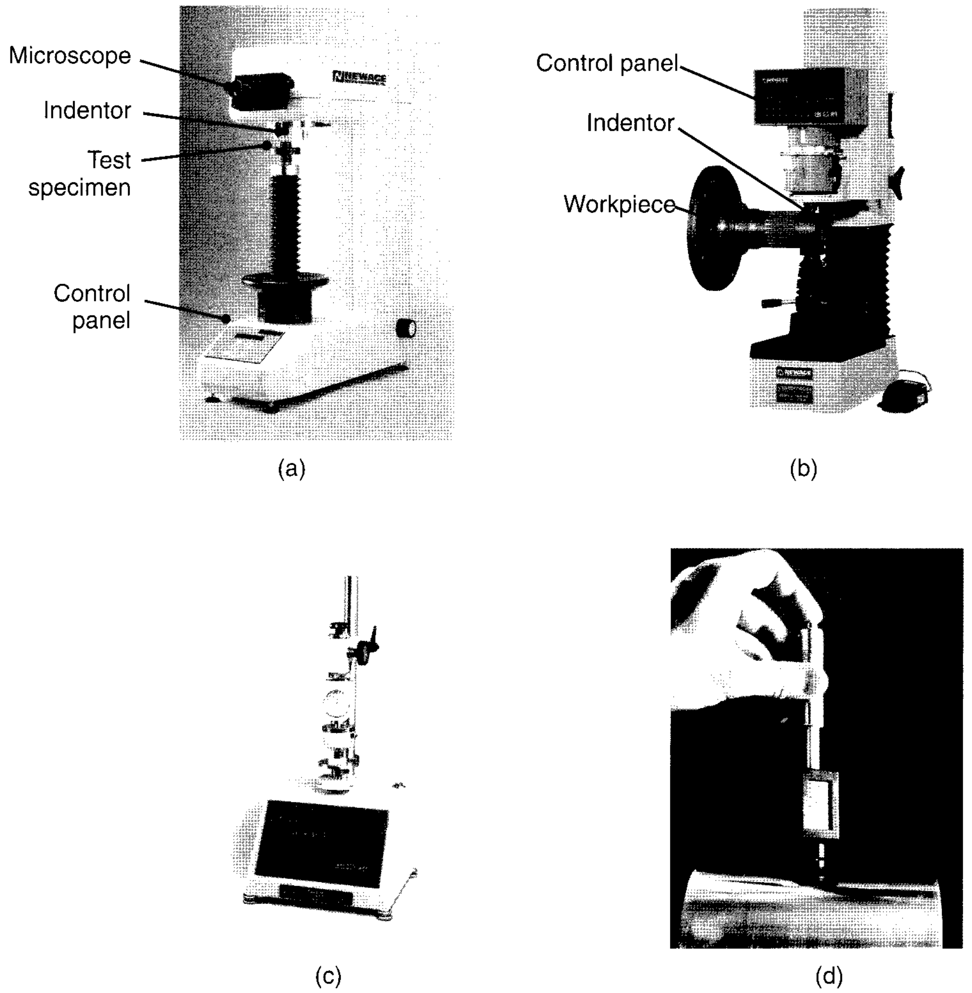
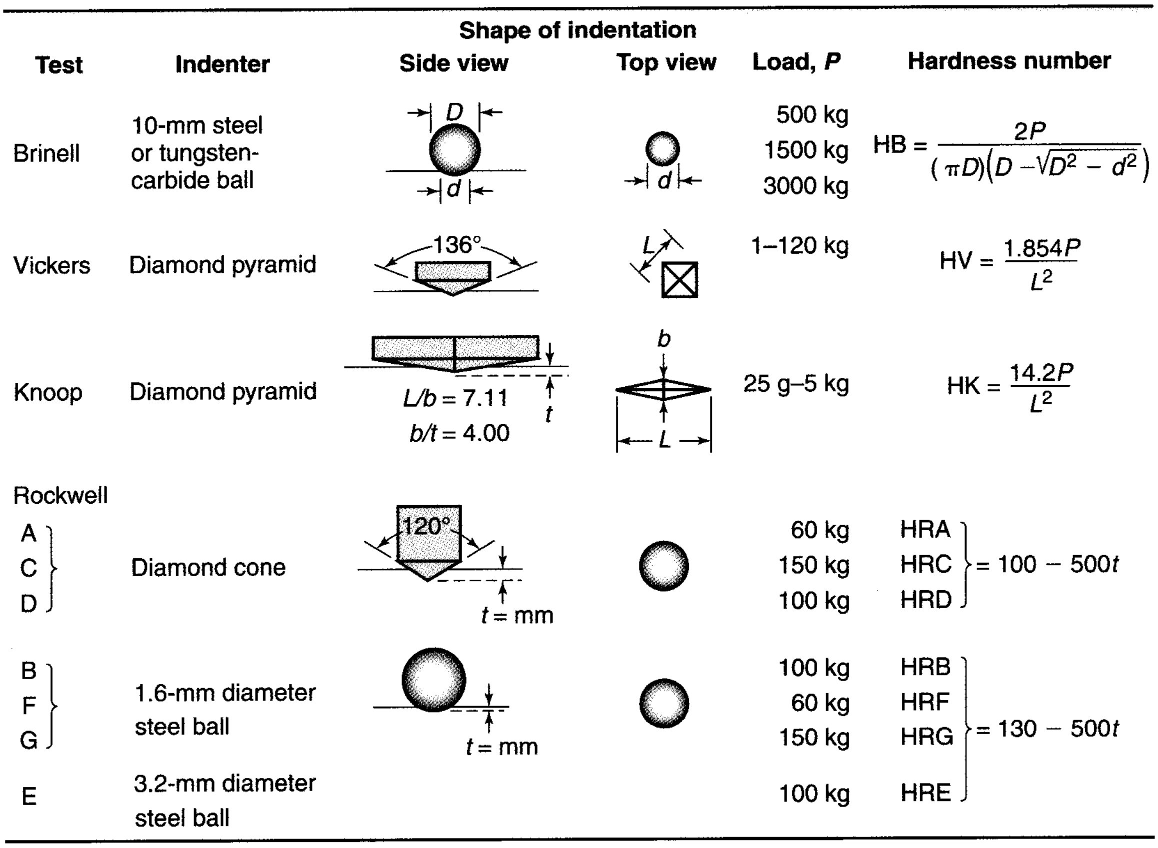


FIGURE 2.12 Selected hardness testers. (a) A Micro Vickers hardness tester; (b) Rockwell hardness tester; (c) Durometer; (d) Leeb tester. *Source*: (a) through (c) Courtesy of Newage Testing Instruments, Inc.; (d) Courtesy of Wilson® Instruments.

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, with a load of 500, 1500, or 3000 kg (Fig. 2.13). The Brinell hardness number (HB) is defined as the ratio of the load P to the curved surface area of the indentation. The harder the material to be tested, the smaller the impression; hence, a 1500-kg or 3000-kg load is usually recommended in order to obtain impressions sufficiently large for accurate measurement. Depending on the condition of the material, one of two types of impression develops on the surface after the performance of this test (Fig. 2.14) or of any of the other tests described in this section. The impressions in annealed metals generally have a rounded profile (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, which has a finite elastic modulus, also undergoes *elastic* deformation under the applied load; as a result, hardness measurements may not be as accurate as expected. One method for minimizing this effect is to use tungsten-carbide balls (Section 22.4); because of their higher modulus of elasticity, they



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

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

distort less than steel balls do. Tungsten-carbide balls are usually recommended for Brinell hardness numbers greater 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 then with a major load; the difference in the depths of penetration is a measure of the hardness of the material. Some of the more common Rockwell hardness scales and the indenters used are shown in Fig. 2.13. Rockwell **superficial hardness** tests using the same type of indenters, but at lighter loads, have also been developed.

Vickers Test. This test, developed in 1922 and formerly known as the *diamond pyramid hardness* test, uses a pyramid-shaped diamond indenter (Fig. 2.13) and a load that ranges from 1 kg to 120 kg. The Vickers hardness number is indicated by HV. The impressions obtained are typically less than 0.5 mm on the diagonal. The Vickers 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. More recently, test procedures have been developed to perform Vickers-type tests in atomic force microscopes and nanoindenters, to estimate hardness at penetration depths 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), with applied loads ranging generally from 25 g to 5 kg. The Knoop hardness number is indicated by HK. Because of the light loads that are applied, the Knoop test is a **microhardness** test; therefore, it is suitable for very small or very thin specimens, and for brittle materials such as carbides, ceramics, and glass.

This test is also used for measuring the hardness of the individual grains and components in a metal alloy. The size of the indentation is generally in the range from 0.01 to 0.10 mm; consequently, surface preparation is very important. Because the hardness number obtained depends on the applied load, Knoop test results should always cite the load used.

Scleroscope and Leeb Test. 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 the material. The impression made by a scleroscope is very small. Because obtaining reliable results with a scleroscope is difficult, a modern electronic version, called a Leeb, or Equotip, test, has been developed (see Fig. 2.12d). In this test, a carbide hammer impacts the surface, and incident and rebound velocities are electronically measured. A *Leeb number* is then calculated and usually converted to Rockwell or Vickers hardness.

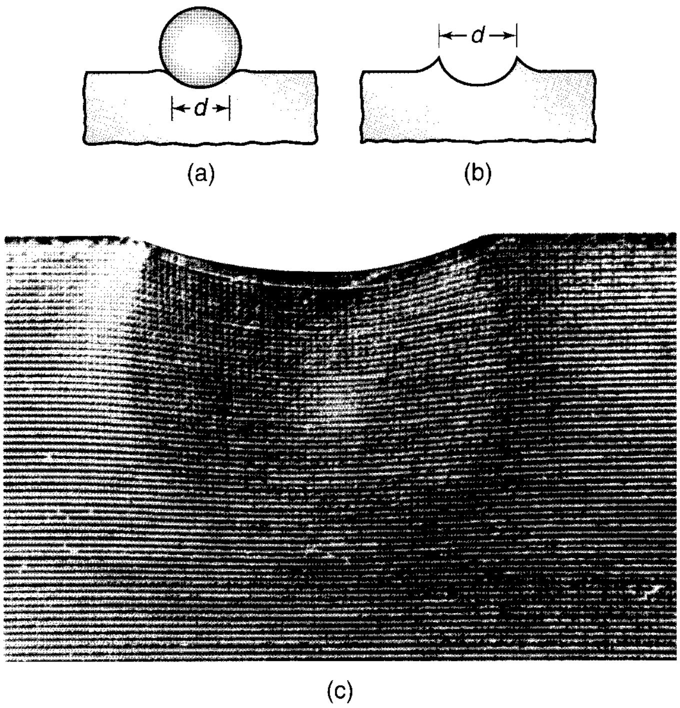


FIGURE 2.14 Indentation geometry in Brinell hardness testing: (a) annealed metal; (b) work-hardened metal; (c) deformation of mild steel under a spherical indenter. Note that the depth of the permanently deformed zone is about one order of magnitude larger than the depth of indentation. For a hardness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

Mohs Hardness. Developed in 1822 by F. Mohs, this test is based on the capability of one material to *scratch* another. The Mohs hardness is based on a scale from 1 to 10, with 1 being the measure for talc and 10 that for diamond (the hardest substance known). A material with a higher Mohs hardness number always scratches one with a lower number. Soft metals have a Mohs hardness of 2 to 3, hardened steels about 6, and aluminum oxide (used for cutting tools and as an abrasive in grinding wheels) of 9. Although the Mohs scale is qualitative and is used mainly by mineralogists, it correlates well with Knoop hardness.

Shore Test and Durometer. The hardness of materials such as rubbers, plastics, and similar 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 pressed against the surface and then a constant load is rapidly applied. The *depth* of penetration is measured after 1 second; the hardness is inversely related to the penetration. There are two different scales for this test. Type A has a blunt indenter and a load of 1 kg; it is used for softer materials. Type D has a sharper indenter and a load of 5 kg, and is used for harder materials. The hardness numbers in these tests range from 0 to 100.

Hot Hardness. The hardness of materials at elevated temperatures (see Fig. 22.1) is important in applications such as cutting tools in machining and dies in hot-working and casting operations. Hardness tests can be performed at elevated temperatures with 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 volume on the surface of a material (Fig. 2.14c). Studies have shown that (in the same units) the hardness of a cold-worked metal is about three times its yield stress, Y; for annealed metals, the hardness is about five times Y.

A relationship has been established between the ultimate tensile strength (UTS) and the Brinell hardness (HB) for steels measured for a load of 3000 kg. In SI units, the relationship is

$$\text{UTS} = 3.5(\text{HB}), \tag{2.13}$$

where UTS is in MPa. In traditional units,

$$\text{UTS} = 500(\text{HB}), \tag{2.14}$$

where UTS is in psi.

2.6.3 Hardness-testing Procedures

For a hardness test to be meaningful and reliable, the **zone of deformation** under the indenter (see Fig. 2.14c) must be allowed to develop freely. Consequently, the *location* of the indenter (with respect to 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 indenter 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.

Moreover, the indentation should be sufficiently large to give a representative hardness value for the bulk material. If hardness variations need to be detected in a small area, or if the hardness of individual constituents in a matrix or an alloy is to be determined, the indentations should 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 values obtained from different hardness tests, on different scales, can be interrelated, and can be converted using Fig. 2.15. Care should be exercised in using these charts because of the many variables in material characteristics and in the shape of the indentation.

EXAMPLE 2.2 Calculation of Modulus of Resilience from Hardness

A piece of steel is highly deformed at room temperature. Its hardness is found to be 300 HB. Estimate the area under 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,

$$Y = \frac{300}{3} = 100 \text{ kg/mm}^2$$

The area under the stress–strain curve is

$$\text{Modulus of Resilience} = \frac{Y^2}{2E}.$$

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

$$\text{Modulus of Resilience} =$$

$$\frac{(100)^2}{2(210{,}000)} \times 9.81 = 0.2336 \text{ mm-kg/mm}^3.$$

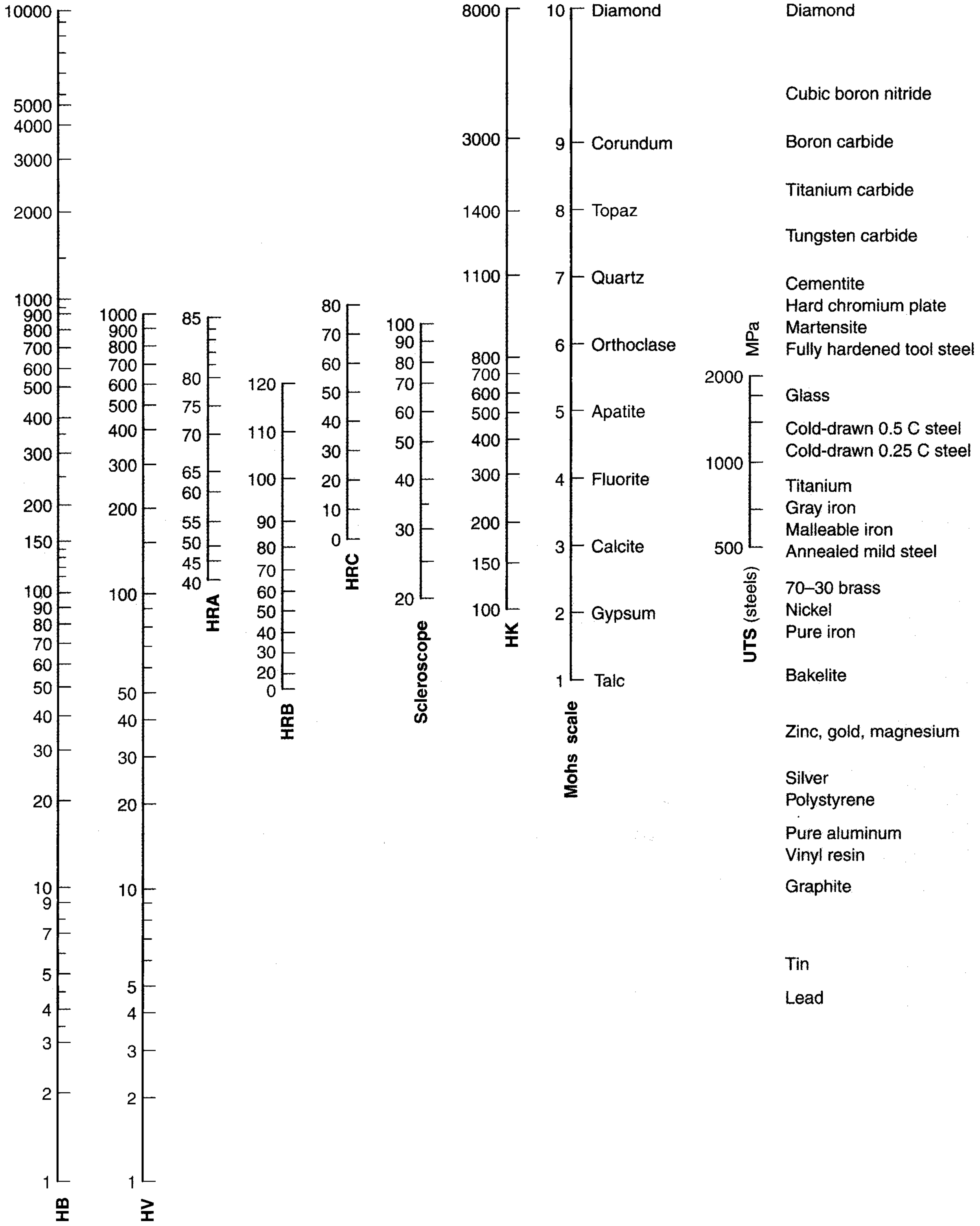


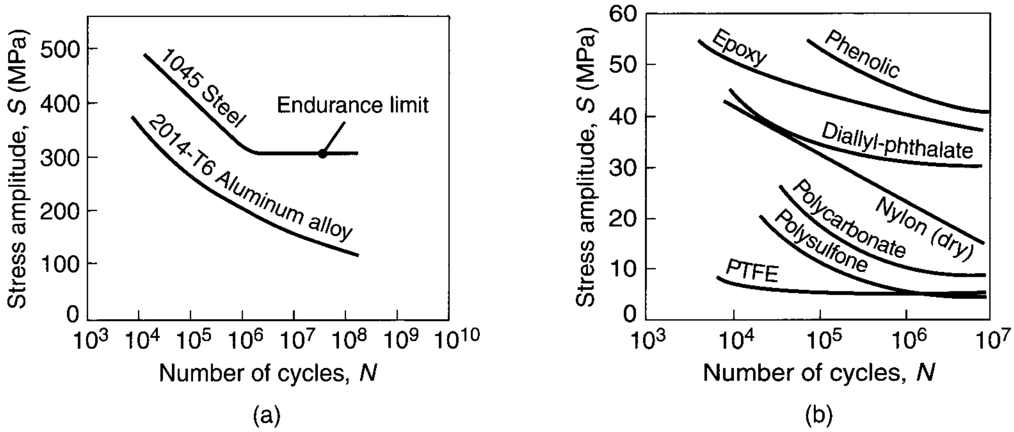
FIGURE 2.15 Chart for converting various hardness scales. Note the limited range of most of the scales. Because of the many factors involved, these conversions are approximate.

2.7 Fatigue

Various components in manufacturing equipment, such as tools, dies, gears, cams, shafts, and springs, are subjected to rapidly fluctuating (cyclic or periodic) loads, in addition to static loads. **Cyclic stresses** may be caused by fluctuating mechanical loads, such as (a) on gear teeth or reciprocating sliders, (b) by rotating machine elements under constant bending stresses, as is commonly encountered by shafts, or (c) by thermal stresses, as when a die comes into repeated contact with hot workpieces and cools between successive contacts. Under these conditions, the part fails 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 grow with every stress cycle and that propagate through the material until a critical crack length is reached, when the material fractures. Known as **fatigue failure**, this phenomenon is responsible for the majority of failures in mechanical components.

Fatigue *test methods* involve testing specimens under various 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) it takes to cause total failure of the specimen or part is 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 rectangular eraser or a piece of wire alternately in one direction and then the other. The test can also be performed on a rotating shaft in four-point bending. (See Fig. 2.11b.) With some materials, the S–N curve becomes horizontal at low stresses, indicating that the material will not fail at stresses below this limit. 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**.

Although many materials, especially steels, have a definite endurance limit, others, such as aluminum alloys, do not have one, 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. In this way, the useful service life of the component can be specified. The endurance limit for metals can be approximately related to their ultimate tensile strength (Fig. 2.17). For carbon steels, the endurance limit is usually 0.4–0.5 times the tensile strength, although particular values can vary.



(a) (b)

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.

2.8 Creep

Creep is the permanent elongation of a component under a static load maintained for a period of time. This phenomenon occurs in metals and certain 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. However, 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** (Section 1.4).

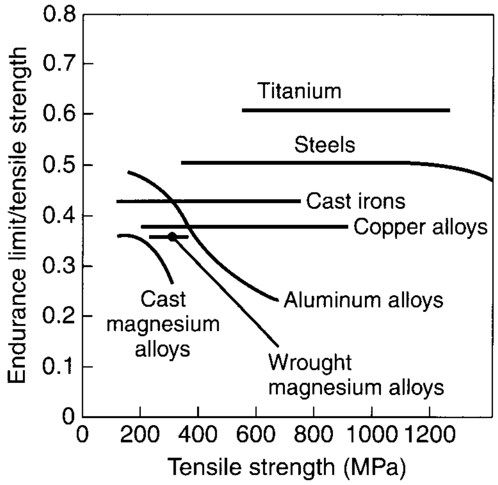


FIGURE 2.17 Ratio of endurance limit to tensile strength for various metals, as a function of tensile strength. Because aluminum does not have an endurance limit, the correlations for aluminum are based on a specific number of cycles, as is seen in Fig. 2.16.

Creep is especially important in high-temperature applications, such as gas-turbine blades and similar components in jet engines and rocket motors; high-pressure steam lines, nuclear-fuel elements, and furnace components are also 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 typical creep curve usually consists of *primary*, *secondary*, and *tertiary* stages (Fig. 2.18). The specimen eventually fails by necking and fracture, called **rupture** or **creep rupture**. As expected, the creep rate increases with specimen temperature and applied load.

Design against creep usually involves a 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. Generally, resistance to creep increases with the melting temperature of a material. Stainless steels, superalloys, and refractory metals and alloys are thus commonly used in applications where resistance to creep is required.

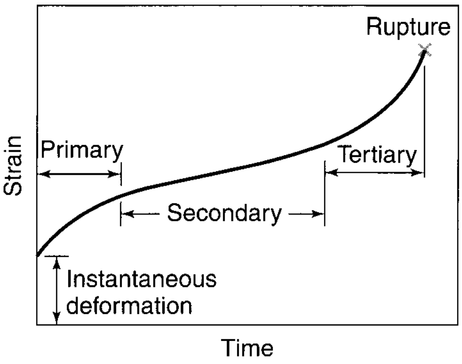


FIGURE 2.18 Schematic illustration of a typical creep curve. The linear segment of the curve (secondary) is used in designing components for a specific creep life.

Stress Relaxation. Stress relaxation is closely related to creep. In *stress relaxation*, the stresses resulting from loading of a structural component decrease in magnitude over a period of time, even though the dimensions of the component remain constant. An example is the decrease in tensile stress of a wire in tension between two fixed ends (as in the wires in a piano); other examples include stress relaxation in rivets, bolts, guy wires, and similar parts under either tension, compression, or flexure. Stress relaxation is particularly common and important in thermoplastics (Section 7.3).

2.9 Impact

In many manufacturing operations and machinery components, materials are subjected to **impact**, or **dynamic loading**—for example, in high-speed metalworking operations such as heading to make bolt heads, and in drop forging (Section 14.9). A typical *impact test* consists of placing a notched specimen in an impact tester and breaking the specimen with a swinging pendulum.

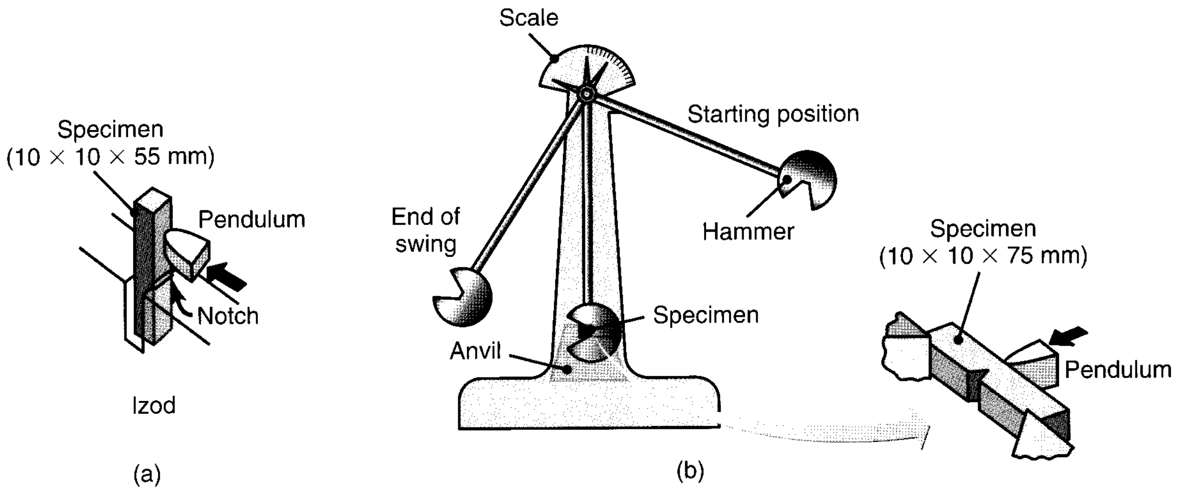


FIGURE 2.19 Impact test specimens. (a) Izod; (b) Charpy.

In the **Charpy** test, the specimen is supported at both ends (Fig. 2.19); in the **Izod** test, it is supported at one end like a cantilever beam (Fig. 2.19a). From the amount of swing of the pendulum, the energy dissipated in breaking the specimen can be obtained; this energy is the **impact toughness** of the material. Unlike hardness-test conversions (Fig. 2.15), no quantitative relationships have yet been established between 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 have high strength, high ductility, and, hence, high toughness. Sensitivity to surface defects (**notch sensitivity**) is important, as it significantly lowers impact toughness, particularly in heat-treated metals and in ceramics and glasses.

2.10 Failure and Fracture of Materials in Manufacturing and in Service

Failure is one of the most important aspects of material behavior, because it directly influences the selection of a material for a particular application, the methods of manufacturing, and the service life of the 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. 2.20 and 2.21).
2. **Buckling**, as shown in Fig. 2.20b.

Although failure of materials is generally regarded as undesirable, some products are designed in such a way that failure is essential for their function.

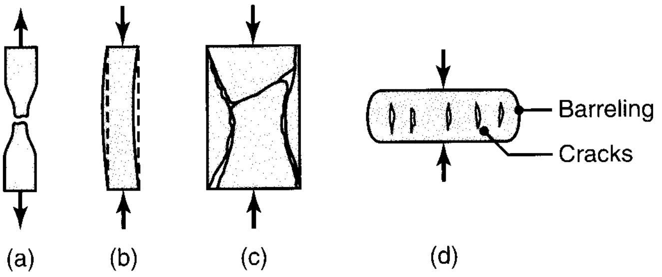


FIGURE 2.20 Schematic illustration of types of failures in materials: (a) necking and fracture of ductile materials; (b) buckling of ductile materials under a compressive load; (c) fracture of brittle materials in compression; (d) cracking on the barreled surface of ductile materials in compression.

Typical examples are (a) food and beverage containers with tabs (or entire tops) which are removed by tearing the sheet metal along a prescribed path; (b) shear pins on shafts that prevent machinery damage in the case of overloads; (c) perforated paper or metal, as in packaging; and (d) metal or plastic screw caps for bottles.

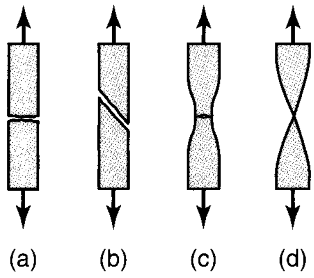


FIGURE 2.21 Schematic illustration of the types of fracture in tension: (a) brittle fracture in polycrystalline metals; (b) shear fracture in ductile single crystals—see also Fig. 1.5a; (c) ductile cup-and-cone fracture in polycrystalline metals; (d) complete ductile fracture in polycrystalline metals, with 100% reduction of area.

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 area and then fail. Ductile fracture generally takes place along planes on which the *shear stress is a maximum*. Thus 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 simple shear, by contrast, is a result of extensive slip along slip planes within the grains. (See Fig. 1.6.)

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 grow in size and eventually lead to fracture.

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

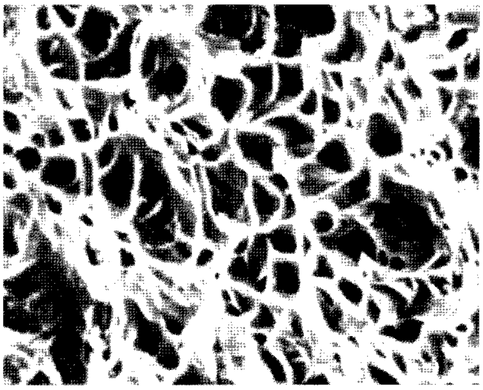


FIGURE 2.22 Surface of ductile fracture in low-carbon steel, showing dimples. Fracture is usually initiated at impurities, inclusions, or preexisting voids (microporosity) in the metal. *Source*: Courtesy of K.-H. Habig and D. Klaffke.

Effects of Inclusions. Because they are nucleation sites for voids, *inclusions* have an important influence on ductile fracture and, consequently, on the workability of materials. 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 fraction of 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 the voids resulting from casting (Section 10.6) and

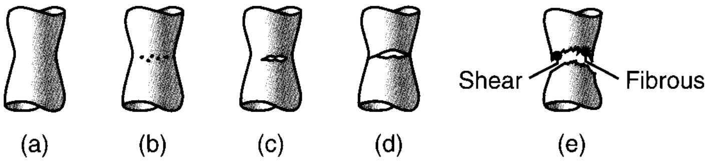


FIGURE 2.23 Sequence of events in the necking and fracture of a tensile-test specimen: (a) early stage of necking; (b) small voids begin to form within the necked region; (c) voids coalesce, producing an internal crack; (d) the rest of the cross section begins to fail at the periphery, by shearing; (e)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

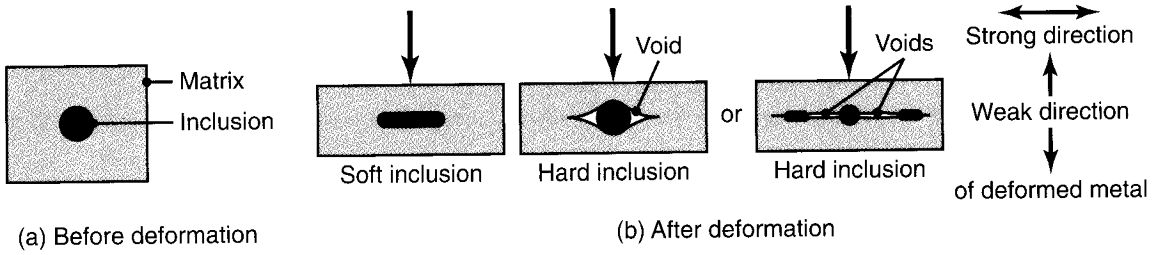


FIGURE 2.24 Schematic illustration of the deformation of soft and hard inclusions and of their effect on void formation in plastic deformation. Note that, because they do not conform to the overall deformation of the ductile matrix, hard inclusions can cause internal voids.

metalworking processes such as drawing and extrusion (Chapter 15). Two factors affect void formation:

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

b. The hardness of the inclusion. If the inclusion is soft, such as one of 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 inclusions, because of their brittle nature, may also break up into smaller particles during deformation.

The alignment of inclusions during plastic deformation leads to **mechanical fibering** (Section 1.5). Subsequent processing of such a material must, therefore, involve considerations of the proper direction of working for maximum ductility and strength.

Transition Temperature. Many metals 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 the composition, microstructure, grain size, surface finish, and shape of the specimen, and the deformation rate. High rates, abrupt changes in workpiece shape, and the presence of surface notches raise the transition temperature.

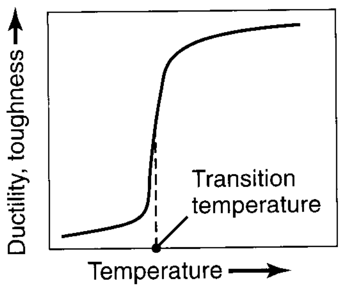


FIGURE 2.25 Schematic illustration of transition temperature in metals.

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; the result is increased strength and reduced ductility. Instead of taking place over several days at room temperature, this phenomenon can occur in just a few hours at a higher temperature; it is then 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 marked 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 gross plastic deformation; in tension, fracture takes place along the crystallographic plane (**cleavage plane**) on which the normal tensile stress is a maximum. Face-centered cubic metals usually do not fail by brittle fracture, whereas body-centered cubic and some hexagonal close-packed metals fail by cleavage. 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, because of the changes in the direction of the cleavage planes as the crack propagates from one grain to another (Fig. 2.26). Brittle fracture of a specimen in compression is more complex, and 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. Tensile stresses normal to the cleavage plane, caused by pulling, initiate and control the propagation of fracture. Another example is the behavior of brittle materials, such as chalk, gray cast iron, and concrete; in tension, they fail 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.

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

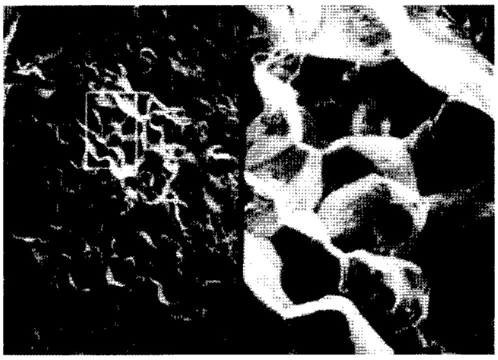


FIGURE 2.27 Intergranular fracture, at two different magnifications. Grains and grain boundaries are clearly visible in this micrograph. The fracture path is along the grain boundaries. Magnification: left, 100×; right, 500×. *Source*: Courtesy of B.J. Schulze and S.L. Meiley and Packer Engineering Associates,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

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 propagate the crack rapidly.

The presence of defects explains why brittle materials exhibit such weakness in tension compared with their strength in compression; see Table 8.2. The ratio of compressive to tensile strength is on the order of 10 for rocks and similar materials, about 5 for glass, and about 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 have been weakened by liquid- or solid-metal embrittlement (Section 1.5.2).

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

FIGURE 2.28 Typical fatigue-fracture surface on metals, showing beach marks. Magnification: left, 500×; right, 1000×. *Source*: Courtesy of B.J. Schulze and S.L. Meiley and Packer Engineering Associates, Inc.

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

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

b. Case hardening (surface hardening) by various means (Section 4.10);

c. Providing a fine surface finish and thereby reducing the effects of notches and other surface imperfections; and

d. Selecting appropriate materials and ensuring that they are free from significant amounts of inclusions, voids, and impurities.

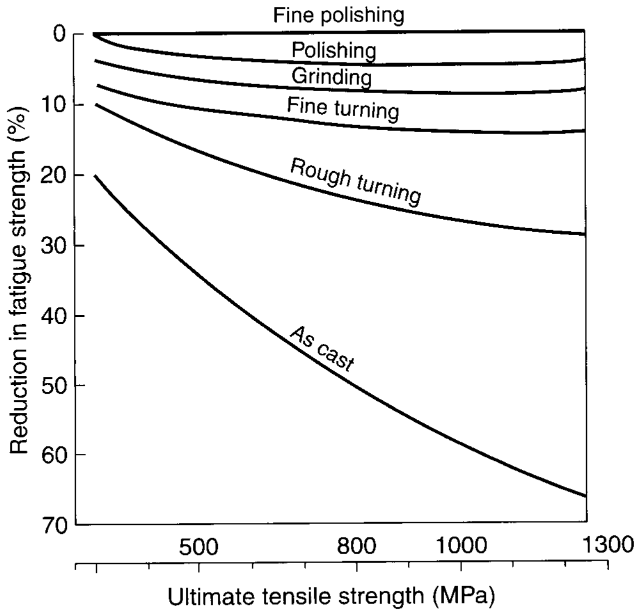


FIGURE 2.29 Reductions in the fatigue strength of cast steels subjected to various surface-finishing operations. Note that the reduction becomes greater as the surface roughness and the strength of the steel increase. *Source*: M.R. Mitchell.

Conversely, the following factors and processes can reduce fatigue strength: tensile residual stresses on the surface (see Section 2.11), decarburization; surface pits (due to corrosion) that act as stress raisers; hydrogen embrittlement; galvanizing; and 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 free from defects may develop cracks, either over time or soon after being manufactured into a product. Crack propagation may be either intergranular or transgranular. The susceptibility of metals to stress–corrosion cracking depends mainly on the material, on the presence and magnitude of *tensile residual stresses*, and on the environment. Brass and austenitic stainless steels are among metals that are highly susceptible to stress cracking; the environment includes corrosive media such as salt water or other chemicals. 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 many metals, alloys, and nonmetallic materials. Known as *hydrogen embrittlement*, this phenomenon is especially severe in high-strength steels. Possible sources of hydrogen arise during melting of the metal, pickling (removing of surface oxides by chemical or electrochemical reaction), and electrolysis in electroplating; other sources are water vapor in the atmosphere and moist electrodes and fluxes used during welding. Oxygen can also cause embrittlement, particularly in copper alloys.

2.11 Residual Stresses

When workpieces are subjected to plastic deformation that is not uniform throughout the part, they develop **residual stresses**. These are stresses that remain within a part after it has been formed and all the external forces (applied through tools and dies) are removed; a typical example is the bending of a metal bar (Fig. 2.30). The bending moment first produces a linear elastic stress distribution (Fig. 2.30a). As the external moment is increased, the outer fibers in the bar reach a stress level high enough to cause yielding. For a typical strain-hardening material, the stress distribution shown in Fig. 2.30b is eventually reached, and the bar has undergone permanent bending.

Let's now remove the external bending moment on the bar. Note that this operation is equivalent to applying an equal but opposite moment to the bar; consequently, 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 the tensile residual stresses in layers *do* and *ef*. Because there are no external forces applied, the internal forces resulting from these residual stresses must be in static equilibrium. Although this example involves only residual stresses in the longitudinal direction of the bar, in most cases these stresses are three dimensional.

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

Residual stresses can also be caused by *temperature gradients* within a body, such as occur during cooling of a casting or a forging. The local expansions and contractions caused by temperature gradients within the material produce a nonuniform deformation, such as is seen in the permanent bending of a beam.

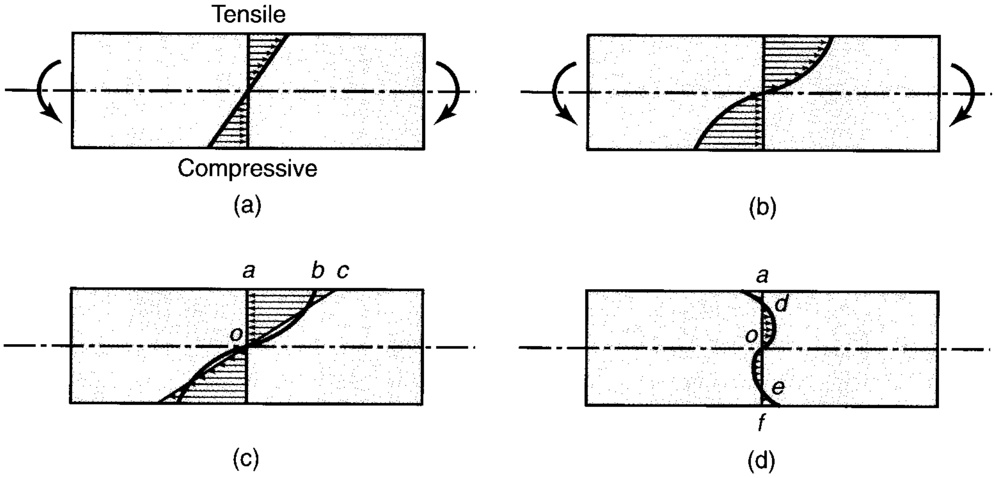


FIGURE 2.30 Residual stresses developed in bending a beam having a rectangular cross section. Note that the horizontal forces and moments caused by residual stresses in the beam must be balanced internally. Because of nonuniform deformation, especially during cold-metalworking operations, most parts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

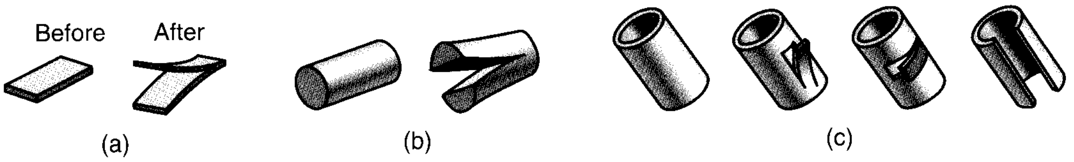


FIGURE 2.31 Distortion of parts with residual stresses after cutting or slitting: (a) flat sheet or plate; (b) solid round rod; (c) thin-walled tubing or pipe.

Tensile residual stresses on the surface of a part are generally undesirable, because they lower the fatigue life and fracture strength of the part. This is due to the fact that a surface with tensile residual stresses cannot sustain additional tensile stresses from external forces as high as those that a surface free from residual stresses can. This 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 to *stress–corrosion cracking* of manufactured products (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 imparted to surfaces by techniques such as shot peening and surface rolling (Section 34.2).

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

2.12 Work, Heat, and Temperature

Almost all the mechanical work in plastic deformation is converted into **heat**. This conversion is not complete, because a portion of this work is stored within the deformed material as elastic energy. Known as **stored energy** (Section 1.7), it is generally 5 to 10% of the total energy input; in some 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, 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 seen that higher temperatures are associated with large areas under the stress–strain curve and with smaller values of specific heat. However, such physical properties (Chapter 3) as specific heat and thermal conductivity can 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 to be as follows: aluminum, 75°C; copper, 140°C; low-carbon steel, 280°C; and titanium 570°C. In actual operations, heat is lost to the environment, to tools and dies, and to lubricants or coolants used, if any. If the deformation process is performed rapidly, the heat losses will be relatively small over that brief period. If the process is carried out slowly, the actual temperature rise will be only a fraction of the calculated value.

SUMMARY

- Many manufacturing processes involve shaping materials by plastic deformation; consequently, such mechanical properties as strength (yield strength, Y, and ultimate tensile strength, UTS); modulus of elasticity, E; ductility (total elongation and reduction of area); hardness; and the energy required for plastic deformation are important factors. These properties, in turn, depend, to various extents, on the particular material and on its condition, temperature, deformation rate, surface condition, and environment.
- The tensile test is the most commonly used test to determine mechanical properties; from these tests, true stress–true strain curves are constructed that are needed to 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 to resultant barreling of the specimen. Torsion tests are conducted on tubular specimens 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 used to determine the resistance of a material to permanent indentation or scratching. Hardness is related to strength and wear resistance of a material, but it is, itself, not a fundamental property.
- Fatigue tests indicate the endurance limit or fatigue limit of materials—that is, 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; and instead, their allowable stress must be 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. The specimen eventually fails by rupture (necking and fracturing).
- Impact tests determine the energy required to completely break a specimen. This energy is called the impact toughness of the material. Impact tests are also useful for determining the transition temperatures of materials.
- Failure and fracture constitute an important aspect of a material's behavior when it is subjected to deformation in manufacturing operations. Ductile fracture is characterized by plastic deformation preceding fracture, and it requires a considerable amount of energy. Brittle fracture can be catastrophic, because it is not preceded by plastic deformation; it requires much less energy than 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 had 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 strength
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**, 3rd ed. Pergamon, 2005.

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

ASM Handbook, Vol. 8: **Mechanical Testing and Evaluation.** ASM International, 2000.

Budinski, K.G., and Budinski, M.K., **Engineering Materials: Properties and Selection**, 8th ed., Prentice Hall, 2004.

Cardelli, F., **Materials Handbook: A Concise Desk Reference**, 2nd ed., Springer, 2008.

Chandler, H., **Hardness Testing**, 2nd ed., ASM International, 1999.

Courtney, T.H., **Mechanical Behavior of Materials**, 2nd ed., Waveland Press, 2005.

Davis, J.R. (ed.), **Tensile Testing**, 2nd ed., ASM International, 2004.

Dowling, N.E., **Mechanical Behavior of Materials: Engineering Methods for Deformation, Fracture, and Fatigue**, 3rd ed., Prentice Hall, 2006.

Hosford, W.F., **Mechanical Behavior of Materials**, Cambridge, 2005.

Wulpi, D.J., **Understanding How Components Fail**, 2nd ed., ASM International, 1999.

REVIEW QUESTIONS

2.1. Distinguish between engineering stress and true stress.

2.2. 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.

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

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

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

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

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

2.8. Describe the differences between brittle and ductile fracture.

2.9. Differentiate between stress relaxation and creep.

2.10. Describe the difference between elastic and plastic behavior.

2.11. Explain what uniform elongation means in tension testing.

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

2.13. Describe the difficulties involved in making a compression test.

2.14. What is Hooke's law? Young's modulus? Poisson's ratio?

2.15. Describe the difference between transgranular and intergranular fracture.

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

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

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

QUALITATIVE PROBLEMS

2.19. 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.

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

2.21. 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.

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

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

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

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

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

2.27. On the basis of Fig. 2.6, can you calculate the percent elongation of the materials listed? Explain.

2.28. If a metal tension-test specimen is pulled and broken rapidly, where would the temperature be highest, and why?

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

2.30. Will the disk test be applicable to a ductile material? Why or why not?

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

2.32. 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 dimples, while others display transgranular fracture surfaces. Explain these observations.

2.33. 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.

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

2.35. 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 nanometers) to indent such coatings in atomic force microscopes. What concerns would you have regarding the appropriateness of the results?

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

a. Cubic boron nitride
b. Lead
c. Cold-drawn 0.5%C steel
d. Diamond
e. Caramel candy
f. Granite

QUANTITATIVE PROBLEMS

2.37. A paper clip is made of wire 0.7 mm in diameter. If the original material from which the wire is made is a rod 25 mm in diameter, calculate the longitudinal engineering and true strains that the wire has undergone during processing.

2.38. A 200-mm-long strip of metal is stretched in two steps, first to 300 mm and then to 400 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.

2.39. Identify the two materials in Fig. 2.6 that have the lowest and the highest uniform elongations. Calculate these quantities as percentages of the original gage lengths.

2.40. Plot the ultimate strength vs. 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.

2.41. 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.)

2.42. 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.)

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

2.44. 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?

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

For material A, $K = 415$ MPa, $n = 0.5$, $A_o = 380\ \text{mm}^2$.

For material B, $K = 210$ MPa, $n = 0.5$, $A_o = 190\ \text{mm}^2$.

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

2.46. On the basis of the information given in Fig. 2.6, calculate the ultimate tensile strength (engineering) of 304 stainless steel.

2.47. In a disk test performed on a specimen 38 mm in diameter and 12 mm thick, the specimen fractures at a stress of 275 Mpa. What was the load on it?

2.48. A piece of steel has a hardness of 330 HB. Calculate its tensile strength, in MPa and in psi.

2.49. A metal has the following properties: UTS = 480 MPa and $n = 0.25$. Calculate its strength coefficient, K.

2.50. Using only Fig. 2.6, calculate the maximum load in tension testing of an annealed copper specimen with an original diameter of 5 mm.

2.51. 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.

2.52. A metal has a strength coefficient $K = 700$ MPa and $n = 0.2$. Assuming that a tensile-test specimen made from this metal begins to neck at a true strain of 0.2, show that the ultimate tensile strength is 415 MPa.

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

2.54. 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 60 μm, is the material acceptable?

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

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

2.57. 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.

2.58. 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.

SYNTHESIS, DESIGN, AND PROJECTS

2.59. List and explain the desirable mechanical properties of (a) an elevator cable, (b) a paper clip, (c) a leaf spring for a truck, (d) a bracket for a bookshelf, (e) piano wire, (f) a wire coat hanger, (g) the clip for a pen, and (h) a staple.

2.60. 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).]

2.61. 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 1 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.

2.62. 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?

2.63. 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?

2.64. 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.

2.65. 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.

2.66. 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.

2.67. 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 Figure 2.5? Comment.

2.68. Figure 2.4 shows the ductility of only four metals. Search the technical literature, and add other metals and alloys to this chart, including their different conditions, such as hot worked, cold worked, annealed, etc.

2.69. 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, etc., and describe your observations.

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

2.71. 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.

▶**2.72.** 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.

2.73. Describe your observations regarding Fig. 2.14c.

2.74. 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.

2.75. 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.

2.76. 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.

2.77. Explain how you would obtain an estimate of the hardness for a carbon nanotube. (See Section 8.6.)

CHAPTER 3

Physical Properties of Materials

EXAMPLE:

- Physical properties can have several important roles in the selection, processing, and use of materials.
- These properties are 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 discussed in the context of lightweight designs, a key consideration in aerospace and automotive industries.
- Thermal, electrical, magnetic, and optical properties are then presented.
- The importance of corrosion and corrosion-resistant materials are described.
- Design and manufacturing implications of all of these properties are considered, with various specific examples given.

3.1 Introduction

Why is electrical wiring generally made of copper? Why are aluminum, stainless steel, and copper so commonly used in cookware? Why are the handles of cookware usually made of wood or plastic, while other types of handles are made of metal? What kind of material should be chosen for the heating elements in toasters? Why are the metallic components in some machines being replaced with ceramics? Why are commercial airplane bodies generally made of aluminum, and why are some airplane components being replaced gradually with those made of various composite materials, including reinforced plastics?

It is apparent from these questions that one important criterion in material selection is consideration of **physical properties**, such as 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 thus keep machines from excessive vibration. Several other examples of the importance of physical properties are discussed in this chapter, with each property being presented from the viewpoint of material selection and manufacturing and its relevance to the service life of the component.

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 in relation to that of water; thus, specific gravity has 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.

TABLE 3.1

Physical Properties of Selected Materials at Room Temperature

	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}
Columbium (niobium)	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	29–234	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	1770–1780	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	$6.2{-}110 \times 10^{-8}$
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}
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	5×10^{-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	—

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	Columbium	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		

Weight saving is particularly important for aircraft and aerospace structures, for automotive bodies and components, and for other products where energy consumption and power limitations are major concerns. Substitution of materials for the sake of weight savings and economy is a major factor in the design both of advanced equipment and machinery and of consumer products such as automobiles.

A significant role that density plays is in the **strength-to-weight ratio** (**specific strength**) and **stiffness-to-weight ratio** (**specific stiffness**) of materials and structures. 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, and

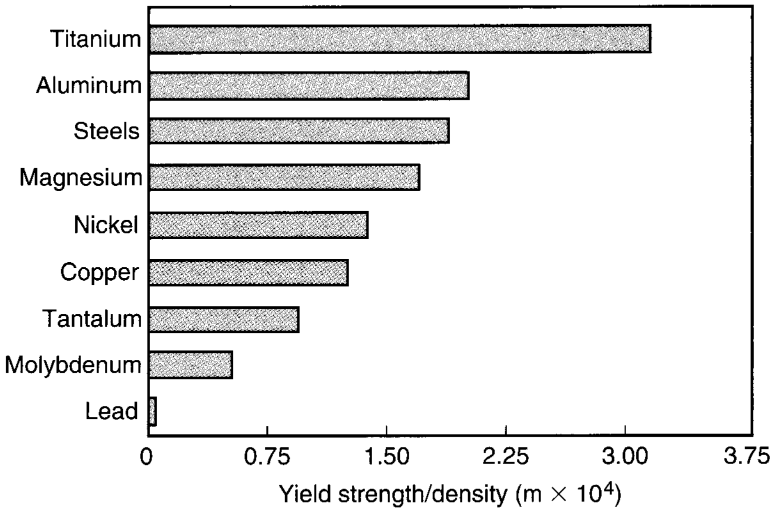


FIGURE 3.1 Ratio of maximum yield stress to density for selected metals.

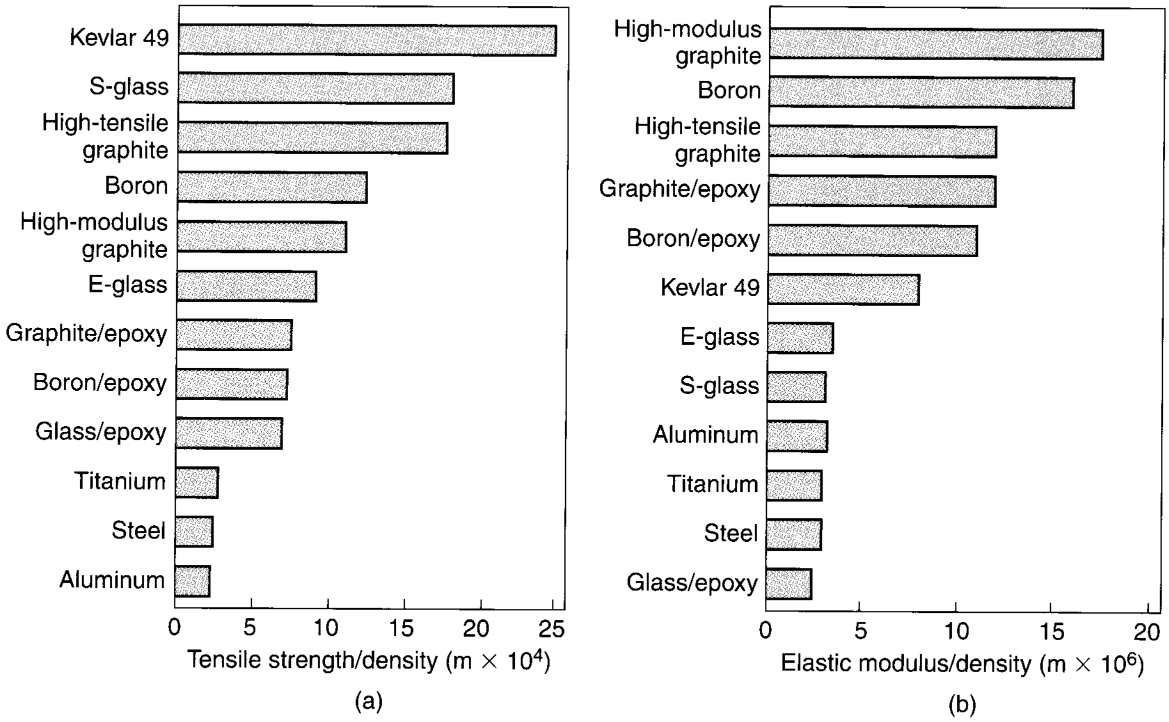


FIGURE 3.2 Specific strength (tensile strength/density) and specific stiffness (elastic modulus/density) for various materials at room temperature. (See also Chapter 9.)

as described in Chapter 6, they are among the most commonly used metals for aircraft and aerospace applications.

The range for specific tensile strength and specific stiffness at room temperature for a variety of metallic and nonmetallic materials is 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 (see Chapter 9). At elevated temperatures, specific strength and specific stiffness are likewise important considerations, especially for systems that operate at these temperatures, such as automotive and jet engines and gas turbines. 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 usually operate at very high speeds. To obtain exposure times of 1/4000 s in cameras without sacrificing accuracy, the focal plane shutters of some high-quality digital cameras are made of titanium. Aluminum is used with some cameras for better performance in cold weather. The low resulting mass of the components in these high-speed operations reduces inertial forces that otherwise could lead to vibrations, to inaccuracies, and even, over time, to part failure. Because of their low density, ceramics (Chapter 8) are used for components in high-speed automated machinery and in 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 yachts and aircraft, and weights on golf clubs (using high-density materials such as tungsten).

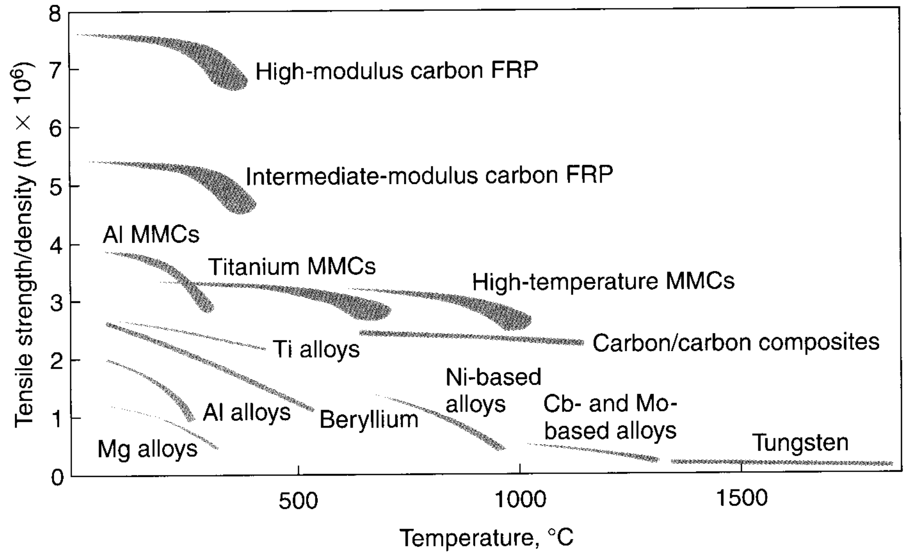


FIGURE 3.3 Specific strength (tensile strength/density) for a variety of materials as a function of temperature. Note the useful temperature range for these materials and the high values for composite materials. MMC—metal-matrix composite; FRP—fiber-reinforced plastic.

3.3 Melting Point

The **melting point** of a metal depends on the energy required to separate its atoms. As Table 3.1 shows, the melting temperature of a metal alloy can have a wide range, depending on its composition, and is unlike that of a pure metal, which has a definite melting point. The temperature range within which a component or structure is designed to function is an important consideration in the selection of materials. Plastics, for example, have the lowest useful temperature range, while graphite and refractory-metal alloys have the highest useful range.

The melting point of a metal has a number of 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 heat treating (Chapter 4), and hot-working (Part III), require a knowledge of the melting points of the materials involved. These considerations are also important in the selection of tool and die materials. Melting point also plays a major role in the selection of the equipment and the melting practice employed in casting operations (Part II). The higher the melting point of the material, the more difficult the operation becomes. 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 1 degree. Alloying elements have a relatively minor effect on the specific heat of metals. The temperature rise in a workpiece, resulting from forming or machining operations (Parts III and IV, respectively), is a function of the work done and of the specific heat of the workpiece material (Section 2.12). An excessive temperature rise in a workpiece can decrease product quality by adversely affecting its surface finish and dimensional accuracy, can cause excessive tool and die wear, and can result in undesirable metallurgical changes in the material.

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 the 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 many applications. For example, high thermal conductivity is desirable in cooling fins, cutting tools, and die-casting molds to extract heat. In contrast, materials with low thermal conductivity are used, for instance, in furnace linings, insulation, coffee cups, and handles for pots and pans.

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 expansions. 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.

Shrink fits utilize thermal expansion and contraction. A shrink fit is a part, often a tube or 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.

Thermal expansion in conjunction with thermal conductivity plays the most significant role in causing **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, 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 can also 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 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 metalworking operations (*heat checking*). **Thermal shock** is the term generally used to describe development of cracks after 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 has been developed, called **low-expansion alloys**. The low thermal expansion characteristic of these alloys is often referred to as the **Invar effect**, after the metal *Invar*. The thermal coefficient of expansion is typically in the range of from 2×10^{-6} to 9×10^{-6} per °C. 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 good ductility; as a result, they can easily be formed into various shapes. Applications include (a) bimetallic strips consisting of a low-expansion alloy metallurgically bonded to a high-expansion alloy (the strip bends when subjected to temperature changes) and (b) high-quality glass-to-metal seals in which the thermal expansions are matched.

3.7 Electrical, Magnetic, and Optical Properties

Electrical conductivity and the *dielectric* properties of materials are important not only in electrical equipment and machinery, but also in such manufacturing processes as the magnetic-pulse forming of sheet metals (Section 16.11), resistance welding (Section 31.5), and the 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 reverse of *ohm*, the unit of electrical resistance. The influence of the type of atomic bonding on the electrical conductivity of materials is the same as that for thermal conductivity. 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 it can encounter 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, or –273°C). The highest temperature at which superconductivity has been exhibited to date is –123°C, but advances in high temperature superconductivity continue to be made.

The main application of superconductors is as high-power magnets. Also, superconductors are the enabling technology for magnetic resonance imaging (MRI), which is widely used for medical imaging. Other applications proposed 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 (**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 what are called **smart materials**. Two basic behaviors are involved: (a) When subjected to an electric current, these materials undergo a reversible change in shape, by as much as 4%, and (b) when deformed by an external force, the materials emit a small electric current. This unique property is utilized in making transducers, which are devices that convert the strain from an external force into electrical energy. Piezoelectric materials include quartz crystals and some ceramics and polymers, Typical applications are sensors, force or pressure transducers, inkjet printers, strain gages, sonar detectors, and microphones. For example, an air bag in an automobile has a sensor that, when subjected to an impact force, sends an electric charge which deploys the bag.

Magnetostriction. The phenomenon of expansion or contraction of a material when it is subjected to a magnetic field is called *magnetostriction*. Materials such as 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**, these materials are being developed for applications such 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. (These two properties are described in Sections 7.2.2 and 8.4.3, respectively.)

3.8 Corrosion Resistance

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 are generally called **degradation**. Corrosion not only leads to surface deterioration of components and structures (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 3% of the gross domestic product; indirect costs of corrosion are estimated at twice this amount.

Corrosion resistance is an important aspect of material selection for applications in the chemical, food, and petroleum 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 components and structures is a major concern, particularly at elevated temperatures and in automobiles and other transportation vehicles.

Resistance to corrosion depends on the composition of the material and on the particular environment. Corrosive media may be chemicals (acids, alkalis, and salts), the environment (oxygen, moisture, pollution, and acid rain), and water (fresh or salt water). Nonferrous metals, stainless steels, and nonmetallic materials generally have high corrosion resistance. Steels and cast irons generally have poor resistance and must be protected by various 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 33.2.) 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 (Chapter 4) 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, had residual stresses. Likewise, cold-worked metals are likely to have residual stresses; hence, they are 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 by coolants; the chemical reaction alters their surface finish and adversely influences the metalworking operation. One example is that of carbide tools and dies having cobalt as a binder (Section 22.5). The cobalt is attacked by elements in the metalworking fluid (**selective leaching**). Thus, compatibility of the tool, die, and workpiece materials with the metalworking fluid under actual operating conditions is an important consideration in the selection of materials.

Chemical reactions should not be regarded as having only adverse effects. Advanced machining processes such as chemical and electrochemical machining are indeed based on controlled chemical reactions (Chapter 27). These processes remove material by chemical action, in a manner similar to the etching of metallurgical specimens. The usefulness of some level of **oxidation** is exhibited in the corrosion resistance of aluminum, titanium, and stainless steel. Aluminum 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 on their surfaces. These processes are known as **passivation.** When the protective film is scratched and exposes the metal underneath, a new oxide film begins to form.

EXAMPLE 3.1 Selection of Materials for Coins

There are five general criteria in the selection of materials for coins.

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 difficult to describe, because it combines many human factors. It is similar in effect to the feel of a fine piece of wood, polished stone, or fine leather.

2. The intended *life of the coin* is another consideration; this duration will reflect resistance to corrosion and to wear (Chapter 32) while the coin is in circulation. These two factors basically determine the span over which the surface imprint of the coin will remain identifiable and 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 14.3.2), and the capability of the materials and processes to resist counterfeiting.
4. Another consideration is the *suitability of the coin for use* in coin-operated devices, such as vending machines, turnstiles, and pay telephones. 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 these tests.
5. A final consideration is the *cost* of raw materials and processing, and whether there is a sufficient *supply* of the coin materials.

SUMMARY

- Physical and chemical properties can have several important influences on materials selection, manufacturing, and on the service life of components. These properties and other material characteristics should be considered because their effects on product design, service requirements, and compatibility with other materials, including tools, dies, and workpieces.
- The combined properties of 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 which are important in tool and die life in manufacturing operations.
- Chemical reactions, including oxidation and corrosion, are considerations in material selection, design, and manufacturing, as well as in the service life of components. Passivation and stress-corrosion cracking are two important phenomena.
- Certain physical properties are utilized in manufacturing processes and their control, such as the magnetostriction effect (for ultrasonic machining of metals and nonmetallic materials) and the piezoelectric effect (for force transducers and various other sensors).

KEY TERMS

Corrosion
Conductors
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., **Materials Selection in Mechanical Design**, 3rd ed., Butterworth-Heineman, 2005.

ASM Handbook, Vol, 13A: **Corrosion: Fundamentals, Testing, and Protection**, 2003; Vol. 13B: *Corrosion Materials*, 2005; Vol. 13B: *Corrosion: Environments and Industries*, 2006, ASM International.

Cardelli, F., **Materials Handbook: A Concise Desk Reference**, 2nd ed., Spinger, 2008.

Hummel, R.E., **Electronic Properties of Materials**, 3rd ed., Springer, 2004.

Pollock, D.D., **Physical Properties of Materials for Engineers**, 2nd ed., CRC Press, 1993.

Schweitzer, P.A., **Encyclopedia of Corrosion Technology**, 3rd ed., Marcel Dekker, 2004.

Solymar, L., and Walsh, D, **Electrical Properties of Materials**, Oxford, 2004.

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. What is the piezoelectric effect?

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

3.6. What is passivation? What is its significance?

3.7. What is the difference between thermal conductivity and specific heat?

3.8. Explain stress-corrosion cracking. Why is it also called season cracking?

3.9. What is the difference between a superconductor and a semiconductor?

3.10. What are smart materials?

QUALITATIVE PROBLEMS

3.11. Describe the significance of structures and machine components made of two materials with different coefficients of thermal expansion.

3.12. 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.

3.13. 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.

▶**3.14.** Rank the following in order of increasing thermal conductivity: aluminum, copper, silicon, titanium, ceramics, and plastics. Comment on their applications vis-à-vis these materials.

3.15. Does corrosion have any beneficial effects? Explain.

3.16. Explain how thermal conductivity can play a role in the development of residual stresses in metals.

▶**3.17.** What material properties are desirable for heat shields such as those placed on the space shuttle?

3.18. List examples of products where materials that are transparent are desired. List applications for opaque materials.

3.19. Refer to Fig. 3.2 and explain why the trends seen are to be expected.

QUANTITATIVE PROBLEMS

▶**3.20.** 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. 2.6, 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.

3.21. 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.

▶**3.22.** 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.

▶**3.23.** Add a column to Table 3.1 that lists the volumetric heat capacity of the materials listed, expressed in units of J/cm^3 K. Compare the results to the value for liquid water (4.184 J/cm^3 K). Note that the volumetric heat capacity of a material is the product of its density and specific heat.

3.24. Conduct a literature search and add the following materials to Table 3.1: cork, cement, ice, sugar, lithium, chromium, and platinum.

SYNTHESIS, DESIGN, AND PROJECTS

3.25. From your own experience, make a list of parts, components, or products that have corroded and have had to be replaced or discarded.

3.26. 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.

3.27. Describe several applications in which both specific strength and specific stiffness are important.

3.28. 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.

3.29. For the materials listed in Table 3.1, determine the specific strength and specific stiffness. Describe your observations.

3.30. 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.

3.31. Describe possible applications and designs using alloys exhibiting the Invar effect of low thermal expansion.

3.32. 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.

3.33. 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.

CHAPTER

4

Metal Alloys: Their Structure and Strengthening by Heat Treatment

- This chapter examines the structures of alloys, including solid solutions, intermetallic compounds, and two-phase systems.
- Phase diagrams show graphically the various phases that develop as a function of alloy composition and temperature.
- The important system of iron and carbon and the phases that can be developed are described in detail.
- Heat treatment is a common strategy to improve a material's mechanical properties; it involves establishing a desired phase at elevated temperatures, followed by controlled cooling of the material before the microstructure can transform into a different phase. This procedure can profoundly affect the behavior of steel and produce special forms, including pearlite, martensite, spheroidite, and bainite, each with its own unique characteristics.
- Some materials can be heat treated only by precipitation hardening or aging.
- Often, improving the ductility of a material is at the expense of properties such as strength or hardness. Exposure of a metal to elevated temperatures for a certain period of time is often practiced to achieve desired properties.
- The chapter ends with a discussion 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 processing history as well as on 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 heat-treatment processes. The properties of non-heat-treatable alloys are improved by mechanical working operations, such as rolling, forging, and extrusion (Part III).

The most common example of a process that improves properties is *heat treatment* (Sections 4.7–4.10), which modifies microstructures and thereby produces a variety of mechanical properties that are important in manufacturing, such as improved formability and machinability or increased strength and hardness for tools and dies. These properties also enhance the service performance of the metals when used in machine components (such as gears, cams, and shafts; Fig. 4.1) or in tools, dies, and molds.

This chapter follows the outline shown in Fig. 4.2, beginning with the effects 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 discusses methods and techniques of heating, quenching, tempering, and annealing, and describes the characteristics of the equipment involved.

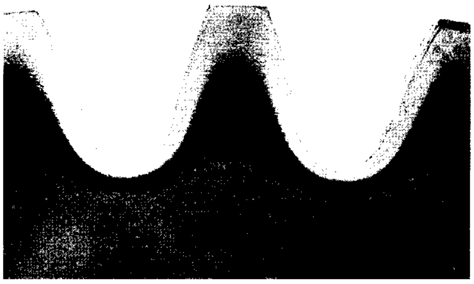


FIGURE 4.1 Cross section of gear teeth showing induction-hardened surfaces. *Source*: TOCCO Div., Park-Ohio Industries, Inc.

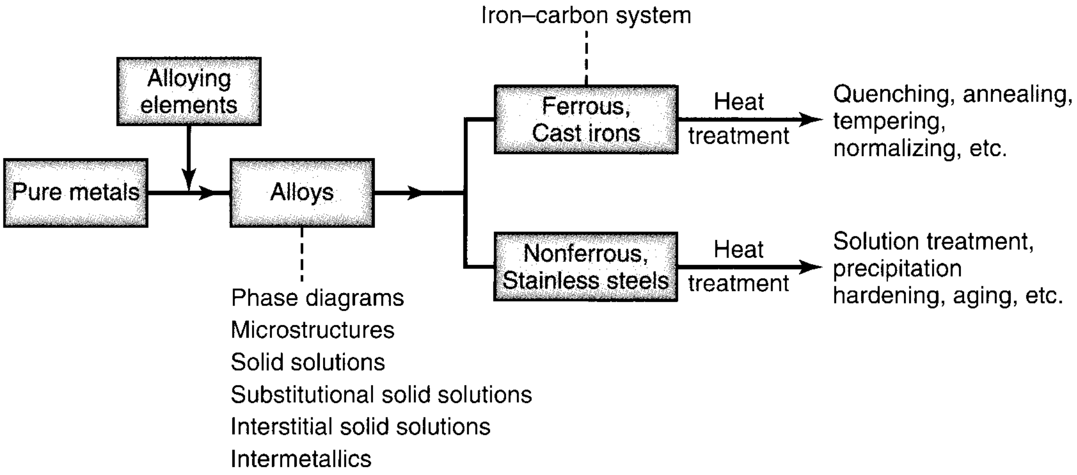


FIGURE 4.2 Outline of topics described in Chapter 4.

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 rare 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: aluminum for foil, copper for electrical conductors, nickel or chromium for plating, and gold for electrical contacts. However, pure metals have somewhat limited properties; these properties can be enhanced and modified by **alloying**. An **alloy** is composed of two or more chemical elements, at least one of which is a metal. The majority of metals used in engineering applications are some form of alloy. 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 involved 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* (see Fig. 1.7). An example is brass (Section 6.4), an alloy of zinc and copper, in which zinc (the solute atom) is introduced into the lattice of copper (solvent atoms). The properties of brasses can thus be altered by controlling the amount of zinc in copper.

Two conditions (known as *Hume-Rothery rules*, after W. Hume-Rothery, 1899–1968) are generally required to form complete substitutional solid solutions:

1. The two metals must have similar crystal structures, and
2. The difference in their atomic radii should be less than 15%.

If these conditions are not satisfied, complete solid solutions will not be obtained and the amount of solid solution formed will be limited.

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; such a process forms an *interstitial solid solution.* There are two conditions necessary for forming interstitial solutions:

1. The solvent atom must have more than one valence electron, and
2. The atomic radius of the solute atom must be less than 59% of the atomic radius for the solvent atom.

If these conditions are not met, interstitial solubility is limited and may not take place at all.

An important family of interstitial solid solutions is **steel** (Chapter 5), 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 less than 59% of the 0.124-nm radius of the iron atom. The properties of carbon steels can be varied over a wide range by controlling the proportion of carbon to iron. This controllability is a major reason that steel is such a versatile and useful material with a wide variety 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 certain proportions. Some intermetallic compounds have solid solubility, and the type of 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, they are candidate materials for applications such as advanced gas-turbine engines. Typical examples are the aluminides of titanium (Ti_3Al), nickel (Ni_3Al), and iron (Fe_3Al).

4.2.3 Two-phase Systems

It has been noted that a solid solution is one in which two or more elements in a solid state form a single homogeneous solid phase in which the elements are uniformly distributed throughout the solid mass. Such a system is limited by some 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 may be regarded as mechanical mixtures; such a system with two solid phases is known as a **two-phase system**.

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. For example, consider a mixture of sand and water as an example of a two-phase system. These two 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; in this case, the two phases are the same chemical compound of exactly the same chemical elements (hydrogen and oxygen), even though their properties are very different. (Note that it is not necessary that one phase be a liquid; for example, sand suspended in ice is also a two-phase system.)

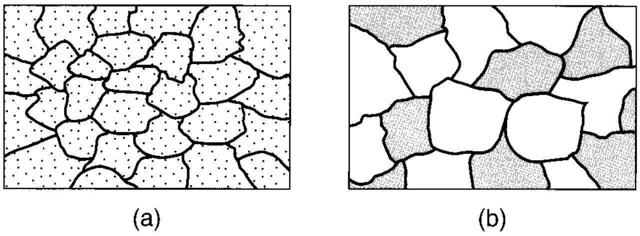


FIGURE 4.3 (a) Schematic illustration of grains, grain boundaries, and particles dispersed throughout the structure of a two-phase system, such as a lead–copper alloy. The grains represent lead in solid solution in copper, and the particles are lead as a second phase. (b) Schematic illustration of a two-phase system consisting of two sets of grains: dark and light. The green and white grains have separate compositions and properties.

A typical 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: one having a small amount of lead in solid solution in copper, the other having lead particles (roughly spherical in shape) *dispersed* throughout the structure (Fig. 4.3a). The lead particles are analogous to the sand particles in water described above. This copper–lead alloy has properties that are different from those of either copper or lead alone. Lead is also added to steels to obtain leaded steels, which have greatly improved machinability (Section 21.7).

Alloying with finely dispersed particles (**second-phase particles**) is an important method of strengthening alloys and controlling their properties. In two-phase alloys, the second-phase particles present obstacles to dislocation movement and thus increase strength. Another example of a two-phase alloy is the aggregate structure shown in Fig. 4.3b, where there are two sets of grains, each with its own composition and properties. The darker grains may have a different structure from the lighter grains; they may, for example, be brittle, while the lighter grains are ductile. Defects may appear 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* (Fig. 4.4). 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. 4.5). Solidification begins when the temperature of the molten metal drops below the **liquidus**; it 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 described by the particular alloy's phase diagram.

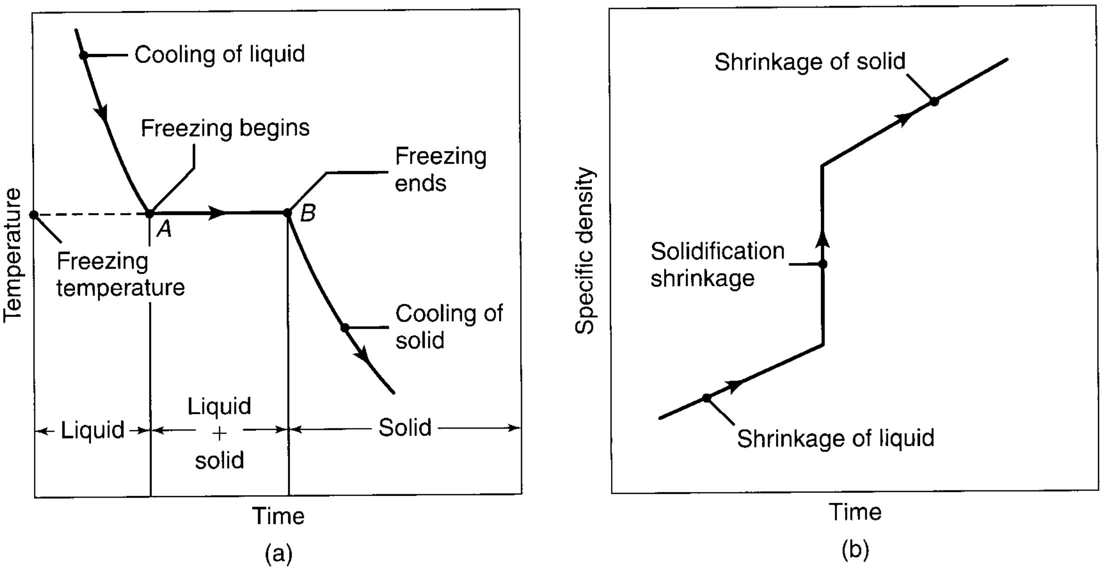


FIGURE 4.4 (a) Cooling curve for the solidification of pure metals. Note that freezing takes place at a constant temperature; during freezing, the latent heat of solidification is given off. (b) Change in density during cooling of pure metals.

A **phase diagram**, also called an **equilibrium** or **constitutional diagram**, shows the relationships among temperature, composition, and phases present in a particular alloy system at equilibrium. *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 (1) complete solid solutions; (2) eutectics, such as cast irons; and (3) eutectoids, such as steels.

One example of a phase diagram is shown in Fig. 4.5 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-hand boundary of this phase diagram (100% Ni) indicates the melting point of nickel; the right-hand boundary (100% Cu) indicates the melting point of copper. (All percentages in this discussion are by weight, not by number of atoms.)

Note that, for a composition of, say, 50% Cu–50% Ni, the alloy begins to solidify at a temperature of 1313°C, and solidification is complete at 1249°C. Above 1313°C, a homogeneous liquid of 50% Cu–50% Ni exists. When cooled slowly to 1249°C, a homogeneous solid solution of 50% Cu–50% Ni results. However, between the liquidus and solidus curves, and at a temperature of 1288°C, there is a two-phase region: a **solid phase** composed of 42% Cu–58% Ni, and a **liquid phase** of 58% Cu–42% Ni. To determine the *solid composition*, we go left horizontally to the solidus curve and read down, obtaining 42% Cu. The *liquid composition* (58%) is obtained similarly, by going to the right to the liquidus curve. The procedure for determining the compositions of various phases in phase diagrams (called the **lever rule**) is described in detail in texts on materials science and metallurgy.

The *completely solidified* alloy in the phase diagram shown in Fig. 4.5 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

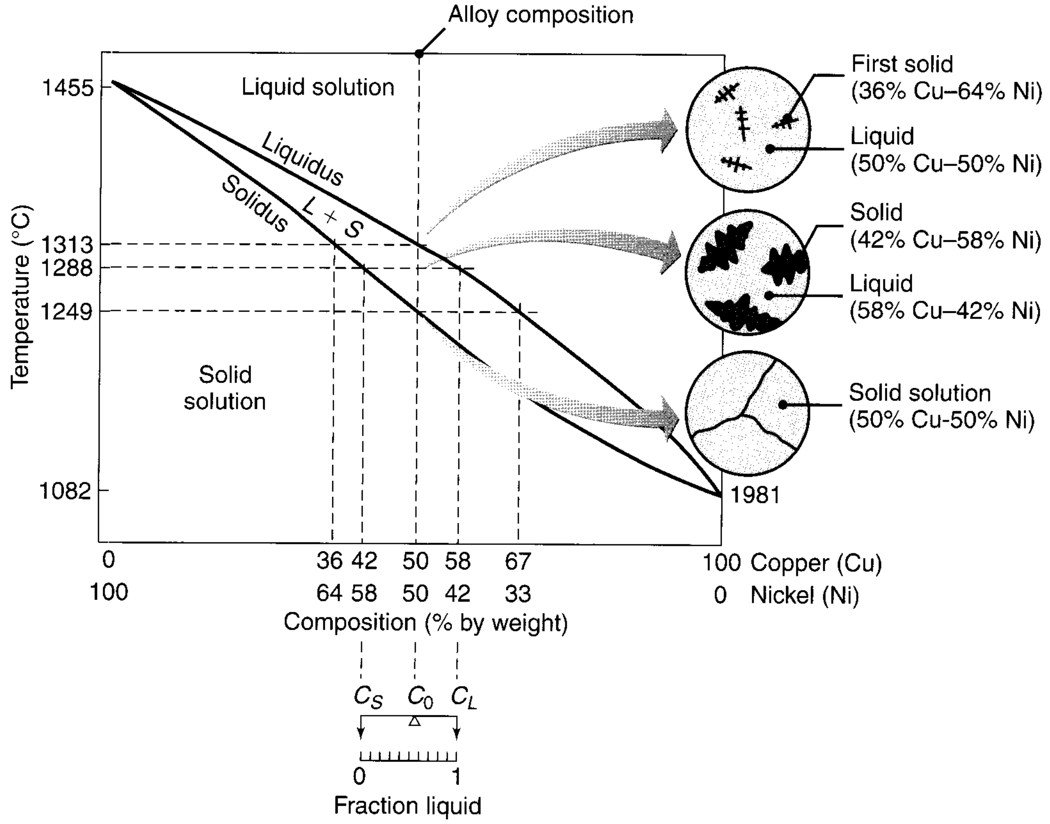


FIGURE 4.5 Phase diagram for nickel–copper alloy system obtained at a slow rate of solidification. Note that pure nickel and pure copper each have one freezing or melting temperature. The top circle on the right depicts the nucleation of crystals. The second circle shows the formation of dendrites (see Section 10.2). The bottom circle shows the solidified alloy, with grain boundaries.

composition. The atomic radius of copper is 0.128 nm and that of nickel is 0.125 nm, and both elements are of face-centered cubic structure; thus, the Hume-Rothery rules are obeyed.

The mechanical properties of solid solutions of Cu–Ni depend on their composition (Fig. 4.6). Up to a point, the properties of pure copper are improved upon by increasing the nickel content: There is an optimal percentage of nickel that gives the highest strength and hardness to the Cu–Ni alloy. Figure 4.6 also 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 at substitutional nickel or zinc atoms, which may also be regarded as impurity atoms; as a result, dislocations cannot move as freely and the strength of the alloy increases.

Another example of a two-phase diagram is shown in Fig. 4.7, for the lead–tin system. The single phases alpha and beta are solid solutions. Note that the single-phase regions are separated from the liquid phase by two two-phase regions: *alpha + liquid* and *beta + liquid*. Figure 4.7 shows the composition of the alloy (61.9% Sn–38.1% Pb) that has the *lowest* temperature at which the alloy is still completely

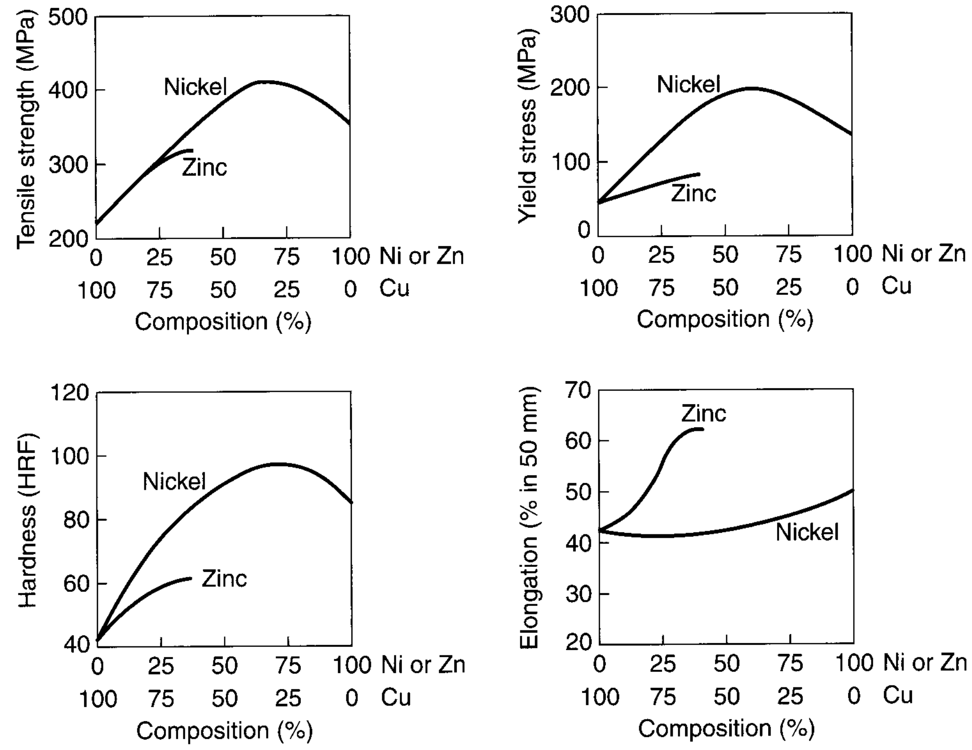


FIGURE 4.6 Mechanical properties of copper–nickel and copper–zinc alloys as a function of their composition. The curves for zinc are short, because zinc has a maximum solid solubility of 40% in copper.

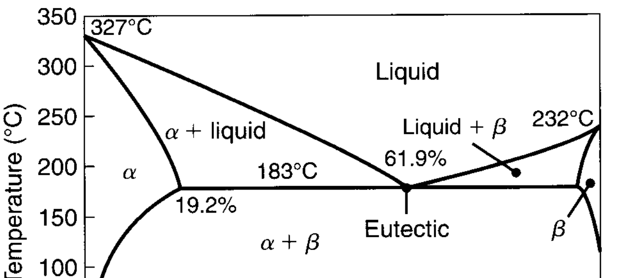


FIGURE 4.7 The lead–tin phase diagram. Note that the composition of the eutectic point for this alloy is 61.9% Sn–38.1% Pb. A composition either lower or higher than this ratio will have a higher liquidus temperature.

liquid, namely, 183°C. This point is known as the **eutectic point**, and it is the point at which the liquid solution decomposes into the components alpha and beta. The word *eutectic* is from the Greek *eutektos*, meaning "easily melted."

Eutectic points are important in applications such as soldering (Section 32.3), where low temperatures are desirable to prevent thermal damage to parts during joining. Although there are various types of solders, tin–lead solders are commonly used for general applications; they have compositions ranging from 5% Pb–95% Sn to 70% Pb–30% Sn. Each composition has its own melting point.

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. In this section the iron–carbon system is described, 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. 4.8. Although this diagram can be extended to the right—to 100% C (pure graphite); see Fig. 4.12—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. 4.8. 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 iron; it has a maximum solid solubility of 0.022% C at a temperature of 727°C. Delta ferrite (δ-ferrite) is another form that is stable only at very high temperatures and is of no practical significance in engineering. 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), so 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 desirable properties.

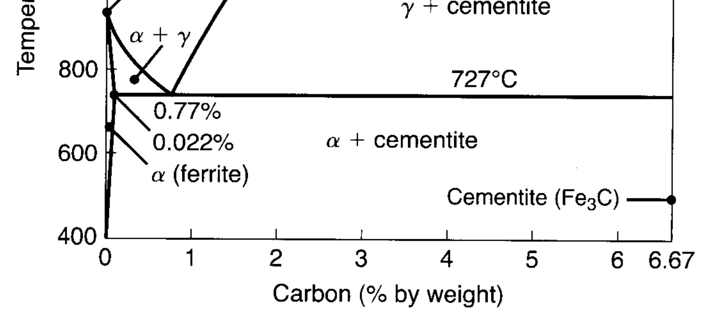
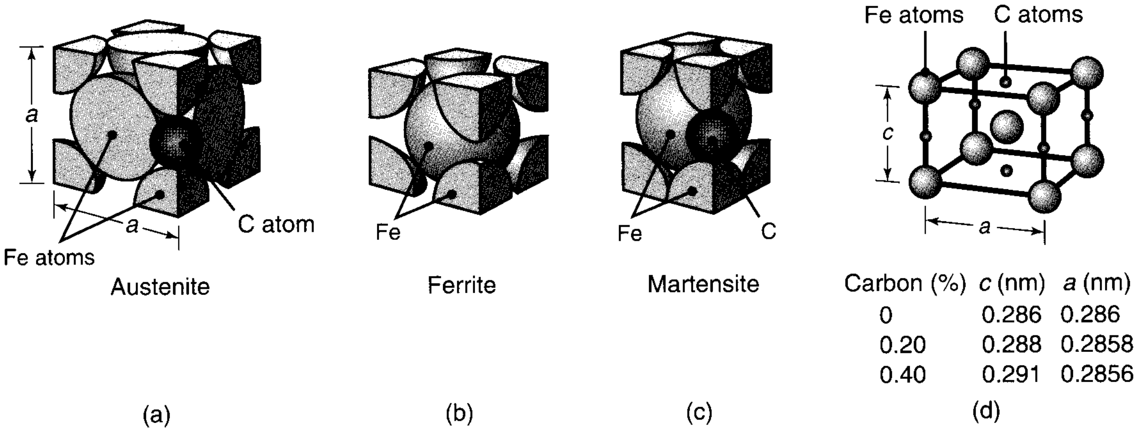


FIGURE 4.8 The iron–iron-carbide phase diagram. Because of the importance of steel as an engineering material, this diagram is one of the most important of all phase diagrams.



Carbon (%)	c (nm)	a (nm)
0	0.286	0.286
0.20	0.288	0.2858
0.40	0.291	0.2856

FIGURE 4.9 The unit cells for (a) austenite, (b) ferrite, and (c) martensite. The effect of percentage of carbon (by weight) on the lattice dimensions for martensite is shown in (d). Note the interstitial position of the carbon atoms. (See Fig. 1.7.) 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. 4.8, 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 (Fig. 4.9a).

Austenite is an important phase in the heat treatment of steels (Section 4.7). It is denser than ferrite, and its single-phase fcc structure is ductile at elevated temperatures; consequently, it possesses good formability. Large amounts of nickel and manganese can also be dissolved in fcc iron to impart various properties. Steel is nonmagnetic in the austenitic form, either at high temperatures or, for austenitic stainless steels, at room temperature.

Cementite. The right boundary of Fig. 4.8 represents **cementite**, which is 100% iron carbide (Fe_3C), having a carbon content of 6.67%. Cementite, from the Latin *caementum* (meaning "stone chips"), is also called **carbide**. This iron carbide should not be confused with other carbides used as dies, cutting tools, and abrasives, such as tungsten carbide, titanium carbide, and silicon carbide, described in Chapters 8 and 22. Cementite is a very hard and brittle intermetallic compound and has a significant influence on the properties of steels. It can include other alloying elements, such as chromium, molybdenum, and manganese.

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. 4.10, which is an enlargement of the lower left-hand portion of Fig. 4.8. Various microstructures can be developed, depending on the carbon content, the amount of plastic deformation (working), and the method of heat treatment. For example, let's 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 *slow* cooling is to maintain equilibrium; higher rates of cooling are employed in heat treating, as explained in Section 4.7. 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. This reaction is called a **eutectoid** (meaning *eutecticlike*) **reaction**, which means 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. 4.11). 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).

In iron with less than 0.77% C, the microstructure formed consists of a pearlite phase (ferrite and cementite) and a ferrite phase. The ferrite in the pearlite is called **eutectoid ferrite**, and the ferrite phase is called **proeutectoid ferrite** (*pro* meaning before). Poreutectoid ferrite forms at a temperature higher than the eutectoid temperature of 727°C in the *alpha* + *gamma* region. If the carbon content is higher than 0.77%, the austenite transforms into pearlite and cementite. The cementite in the pearlite is called **eutectoid cementite**, and the cementite phase is called **proeutectoid cementite** because it forms in the alpha + Fe_3C region, at a temperature higher than the eutectoid temperature.

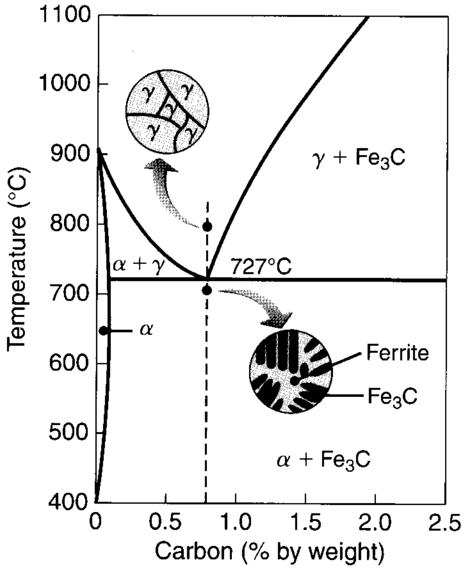


FIGURE 4.10 Schematic illustration of the microstructures for an iron-carbon alloy of eutectoid composition (0.77% carbon), above and below the eutectoid temperature of 727°C (1341°F).

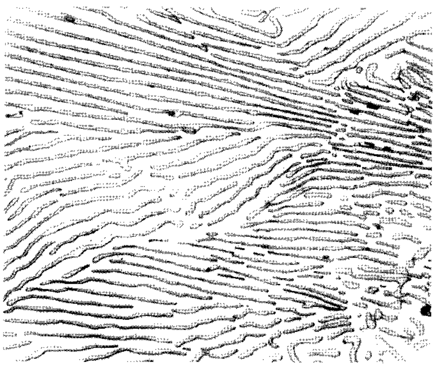


FIGURE 4.11 Microstructure of pearlite in 1080 steel, formed from austenite of eutectoid composition. In this lamellar structure, the lighter regions are ferrite and the darker regions are carbide. Magnification: 2500×.

4.5.1 Effects of Alloying Elements in Iron

Although carbon is the basic element that transforms iron into steel, other elements are also 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); they shift other phase boundaries as well.

The eutectoid temperature may be raised or lowered from 727°C, depending on the particular alloying element. On the other hand, 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 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 usually classified according to their solidification morphology from the eutectic temperature (see also Section 12.3):

- **a.** Gray cast iron, or gray iron;
- **b.** Ductile cast iron, nodular cast iron, or spheroidal graphite cast iron;
- **c.** White cast iron;
- **d.** Malleable iron;
- **e.** Compacted graphite iron.

Cast irons are also classified by their structure: ferritic, pearlitic, quenched and tempered, or austempered.

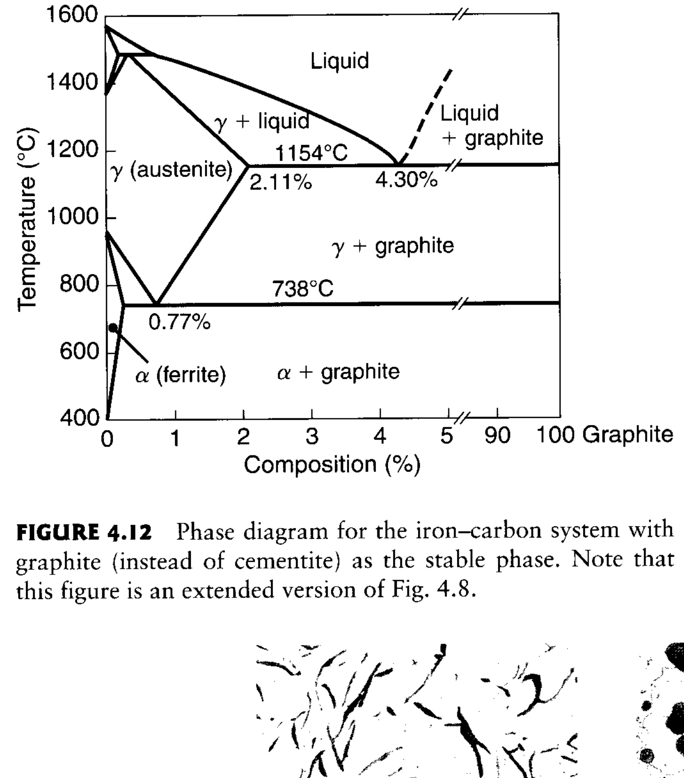


FIGURE 4.12 Phase diagram for the iron–carbon system with graphite (instead of cementite) as the stable phase. Note that this figure is an extended version of Fig. 4.8.

The equilibrium phase diagram relevant to cast irons is shown in Fig. 4.12, in which the right boundary is 100% C—that is, pure graphite. The eutectic temperature is 1154°C; thus, 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 not completely stable; it is **metastable**, 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.

Gray Cast Iron. In this structure, graphite exists largely in the form of *flakes* (Fig. 4.13a). It is called **gray cast iron**, or **gray iron**, because, when it is

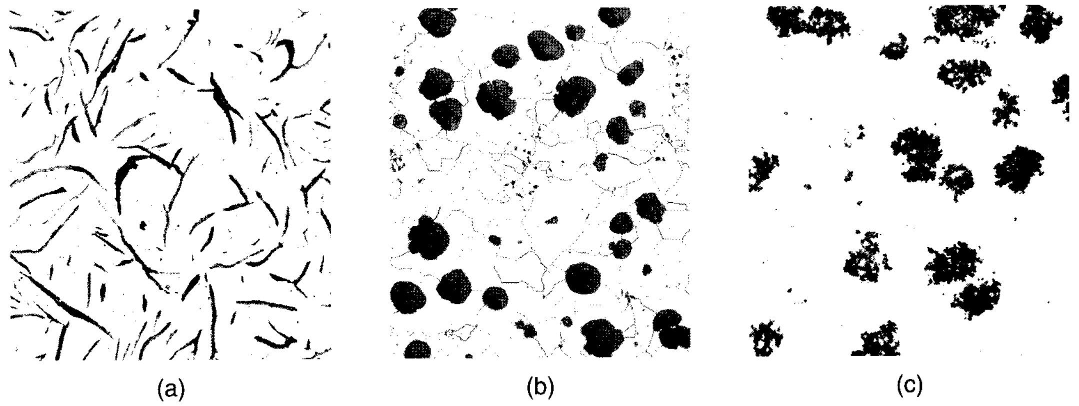


FIGURE 4.13 Microstructure for cast irons. Magnification: 100×. (a) Ferritic gray iron with graphite flakes. (b) Ferritic ductile iron (nodular iron), with graphite in nodular form. (c) Ferritic malleable iron; this cast iron solidified as white cast iron, with the carbon present as cementite, and was heat treated to graphitize the carbon.

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, as are other brittle materials. On the other hand, the presence of graphite flakes gives this material the capacity to dampen vibrations (by internal friction). This capacity makes gray cast iron 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 the 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; although still brittle, it is stronger than fully gray iron. Martensitic gray iron is obtained by austenitizing a pearlitic gray iron and then rapidly quenching it, to produce a structure of graphite in a martensite matrix; as a result, this cast iron is very hard.

Ductile (Nodular) Iron. In the ductile-iron structure, graphite is in a **nodular** or **spheroid** form (Fig. 4.13b); this shape permits the material to be somewhat ductile and shock resistant. The shape of graphite flakes is changed 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, alternatively, a structure of tempered martensite (Section 4.7).

White Cast Iron. The white cast iron structure is very hard, wear resistant, and brittle because of the presence of large amounts of iron carbide (instead of graphite). **White cast iron** is obtained either by cooling gray iron rapidly or by adjusting the composition by keeping the carbon and silicon content low. This type of cast iron is also called *white iron* because of the white crystalline appearance of the fracture surface.

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. 4.13c) in a ferrite or pearlite matrix; consequently, malleable iron has a structure similar to that of nodular iron. This structure promotes 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 having 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 various microstructures described thus far can be modified by **heat-treatment** techniques—that is, by controlled heating and cooling of the alloys at various rates. These treatments induce **phase transformations** that greatly influence such mechanical properties as the strength, hardness, ductility, toughness, and wear resistance of the alloys. The 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. The processes of recovery, recrystallization, and grain growth (Section 1.7) are examples of thermal treatment, involving changes in the grain structure of the alloy.

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. 4.11 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 it is in air), fine pearlite is produced; if cooling is slow (as it is 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 (**subcritical annealing**, Section 4.11), such as for a day at 700°C, the cementite lamellae transform to roughly *spherical* shapes (Fig. 4.14). Unlike the lamellar shapes of cementite, which act as stress raisers, **spheroidites** (spherical particles) are less conducive to stress concentration because of their rounded shapes. Consequently, this structure has higher toughness and lower hardness than the pearlite structure. In this form, it can be cold worked, because the ductile ferrite has high toughness and the spheroidal carbide particles prevent the initiation of cracks within the material.

Bainite. Visible only through electron microscopy, **bainite** is a very fine microstructure consisting of ferrite and cementite, similar to pearlite, but having a different morphology. Bainite can be produced in steels with alloying elements and at cooling rates that are higher than those required for transformation to pearlite. This structure, called **bainitic 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 it 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. 4.9d). 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. 4.15); it lacks toughness, and therefore has limited use. Martensite transformation takes place almost instantaneously because it involves, not the diffusion process, but a slip mechanism (and thus plastic deformation), which is a time-dependent phenomenon that is the mechanism in other transformations as well.

The material undergoes volume changes because of the differences in densities that result from phase transformations. For example, when austenite transforms to martensite, its volume increases (and hence its density decreases) by as much as 4%. A similar, but smaller, volume expansion also occurs when austenite transforms to pearlite. These expansions, and the thermal gradients present in a quenched part, cause internal stresses within the body. They may cause parts to undergo *distortion* or even to crack during heat treatment; **quench cracking** of steels is caused by rapid cooling during quenching.

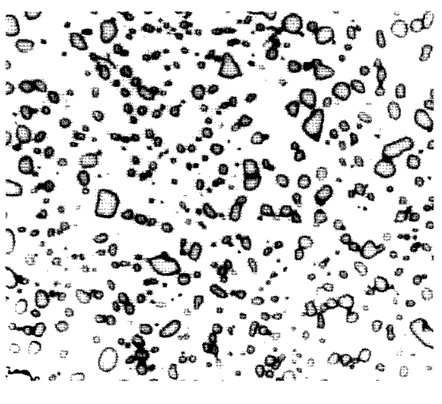


FIGURE 4.14 Microstructure of eutectoid steel. Spheroidite is formed by tempering the steel at 700°C. Magnification: 1000×.

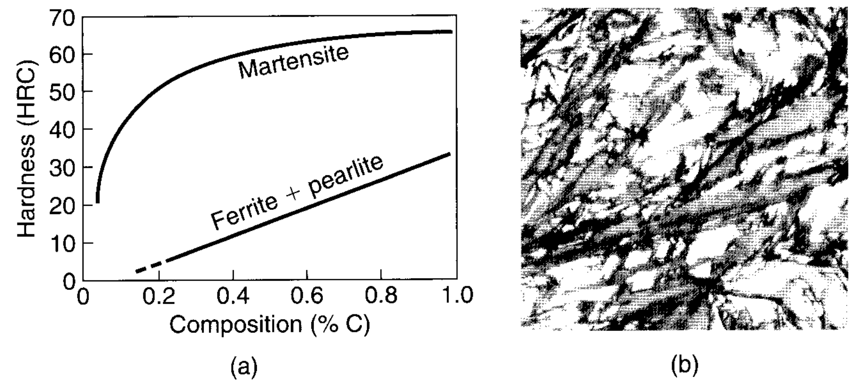


FIGURE 4.15 (a) Hardness of martensite as a function of carbon content. (b) Micrograph of martensite containing 0.8% carbon. The gray platelike regions are martensite; they have the same composition as the original austenite (white regions). Magnification: 1000×.

Distortion is an irreversible dimensional change of the part during heat treatment. It is a general phenomenon and may consist of size distortion or shape distortion. *Size distortion* involves changes in the dimensions of the part without a change in shape, whereas *shape distortion* involves bending, twisting, and similar nonsymmetrical dimensional changes. Distortion can be reduced by proper control of heating and cooling cycles, by improved part design, and by more localized heat treatment of the part (see also Section 4.13).

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 it lowers the hardness and strength of the alloy.

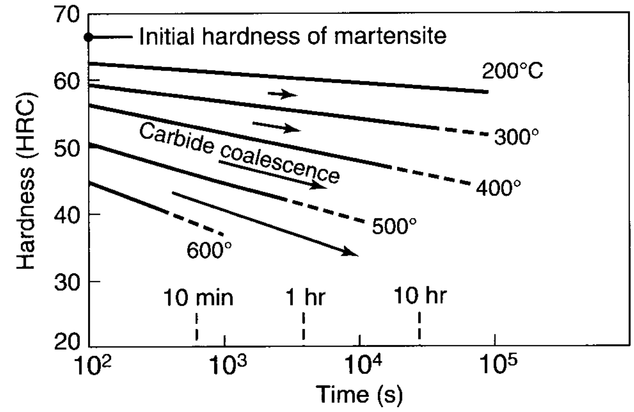


FIGURE 4.16 Hardness of tempered martensite as a function of tempering time for 1080 steel quenched to 65 HRC. Hardness decreases because the carbide particles coalesce and grow in size, thereby increasing the interparticle distance of the softer ferrite.

Tempered Martensite. Martensite is tempered in order to improve its mechanical properties. **Tempering** is a heating process by which hardness is reduced and toughness is improved. The body-centered tetragonal martensite is heated to an intermediate temperature, typically 150°–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. 4.16). 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

Figure 4.17a shows the percentage of austenite transformed into pearlite as a function of temperature and time. The transformation from austenite to pearlite (among other structures) is best illustrated by Figs. 4.17b and c, in diagrams called **isothermal**

transformation (IT) diagrams, or *time–temperature-transformation* (TTT) *diagrams*, constructed from the data given in Fig. 4.17a. The higher the temperature or the longer the time, the more austenite that transforms 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. This transformation can be traced in Figs. 4.17b and c.

The TTT diagrams shown allow metallurgists to design heat treatment schedules to obtain desirable microstructures. For example, consider the TTT curves shown in Fig. 4.17c. The steel can be raised to a very high temperature (above the eutectic

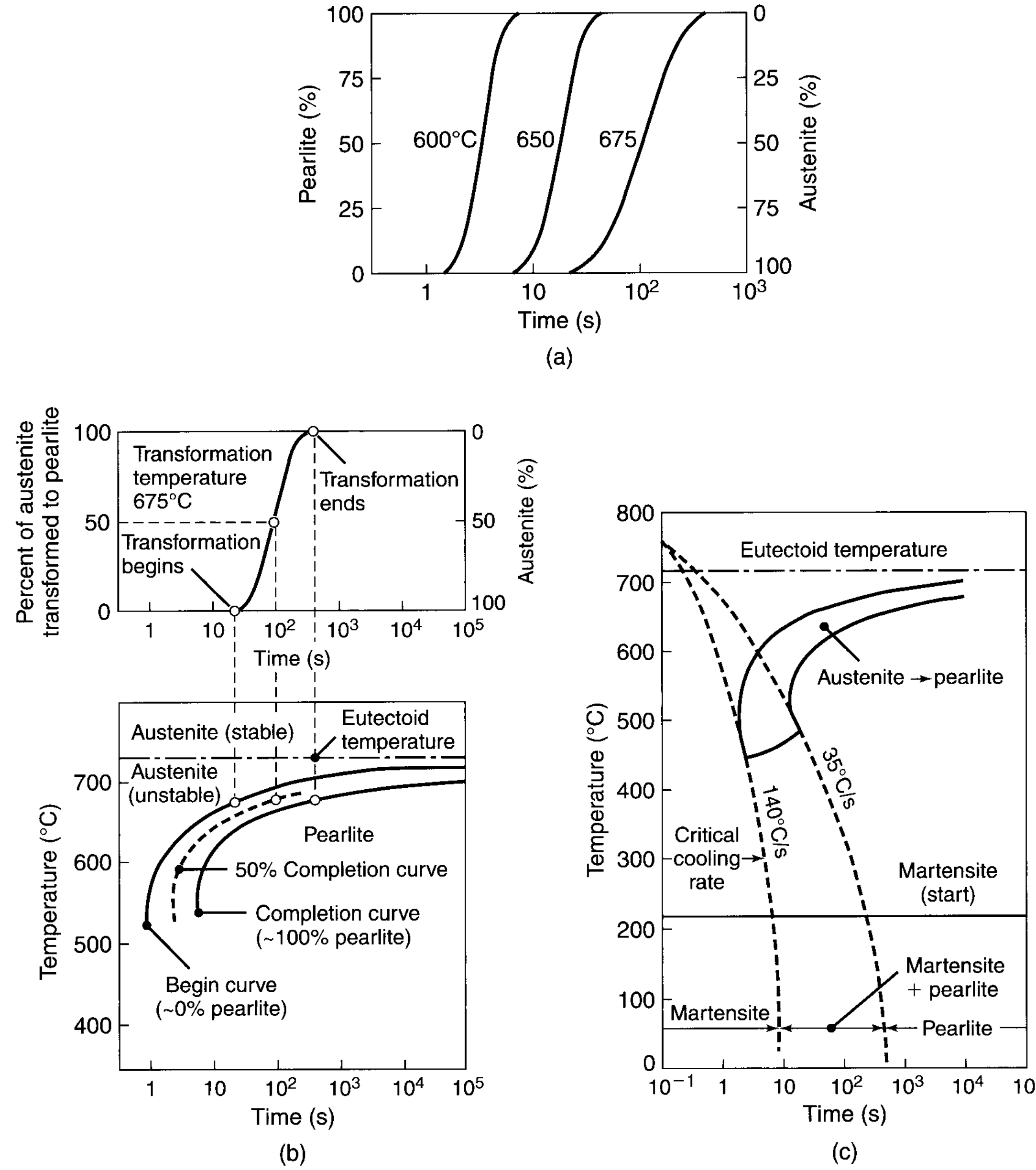


FIGURE 4.17 (a) Austenite-to-pearlite transformation of iron–carbon alloy as a function of time and temperature. (b) Isothermal transformation diagram obtained from (a) for a transformation temperature of 675°C. (c) Microstructures obtained for a eutectoid iron–carbon alloy as a function of cooling rate.

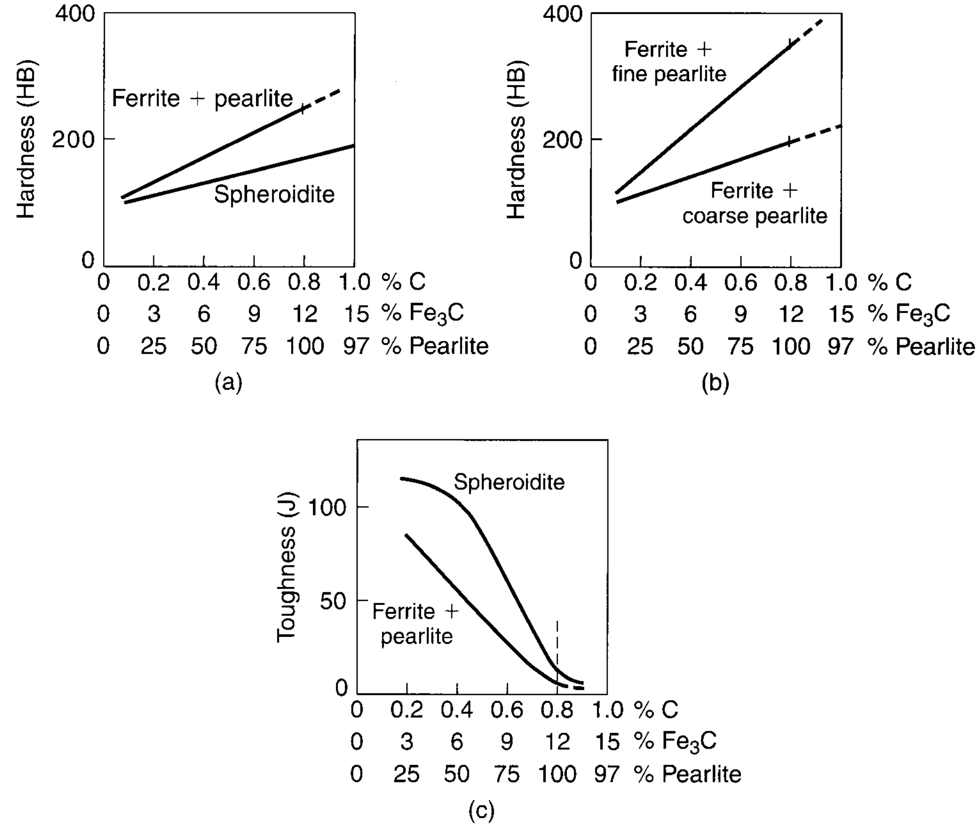


FIGURE 4.18 (a) and (b) Hardness and (c) toughness for annealed plain-carbon steels as a function of carbide shape. Carbides in the pearlite are lamellar. Fine pearlite is obtained by increasing the cooling rate. The spheroidite structure has spherelike carbide particles.

temperature) to start with a state of austenite. If the material is cooled very quickly, it can follow the 140°C/s cooling rate trajectory shown that results 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. If tempered martensite is desired, the heat treat and quench stages will be followed by a tempering process.

The differences in hardness and toughness of the various structures obtained are shown in Fig. 4.18. 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. 4.19.

4.8 Hardenability of Ferrous Alloys

The capability of an alloy to be hardened by heat treatment is called its **hardenability**. It is a measure of the *depth* of hardness that can be obtained by heating and subsequent quenching. The term "hardenability" should not be confused with "hardness," which is the resistance of a material to indentation or scratching (Section 2.6). From the discussion thus far, it can be seen that hardenability of ferrous alloys depends on the carbon content, the grain size of the austenite, the alloying elements present in the material, and the cooling rate.

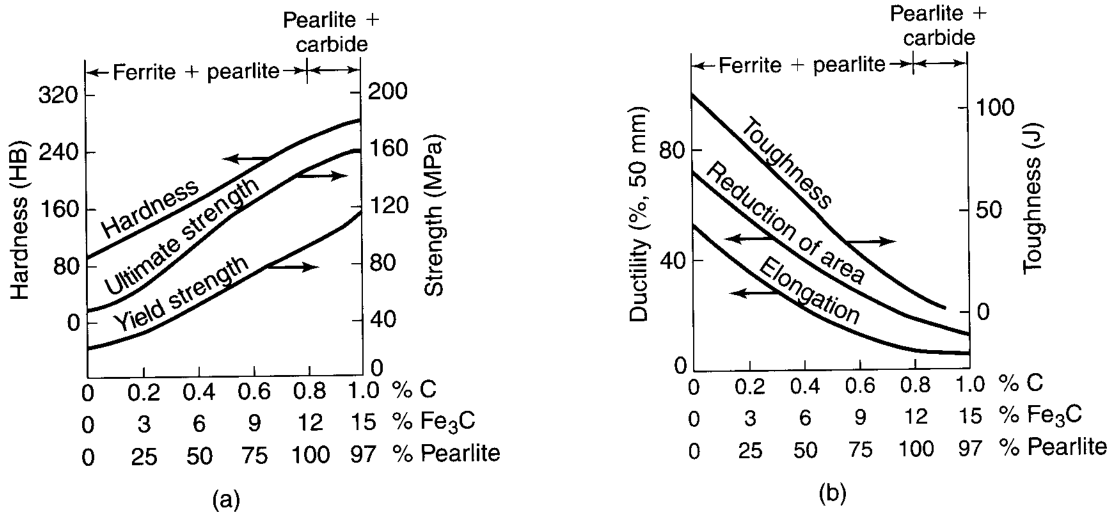


FIGURE 4.19 Mechanical properties of annealed steels as a function of composition and microstructure. Note in (a) the increase in hardness and strength, and in (b), the decrease in ductility and toughness, with increasing amounts of pearlite and iron carbide.

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. 4.20a) 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, which is 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 discussion of the effects of cooling rates in Section 4.7, hardness decreases away from the quenched end of the bar (Fig. 4.20b). The greater the depth to which the hardness penetrates, the greater 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; note also 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.

Because small variations in composition and in grain size can affect the shape of hardenability curves, each lot of an alloy should be tested individually. The data may be plotted as a band, rather than as a single curve. Hardenability curves are essential in predicting the hardness of heat-treated parts (such as gears, cams, and various other components) as a function of their composition.

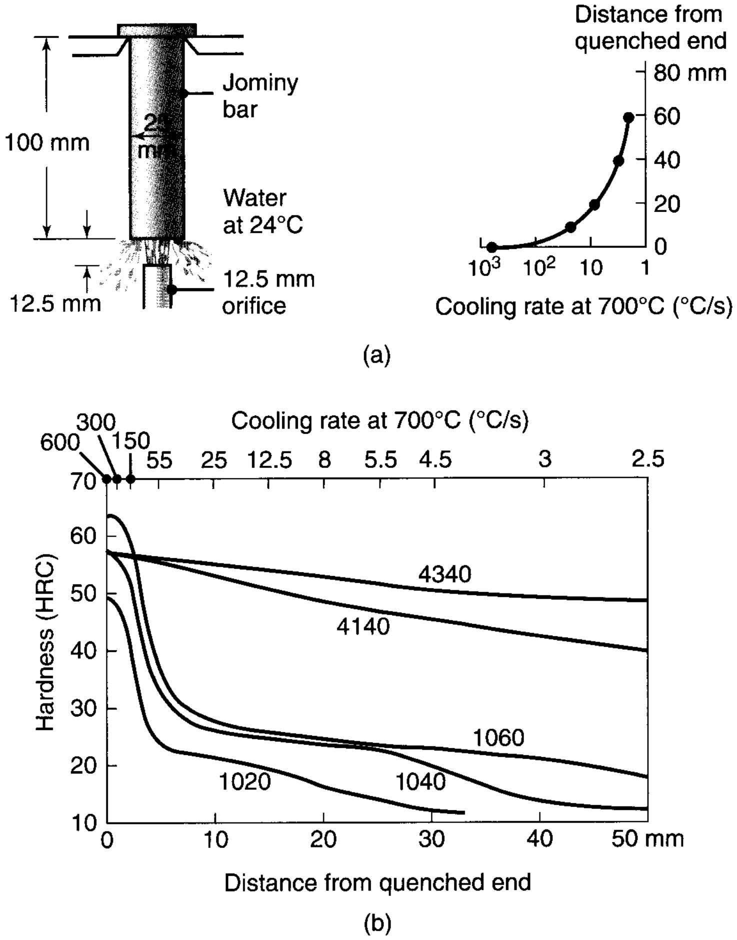
4.8.2 Quenching Media

The fluid used for quenching the heated alloy also has an effect on hardenability. Quenching may be carried out in water, brine (salt water), oils, molten salts, or air; caustic solutions, polymer solutions, and gases are also used. Because of the differences in thermal conductivities, specific heats, and heats of vaporization of these media, the rate of cooling of the alloy (**severity of quench**) is also different. In relative terms and in decreasing order, the cooling capacities of several quenching media are as follows: agitated brine, 5; still water, 1; still oil, 0.3; cold gas, 0.1; still air, 0.02.

Agitation is also 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 5.6), 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.

Water is a common medium for rapid cooling. However, the heated metal may form a **vapor blanket** along its surfaces due to the water-vapor bubbles that form when water boils at the metal–water interface. 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 high pressure. Brine is an effective quenching medium, because salt helps to nucleate bubbles at the interfaces, which improves agitation; however, brine can corrode the part.

Polymer quenchants can be used for ferrous as well as for nonferrous alloy quenching, and new quenchants are developed regularly. They 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 results, 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). The quenching rate can be controlled by varying the concentration of the solutions.



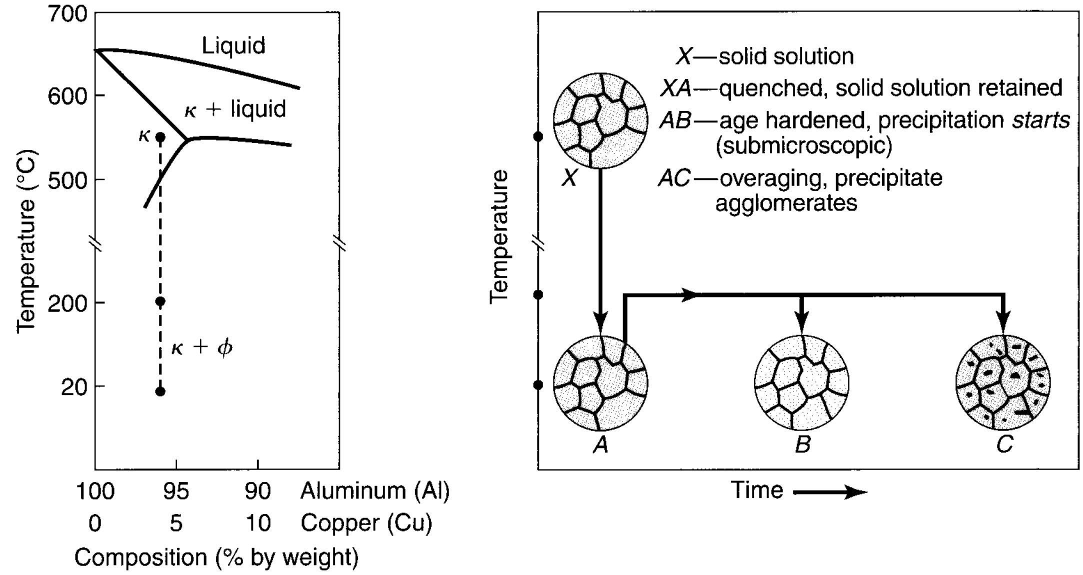
(a)

(b)

FIGURE 4.20 (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.

4.9 Heat Treatment of Nonferrous Alloys and Stainless Steels

Nonferrous alloys and some stainless steels cannot be heat treated by the techniques used for ferrous alloys. The reason is that nonferrous alloys do not undergo phase transformations like those in steels; the hardening and strengthening mechanisms for these alloys are fundamentally different. Heat-treatable aluminum alloys, copper alloys, martensitic stainless steels, and some other stainless steels are hardened and strengthened by a process called **precipitation hardening.** This heat treatment is a technique in which small particles of a different phase, called **precipitates,** are uniformly dispersed in the matrix of the original phase (Fig. 4.3a). In this process,



(a) (b)

FIGURE 4.21 (a) Phase diagram for the aluminum–copper alloy system. (b) Various microstructures obtained during the age-hardening process.

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; they can best be described by reference to the phase diagram for the aluminum–copper system (Fig. 4.21a). For an alloy with the composition 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. This *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 are 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—for instance, by quenching it in water. The structure obtained soon after quenching (*A* in Fig. 4.21b) consists only of the single phase kappa; this alloy has moderate strength and considerable ductility.

4.9.2 Precipitation Hardening

The structure obtained in *A* in Fig. 4.21b can be made stronger by **precipitation hardening.** In this process, the alloy is reheated to an intermediate temperature and then held there for a period of time, during which precipitation takes place. The copper atoms diffuse to nucleation sites and combine with aluminum atoms; this process produces the theta phase, which forms as submicroscopic precipitates (shown in *B* by the small dots within the grains of the kappa phase). This structure is stronger than that in *A*, although it is less ductile. The increase in strength is due to 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 at room temperature; this process is 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 gain strength and hardness by aging naturally. Natural aging can be slowed by refrigerating the quenched alloy (**cryogenic treatment**).

In the precipitation process, if the reheated alloy is held at the elevated temperature for an extended period of time, the precipitates begin to coalesce and grow. They become larger, but fewer, as is shown by the larger dots in C in Fig. 4.21b; this process is called **overaging**, and the resulting alloy is softer and weaker.

There is an optimal time–temperature relationship in the aging process that must be observed in order to obtain desired properties (Fig. 4.22). Obviously, an aged alloy can be used only up to a certain maximum temperature in service; otherwise, it will over-age and so lose its strength and hardness. Although weaker, an over-aged part has better dimensional stability.

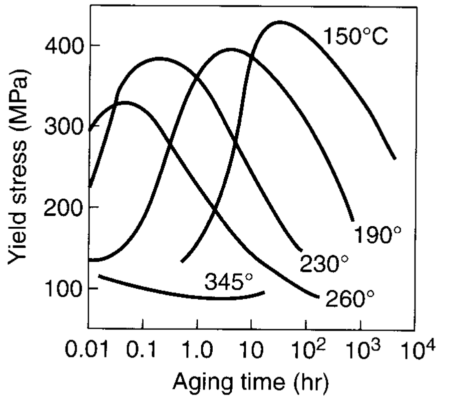


FIGURE 4.22 The effect of aging time and temperature on the yield stress of 2014-T4 aluminum alloy. Note that, for each temperature, there is an optimal aging time for maximum strength.

Maraging. This is a precipitation-hardening treatment for a special group of high-strength iron-base alloys. The word *maraging* is derived from *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 is done at 480°C. Hardening by maraging does not depend on the cooling rate; consequently, uniform and full hardness can be obtained throughout large parts with minimal distortion. Typical uses of maraging steels are in dies and tooling for casting, molding, forging, and extrusion (Parts II and III).

4.10 Case Hardening

The heat-treatment processes described thus far involve microstructural alterations and property changes in the *bulk* of the material or component by means of *through hardening*. It is not desirable to through harden parts, because a hard part lacks the necessary toughness for these applications; a small surface crack could propagate rapidly through such a part and cause total failure. In many cases, however, alteration of only the *surface properties* of a part (hence, the term surface or *case hardening*) is desirable. This method is particularly useful for improving resistance to surface indentation, fatigue, and wear. Typical applications for case hardening are gear teeth, cams, shafts, bearings, fasteners, pins, automotive clutch plates, tools, and dies.

Several case-hardening processes are available (Table 4.1):

a. **Carburizing** (gas, liquid, and pack carburizing);

b. **Carbonitriding**;

c. **Cyaniding**;

d. **Nitriding**;

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
Carbonitriding	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 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

e. **Boronizing;**
f. **Flame hardening;**
g. **Induction hardening;** and
h. **Laser-beam hardening.**

Basically, these are operations in which the component is heated in an atmosphere containing elements (such as carbon, nitrogen, or boron) that alter the composition, microstructure, and properties of surfaces. For steels with sufficiently high carbon content, surface hardening takes place without using any of these additional elements. Only the heat-treatment processes described in Section 4.7 are needed to alter 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 both small and large surfaces, such as gears, valves, punches, and locomotive cylinders. 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 distortion, and the ability to reach areas that would be inaccessible by other means. Capital costs can be high, however, and the depth of the case-hardened layer is usually less than 2.5 mm.

Because case hardening is a localized heat treatment, case-hardened parts have a hardness gradient. Typically, the hardness is a maximum at the surface and decreases below the surface, with a rate of decrease that depends on the composition of the metal and the process variables. Surface-hardening techniques can also be used for *tempering* (Section 4.11), to modify the properties of surfaces that have been subjected to heat treatment. Various other processes and techniques for surface hardening, such as shot peening and surface rolling, improve wear resistance and other characteristics (Section 34.2).

Decarburization is the phenomenon in which alloys containing carbon lose carbon from their surfaces as a result of heat treatment or of hot working in a medium, usually oxygen that reacts with the carbon. Decarburization is undesirable because it affects the hardenability of the surfaces of the part (by lowering its carbon content). It also adversely affects the hardness, strength, and fatigue life of steels, by significantly lowering their endurance limit. Decarburization is best avoided by processing the parts in an inert atmosphere or a vacuum, or by using neutral salt baths during heat treatment.

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 strength, or to modify the microstructure of the 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 glasses and similar products, castings, and weldments.

The annealing process consists of the following steps:

1. Heating the workpiece to a specific range of temperature in a furnace;
2. Holding it at that temperature for a period of time (*soaking*); and
3. Cooling the workpiece, in air or in a furnace.

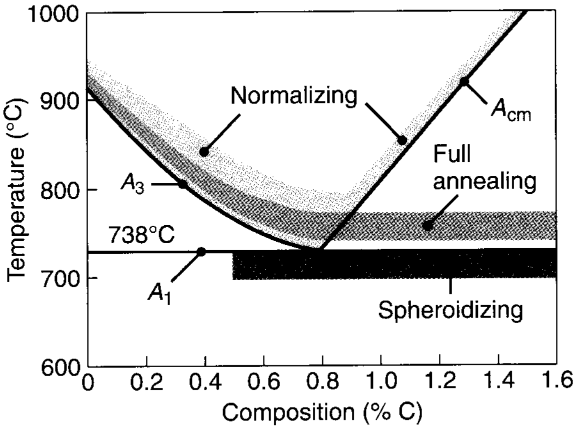


FIGURE 4.23 Heat-treating temperature ranges for plain-carbon steels, as indicated on the iron–iron carbide phase diagram.

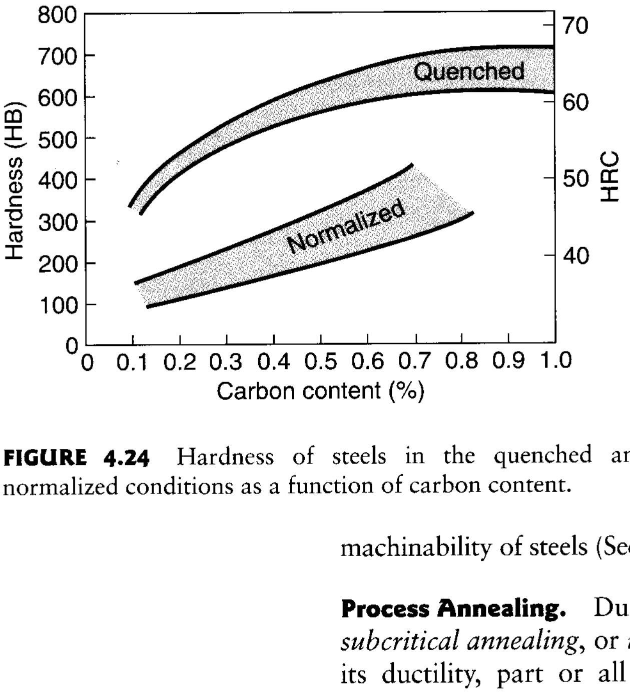


FIGURE 4.24 Hardness of steels in the quenched and normalized conditions as a function of carbon content.

The annealing process may be carried out in an inert or a controlled atmosphere, or it may be performed at lower temperatures to minimize or prevent surface oxidation.

An *annealing temperature* may be higher than the material'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 annealing temperature needed to fully recover the original properties ranges from 260° to 650°C, depending on the degree of prior cold work (see also Section 1.6). **Full annealing** is a term applied to the annealing of ferrous alloys. The steel is heated to above A_1 or A_3 (Fig. 4.23), and the cooling takes place slowly [typically at 10°C per hour], in a furnace, after which it is turned off. The structure obtained through full annealing is coarse pearlite, which is soft and ductile and has small, uniform grains.

To avoid excessive softness from the annealing of steels, the cooling cycle may be done completely in still air. This process is called **normalizing**, to indicate that the part is heated to a temperature above A_3 or A_{cm} in order to transform the structure to austenite. Normalizing results in somewhat higher strength and hardness, and in lower ductility, than does full annealing (Fig. 4.24). The structure obtained 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 it were described in Section 4.7 and shown in Figs. 4.14 and 4.23. *Spheroidizing annealing* improves the cold workability (Section 4.7) and the machinability of steels (Section 21.7).

Process Annealing. During *process annealing* (also called *intermediate annealing, subcritical annealing*, or *in-process annealing*), the workpiece is annealed to restore its ductility, part or all of which may have been exhausted by cold working. Afterwards, the part can be worked further into the final desired shape. If the temperature is high and/or the time of annealing is long, grain growth may result (Section 1.6), with adverse effects on the formability of the annealed parts.

Stress-relief Annealing. To reduce or eliminate residual stresses, a workpiece is generally subjected to *stress-relief annealing*, or simply **stress relieving**. The temperature and time required for this process depend on the material and on the magnitude of the residual stresses present. The 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.

For steels, the part is not heated to as high as A_1 in Fig. 4.23, in order to avoid phase transformations; slow cooling, such as occurs in still air, is generally employed. Stress relieving promotes dimensional stability in situations where subsequent relaxing of residual stresses may cause distortion of the part when it is in service over a period of time. It also reduces the tendency toward stress-corrosion cracking (Sections 2.10 and 3.8).

Tempering. If steels are hardened by heat treatment, then *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. 4.25. Alloy steels may undergo **temper embrittlement**, which is caused by the segregation of impurities along the grain boundaries at temperatures between 480° and 590°C.

Austempering. In *austempering*, the heated steel is quenched from the austenitizing temperature rapidly, to avoid formation of ferrite or pearlite. It is held at a certain temperature until isothermal 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.

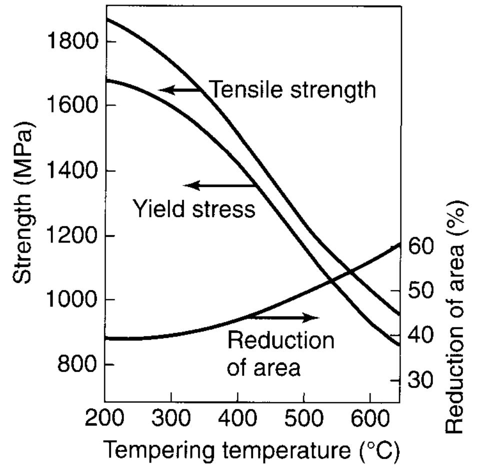


FIGURE 4.25 Mechanical properties of oil-quenched 4340 steel as a function of tempering temperature.

Austempering is often substituted for conventional quenching and tempering, either to reduce the tendency toward cracking and distortion during quenching or to improve ductility and toughness while maintaining hardness. Because of the shorter cycle time involved, this process also is economical for many applications. In **modified austempering**, a mixed structure of pearlite and bainite is obtained. The best example of this practice is **patenting**, which provides high ductility and moderately high strength, such as in patented wire (Section 15.8).

Martempering (Marquenching). In *martempering*, the steel or cast iron is first quenched from the austenitizing temperature in a hot-fluid medium, such as hot oil or molten salt. Next, it is held at that temperature until the temperature is uniform throughout the part; then it is cooled at a moderate rate, such as in air in order to avoid excessive temperature gradients within the part. Usually, the part is then tempered, because the structure obtained is otherwise primarily untempered martensite and thus is not suitable for most applications. Martempered steels have less tendency to crack, distort, or develop residual stresses during heat treatment. In **modified martempering**, the quenching temperature is lower, and thus the cooling rate is higher. The process is suitable for steels with lower hardenability.

Ausforming. In *ausforming*, also called **thermomechanical processing**, the steel is formed into desired shapes within controlled ranges of temperature and time, to avoid formation of nonmartensitic transformation products; the part is then cooled at various rates to obtain the desired microstructures. Ausformed parts have superior mechanical properties.

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 energy, their insulation and efficiency are important design considerations, as are their initial cost, the personnel needed for their operation and 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. Heating-system fuels are usually natural gas, oil, or electricity (for resistance or induction heating). The type of fuel used affects the furnace's 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 it 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. Basically, the furnace consists of an insulated chamber, a heating system, and an access door or doors. Batch furnaces are of the following basic types:

a. A **box furnace** is a horizontal rectangular chamber with one or two access doors through which parts are loaded. This type of furnace is used commonly and is versatile, simple to construct and use, and available in several sizes. A variation is the **car-bottom furnace**. The parts to be heat treated, usually long or large, are loaded onto a flatcar, which then moves on rails into the furnace.

b. A **pit furnace** is a vertical pit below ground level into which the parts are lowered. This type of furnace is particularly suitable for long parts, such as rods, shafts, and tubing, because they can be suspended by one end and, consequently, are less likely to warp during processing than if positioned horizontally within a box furnace.

c. 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.

d. In an **elevator furnace**, the parts to be heat treated are loaded onto a car platform, rolled into position, and then raised into the furnace. This type of furnace saves space in the plant and can be especially suitable for metal alloys that have to be quenched rapidly, because a quenching tank can be placed directly under 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 that use trays, belts, chains, and other mechanisms. Continuous furnaces are suitable for high production runs and can be designed and programmed so that complete heat-treating cycles can be performed under tight control.

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 and wire. Heating rates are high because of the higher thermal conductivity of liquid salts compared with that of air or gases. Depending on the electrical conductivity of the salt, heating may be done externally (for nonconducting salts), or by immersed or submerged electrodes using low-voltage alternating current. Direct current cannot be used because it subjects the salt to electrolysis. Salt baths are available for a wide range of temperatures. Lead can also be used as the heating medium.

Fluidized Beds. Dry, fine, and loose solid particles, usually aluminum oxide, are heated and suspended in a chamber by an upward flow of hot gas at various speeds.

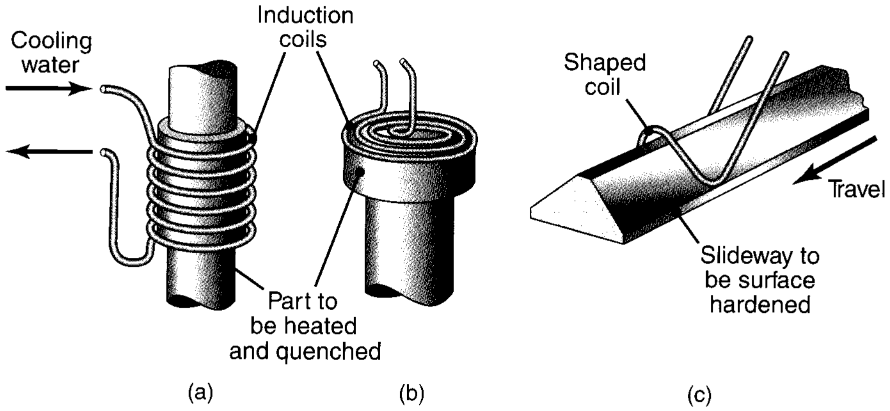


FIGURE 4.26 Types of coils used in induction heating of various surfaces of parts.

The parts to be heat treated are then placed within the floating particles—hence the term *fluidized bed.* Because of the constant agitation, the system is efficient, the temperature distribution is uniform, and the heat-transfer rate is high. These furnaces are used for various batch-type applications.

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. 4.26), is made of copper or of a copper-base alloy. The coil, which is usually water cooled, may be designed to quench the part as well, after heating it. Induction heating is desirable for localized heat treating, such as that required for gear teeth, cams, and similar parts.

Furnace Atmospheres. The atmospheres in furnaces can be controlled so as to avoid (or cause) oxidation, tarnishing, and decarburization of ferrous alloys heated to elevated temperatures. Oxygen causes oxidation (corrosion, rusting, and scaling). Carbon dioxide has various effects; it may be neutral or decarburizing, depending on its concentration in the furnace atmosphere. Water vapor in the furnace causes oxidation of steels, resulting in a blue color. Nitrogen is a common neutral atmosphere, and a vacuum provides a completely neutral atmosphere. The term **bluing** is used to describe formation of a thin blue film of oxide on finished parts to improve their appearance and their resistance to oxidation.

4.13 Design Considerations for Heat Treating

In addition to taking the metallurgical factors into account, successful heat treating involves design considerations for avoiding problems such as cracking, distortion, and nonuniformity of the properties throughout the heat-treated part. The rate of cooling during quenching may not be uniform, particularly in complex shapes having varying cross sections and thicknesses, and this nonuniformity may produce severe temperature gradients in the part. Nonuniformity can lead to variations in contraction, resulting in thermal stresses that may cause cracking of the part; furthermore, nonuniform cooling causes residual stresses in the part, which can lead to stress-corrosion cracking. The quenching method selected, the care taken in quenching, and the selection of a proper quenching medium and temperature are, consequently, important considerations.

As a general guideline for part design for heat treating, sharp internal or external corners should be avoided; otherwise, stress concentrations at these corners may raise the level of stresses high enough to cause cracking. The part should have its thicknesses as nearly uniform as possible; also, the transition between regions of different thicknesses should be made smooth. Parts with holes, grooves, keyways, splines, and asymmetrical shapes may likewise be difficult to heat treat, because they may crack during quenching. Large surfaces with thin cross sections are likely to warp. Hot forgings and hot steel-mill products may have a **decarburized skin** (a layer that has lost its carbon, Section 4.10); as a result, they may not respond successfully to heat treatment.

SUMMARY

- Commercially pure metals generally do not have sufficient strength for most engineering applications; consequently, they must be alloyed with various elements which 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 in order 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 the 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 involve heating the alloy and subsequently quenching it at varying cooling rates. As a result, important phase transformations take place, producing 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.
- The control of the furnace atmosphere, the quenchants used, the characteristics of the equipment, and the shape of the parts to be heat treated are important heat-treatment considerations. Hardenability is the capability of an alloy to be hardened by heat treatment. The Jominy end-quench hardenability 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, among them carburizing, nitriding, induction hardening, and laser-beam hardening.
- Annealing includes several alternative processes (normalizing, process annealing, stress relieving, tempering, austempering, and martempering), each having 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
Distortion
End-quench test
Equilibrium diagram
Eutectic point
Eutectoid reaction
Ferrite
Furnaces
Hardenability
Heat treatment
Intermetallic compounds
Iron–Carbon system
Jominy test
Maraging
Martempering
Martensite
Normalizing
Overaging
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, 1992.
ASM Handbook, Vol. 4: **Heat Treating.** ASM International, 1991.
ASM Handbook, Vol. 9: **Metallography and Microstructures,** ASM International, 2004.
Brooks, C.R., **Principles of the Heat Treatment of Plain Carbon and Low Alloy Steel,** ASM International, 1996.
Bryson, W.E., **Heat Treatment, Selection, and Application of Tool Steels,** 2nd ed., Hanser Gardner, 2005.
Heat Treater's Guide: Practices and Procedures for Irons and Steels. ASM International, 1995.
Heat Treater's Guide: Practices and Procedures for Nonferrous Alloys. ASM International, 1996.
Krauss, G., **Steels: Heat Treatment and Processing Principles.** ASM International, 1990.
Totten, G.E., **Steel Heat Treatment Handbook,** 2 Vols., 2nd ed., CRC Press/Taylor & Francis, 2007.

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. Explain what is meant by "second-phase particle."

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?

QUALITATIVE PROBLEMS

4.15. 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.

4.16. Describe the engineering significance of the existence of a eutectic point in phase diagrams.

4.17. What is the difference between hardness and hardenability?

4.18. Referring to Table 4.1, explain why the items listed under typical applications are suitable for surface hardening.

4.19. It generally is not desirable to use steels in their as-quenched condition. Explain why.

4.20. Describe the differences between case hardening and through hardening, insofar as engineering applications of metals are concerned.

4.21. Describe the characteristics of (a) an alloy, (b) pearlite, (c) austenite, (d) martensite, and (e) cementite.

4.22. Explain why carbon, among all elements, is so effective in imparting strength to iron in the form of steel.

4.23. How does the shape of graphite in cast iron affect its properties?

4.24. 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?

4.25. Why is it important to know the characteristics of heat-treating furnaces? Explain.

4.26. Explain why, in the abscissa of Fig. 4.18c, the percentage of pearlite begins to decrease after 0.8% carbon content is reached.

4.27. What is the significance of decarburization? Give some examples.

4.28. Explain your understanding of size distortion and shape distortion in heat-treated parts, and describe their causes.

4.29. Comment on your observations regarding Fig. 4.20b.

4.30. Design a heat-treating cycle for carbon steel, including temperature and exposure times, to produce (a) pearlite–martensite steels and (b) bainite–martensite steels.

QUANTITATIVE PROBLEMS

▶**4.31.** Using Fig. 4.5, estimate the following quantities for a 20% Cu–80% Ni alloy: (a) the liquidus temperature, (b) the solidus temperature, (c) the percentage of nickel in the liquid at 1400°C, (d) the major phase at 1400°C, and (e) the ratio of solid to liquid at 1400°C.

▶**4.32.** Extrapolating the curves in Fig. 4.16, estimate the time that it would take for 1080 steel to soften to 53 HRC at (a) 200°C and (b) 300°C.

▶**4.33.** 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

4.34. 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?

4.35. 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.

4.36. 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.

4.37. Figure 4.20b 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.

4.38. 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 at the end of this chapter make a survey of the heat-treating literature, and then compile several examples and illustrations of parts that have been heat treated.

4.39. Refer to Fig. 4.26, 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).

4.40. Inspect various parts in your car or home, and identify those that are likely to have been case hardened. Explain your reasons.

Ferrous Metals and Alloys: Production, General Properties, and Applications

CHAPTER 5

- Ferrous metals and alloys are the most widely used structural materials, generally because of their performance, wide availability, 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 which now have become 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, 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, as simple as a mixture of iron and carbon, but also 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 to 85% by weight of structural members and mechanical components. The average U.S. passenger vehicle (including trucks and sport utility vehicles) contains about 1200 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 modern technological developments. Primitive ferrous tools first appeared about 4000 to 3000 B.C. They 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 world, making up about 5% of the earth's crust (in the form of various ores). 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), which is then used to reduce iron oxide to iron. The chemical by-products of coke are used in the synthesis of plastics and of chemical compounds. The gases evolved during the conversion of coal to coke are used as fuel for plant operations.

The function of limestone (calcium carbonate) is to remove impurities from the 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 is 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 used later in making cement, fertilizers, glass, building materials, rock-wool insulation, and road ballast.

5.2.2 Ironmaking

The three raw materials described previously are 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 the height of about a 10-storey building. The charge mixture is melted in a 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 reactions may take place, the basic reaction is that of oxygen with carbon to produce carbon monoxide, which, in turn, reacts with the iron oxide and reduces it to **iron.**

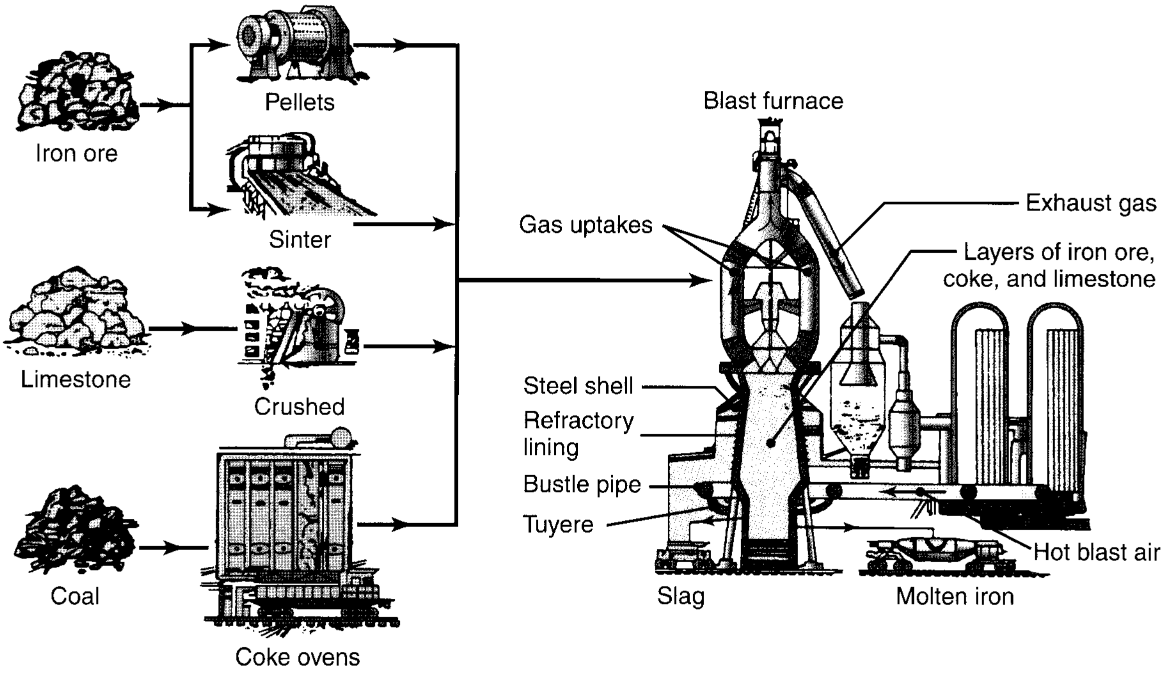


FIGURE 5.1 Schematic illustration of a blast furnace.

Preheating the incoming air is necessary because the burning coke alone does not produce sufficiently high temperatures for these reactions to occur.

The molten metal accumulates at the bottom of the blast furnace, while the impurities float to the top of the metal. 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**; it 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 (pig) is then 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 reducing the percentages of manganese, silicon, carbon, and other elements and by controlling the composition of the output through the addition of various elements. The molten metal from the blast furnace is transported into one of three types of furnaces: **open-hearth, electric,** or **basic-oxygen.** The name "open-hearth" derives from the shallow hearth shape that is open directly to the flames that melt the metal. Developed in the 1860s, the open-hearth furnace has been essentially replaced by electric furnaces and by the basic-oxygen process, because the latter two are more efficient and produce steels of better quality.

Electric Furnace. The source of heat in this furnace is a continuous electric arc that is formed between the electrodes and the charged metal (Figs. 5.2a and b). Temperatures as high as 1925°C are generated in this type of furnace. There are usu-

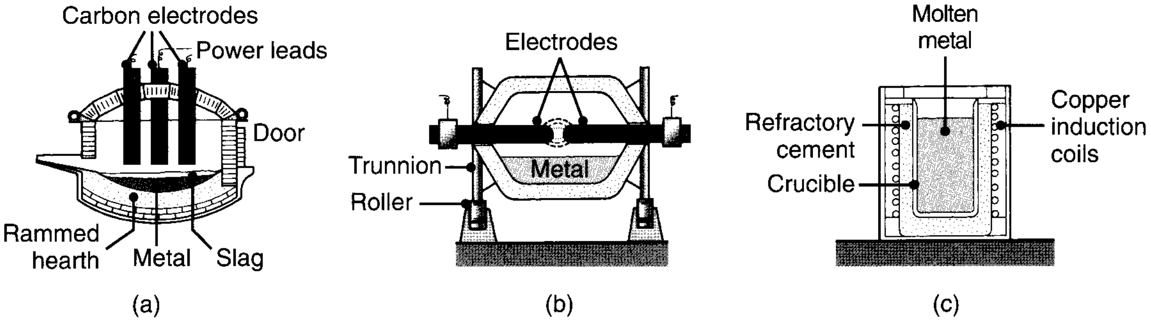


FIGURE 5.2 Schematic illustration of types of electric furnaces: (a) direct arc, (b) indirect arc, and (c) induction.

ally three graphite electrodes, and they can be as large as 750 mm in diameter and 1.5 to 2.5 m in length; their height in the furnace can be adjusted in response to the amount of metal present and the amount of wear of the electrodes.

Steel scrap and a small amount of carbon and limestone first are dropped into the electric furnace through the open roof. The roof then is closed and the electrodes are lowered. The power is then turned on, and within a period of about two hours, temperatures increase sufficiently to melt the metal. The current then is shut off, the electrodes are raised, the furnace is tilted, and the molten metal is poured into a *ladle*, which is a receptacle used for transferring and pouring molten metal. Electric-furnace capacities range from 55 to 82 metric tons of steel per day. The quality of steel produced is better than that from either the open-hearth or the basic-oxygen process.

For smaller quantities, electric furnaces (Fig. 5.2) can be of the **induction** type. 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. These furnaces also are used for remelting metal for casting.

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 82 metric tons of scrap are charged into a vessel (Fig. 5.3a); some units can hold as much as 350 tons. Pure oxygen is then blown into the furnace for about 20 minutes through a water-cooled *lance* (a long tube), under a pressure of about 1250 kPa, as shown in Fig. 5.3b. Fluxing agents (such as lime) are added through a chute. 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 230 metric tons of steel in 35 to 50 minutes. Most BOF steels, which have low impurity levels and are of better quality than open-hearth furnace steels are processed into plates, sheets, and various structural shapes, such as I-beams and channels (see Fig. 13.1).

Vacuum Furnace. Steel also may be melted in induction furnaces from which the air has been removed, similar to the one shown in Fig. 5.2c. Because the process removes gaseous impurities from the molten metal and prevents oxidation, vacuum furnaces produce high-quality steels.

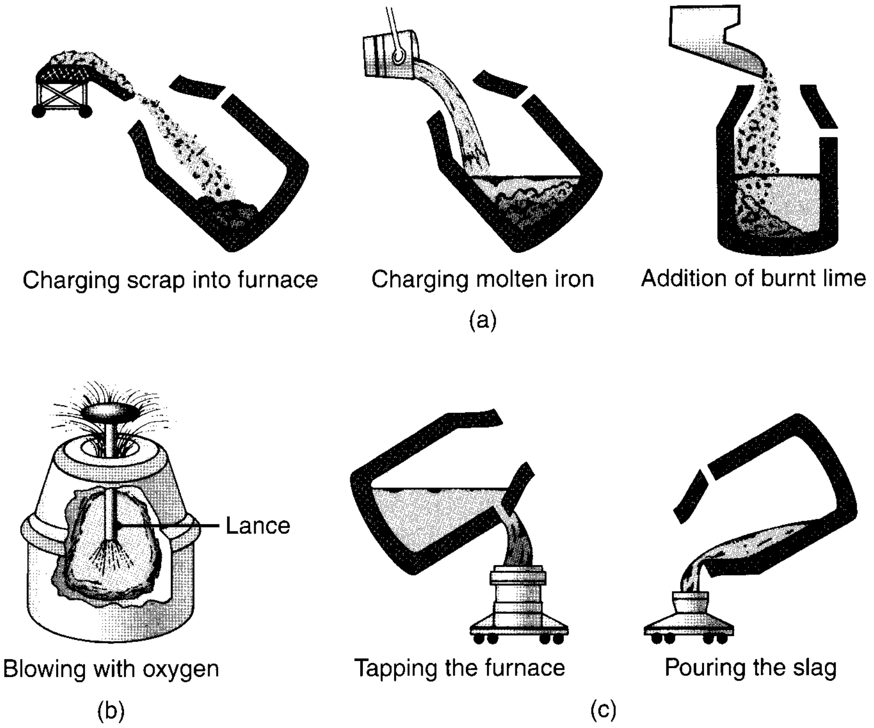


FIGURE 5.3 Schematic illustrations showing charging, melting, and pouring of molten iron in a basic-oxygen process.

5.3 Casting of Ingots

Traditionally, the next step in the steelmaking process is the shaping of the molten steel into a solid form (**ingot**) for such further processing as rolling it into shapes, casting it into semifinished forms, or forging it. 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. The bottoms of the molds may be closed or open; if they are open, the molds are placed on a flat surface. The cooled ingots are removed (stripped) from the molds and lowered into **soaking pits**, where they are reheated to a uniform temperature of about 1200°C for subsequent processing by rolling. Ingots may be square, rectangular, or round in cross section, and their weights range from a few hundred kilograms to 36 metric tons.

Certain reactions take place during the solidification of an ingot that have an important influence on 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. 10.15). Rejected oxygen combines with carbon to form 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: killed, semi-killed, and rimmed.

1. **Killed Steel.** Killed steel is a fully deoxidized steel; that is, oxygen is removed and the associated porosity is thus eliminated. In the deoxidation process, the dissolved oxygen in the molten metal is made to react with elements such as

aluminum, silicon, manganese, and vanadium that have been added to the melt. These elements have an affinity for oxygen and form metallic oxides. If aluminum is used, the product is called *aluminum-killed steel* (see Table 16.4.) The term *killed* comes from the fact that the steel lies quietly after being poured into the mold.

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 thus is 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. Because of shrinkage during the solidification, however, an ingot of this type develops a **pipe** at the top (also called a **shrinkage cavity**). It has the appearance of a funnel-like shape. This pipe can take up a substantial volume of the ingot, as it has to be cut off and scrapped.

2. **Semikilled 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. As a result, scrap is reduced. Although the piping in semi-killed steels is less, this advantage is offset by the presence of porosity in that region. Semi-killed steels are economical to produce.
3. **Rimmed Steel.** In a rimmed steel, which generally has a low carbon content (less than 0.15%), the evolved gases are only partially killed (or controlled) by the addition of 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 they are not controlled properly, blowholes may break through the skin. Furthermore, impurities and inclusions tend to segregate toward the center of the ingot. Thus, products made from this steel may be defective, and thorough inspection is essential.

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. There is an increasing demand for cleaner steels with improved and more uniform properties and a greater consistency of composition.

Refining is important particularly in producing high-grade steels and alloys for high-performance and critical applications, such as aircraft components. Moreover, warranty periods on shafts, camshafts, crankshafts, and similar parts can be increased significantly by using higher quality steels. Such steels are 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. Several processes using controlled atmospheres (such as electron-beam melting, vacuum-arc remelting, argon–oxygen decarburization, and vacuum-arc double-electrode remelting) have been developed.

5.4 Continuous Casting

The inefficiencies and the problems involved in making steels in the traditional form of ingots are alleviated by the *continuous-casting* process, which produces higher quality steels at reduced costs (see also Section 13.5.1 on *minimills*). Conceived in the 1860s, continuous or **strand casting** was first developed for casting nonferrous

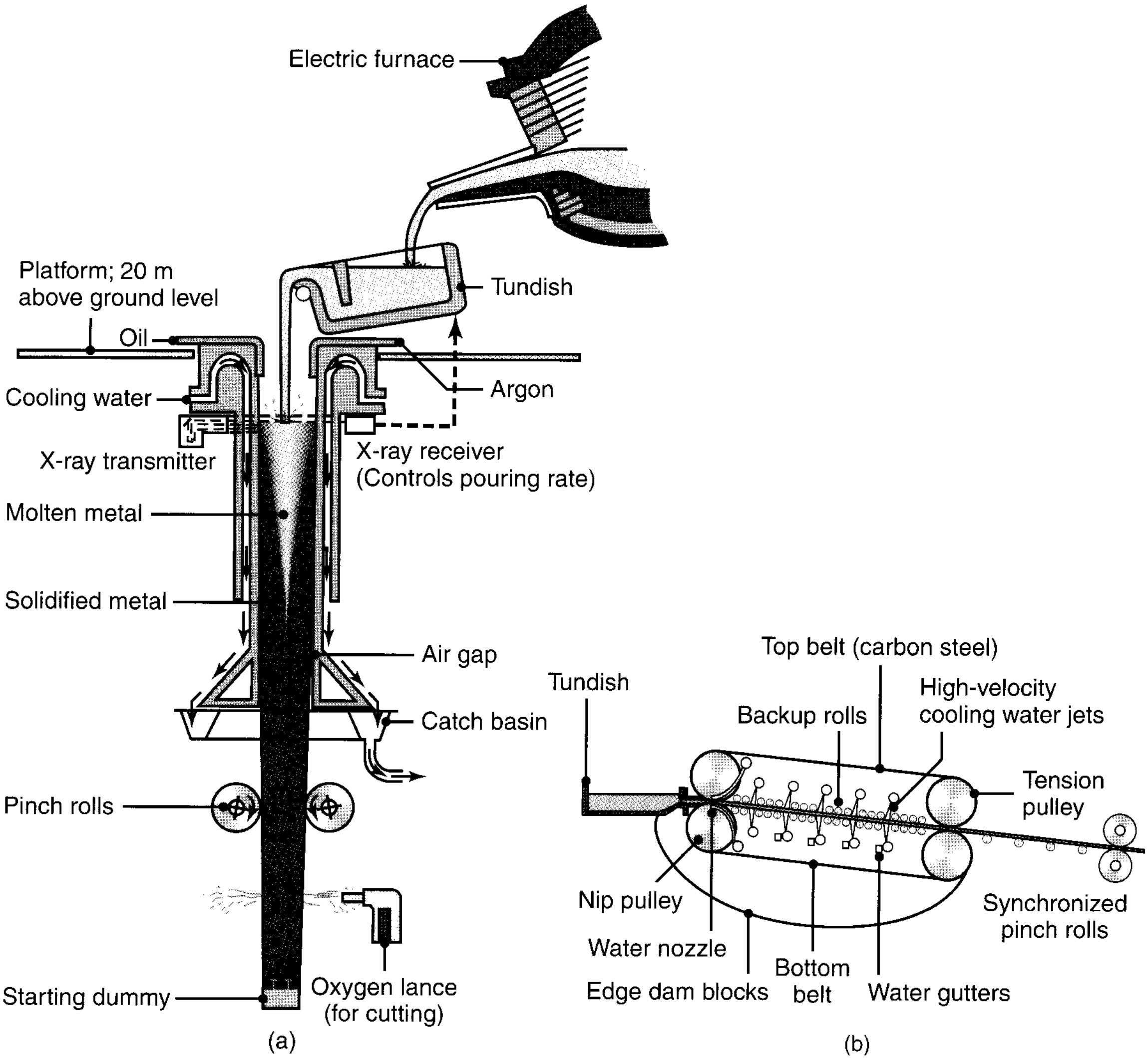


FIGURE 5.4 (a) The continuous-casting process for steel. Typically, the solidified metal descends at a speed of 25 mm/s. Note that the platform is about 20 m above ground level. *Source*: *Metalcaster's Reference and Guide*, American Foundrymen's Society. (b) Continuous strip casting of nonferrous metal strip. *Source*: Courtesy of Hazelett.

metal strips. The process now is used widely for steel 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 cleaned, then it is equalized in temperature by blowing nitrogen gas through it for 5 to 10 minutes. The metal then is poured into a refractory-lined intermediate pouring vessel (**tundish**), where impurities are skimmed off. The tundish holds as much as 3 metric tons of metal. The molten metal is then tapped from the tundish and travels downward through water-cooled copper molds, and begins to solidify; it is drawn through the molds at a constant velocity by rollers (called *pinch rolls*).

Before starting the casting process, a solid *starter bar* (*dummy bar*) is inserted into the bottom of the mold. When the molten metal is first poured, it freezes 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 end 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 generally are coated with graphite or similar solid lubricants in order to reduce both friction and adhesion at the mold–metal interfaces. Also, the molds are vibrated in order to reduce friction and sticking.

The continuously cast metal may be cut into desired 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 of products such as channels and I-beams. In addition to costing less, continuously cast metals have more uniform compositions and properties than those obtained by ingot casting. Modern facilities use computer-controlled hot-rolling operations on continuously cast strands, with final sheet thicknesses on the order of 2 to 6 mm for carbon, stainless, and electrical steels and with capabilities for a rapid switchover from one type of steel to another. Afterwards, steel plates or other shapes undergo one or more further processes, 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) coating (galvanizing or aluminizing) to improve resistance to corrosion.

In **strip casting**, thin slabs, or strips, are produced from molten metal. The metal solidifies in similar fashion to strand casting, but the hot solid then is rolled to form the final shape (Fig. 5.4b). The compressive stresses in rolling (see Section 13.2) serve to reduce porosity and provide better material properties. In effect, 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 and other metals.

5.5 Carbon and Alloy Steels

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). These steels are available in various basic product shapes: plate, sheet, strip, bar, wire, tube, castings, and forgings.

5.5.1 Effects of Various Elements in Steels

Various elements are added to steels in order to impart properties such as hardenability, strength, hardness, toughness, wear resistance, workability, weldability, and machinability. These elements are listed below (in alphabetical order) with summaries of their beneficial and detrimental effects. Generally, the higher the percentages of these elements in steels, 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.

Boron improves hardenability without the loss of (or even with some improvement in) machinability and formability.

Calcium deoxidizes steels, improves toughness, and may improve formability and machinability.

TABLE 5.1

Applications for Selected Carbon and Alloy Steels

Product	Steel	Product	Steel
Aircraft forgings, tubing, fittings	4140, 8740	Gears (car and truck)	4027, 4032
Automobile bodies	1010	Landing gear	4140, 4340, 8740
Axles	1040, 4140	Lock washers	1060
Ball bearings and races	52100	Nuts	3130
Bolts	1035, 4042, 4815	Railroad rails and wheels	1080
Camshafts	1020, 1040	Springs (coil)	1095, 4063, 6150
Chains (transmission)	3135, 3140	Springs (leaf)	1085, 4063, 9260, 6150
Coil springs	4063	Tubing	1040
Connecting rods	1040, 3141, 4340	Wire	1045, 1055
Crankshafts (forged)	1045, 1145, 3135, 3140	Wire (music)	1085
Differential gears	4023		

Carbon improves hardenability, strength, hardness, and wear resistance; it reduces ductility, weldability, and toughness.

Cerium controls the shape of inclusions and improves toughness in high-strength, low-alloy steels; it deoxidizes steels.

Chromium improves toughness, hardenability, wear and corrosion resistance, and high-temperature strength; it increases the depth of hardness penetration resulting from heat treatment by promoting carburization.

Cobalt improves strength and hardness at elevated temperatures.

Copper improves resistance to atmospheric corrosion and, to a lesser extent, increases strength with little loss in ductility; it adversely affects hot-working characteristics and surface quality.

Lead improves machinability; it causes liquid–metal embrittlement.

Magnesium has the same effects as cerium.

Manganese improves hardenability, strength, abrasion resistance, and machinability; it deoxidizes the molten steel, 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 (columbium) imparts fineness of grain size and 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; it decreases magnetic-hysteresis loss, machinability, and cold formability.

Sulfur improves machinability when combined with manganese; it lowers impact strength and ductility and impairs surface quality and weldability.

Tantalum has effects similar to those of niobium.

Tellurium improves machinability, formability, and toughness.

Titanium improves hardenability; it deoxidizes steels.
Tungsten has the same effects as cobalt.
Vanadium improves strength, toughness, abrasion resistance, and hardness at elevated temperatures; it inhibits grain growth during heat treatment.
Zirconium has the same effects as cerium.

5.5.2 Residual Elements in Steels

During steel production, refining, and processing, some *residual elements* (**trace elements**) may remain. Although the elements in the preceding list also can be considered as residuals, the following generally are considered unwanted residual elements:

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.

5.5.3 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 using 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 another designation system, which 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 a few numbers (generally three). The present numbering system is known as the *Unified Numbering System* (UNS) and has been adopted widely by 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 as follows:

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.

5.5.4 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 shown in

TABLE 5.2

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 (%)	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

Table 5.2. The machinability, formability, and weldability of such steels are described in various chapters throughout this text.

- **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 to 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 equipment 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. The higher the carbon content of the steel, the higher is its hardness, strength, and wear resistance after heat treatment.
- 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.

5.5.5 Alloy Steels

Steels containing significant amounts of alloying elements are called **alloy steels**; they usually are made with more care than are carbon steels. **Structural-grade alloy steels** are used mainly in the construction and transportation industries because of their high strength. Other alloy steels are used in applications where strength, hardness, creep and fatigue resistance, and toughness are required. These steels can be heat treated to obtain the desired properties.

5.5.6 High-strength Low-alloy Steels

In order to improve the strength-to-weight ratio of steels, a number of **high-strength, low-alloy steels** (HSLA) have been developed. First developed in the 1930s, HSLA

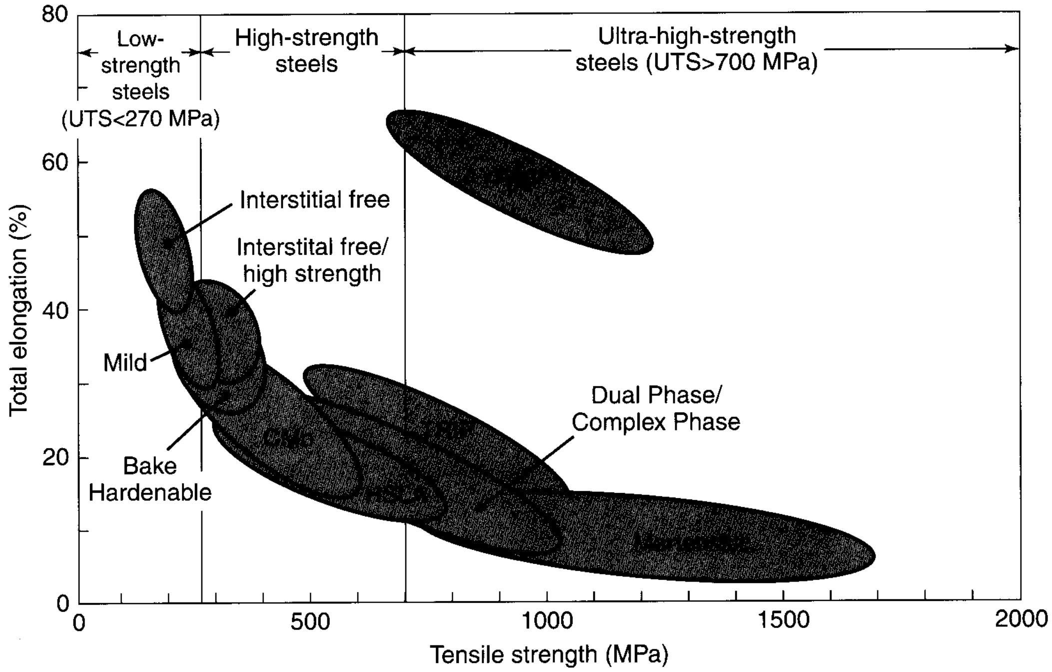


FIGURE 5.5 Comparison of advanced high-strength steels.

steels are typically produced in sheet form by microalloying followed by controlled hot rolling (Chapter 13). Plates, bars, and structural shapes are made from these steels. HSLA steels have a 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. These steels have high strength and energy-absorption capabilities as compared to conventional steels. The ductility, formability, and weldability of HSLA steels, however, generally are inferior to those of conventional low-alloy steels (see Fig. 5.5). To improve these properties, several ultra-high-strength steels have been developed, as described in Section 5.5.7.

Sheet products of HSLA steels typically are used for parts of automobile bodies and other transportation equipment (in order to reduce weight and, hence, fuel consumption) and in mining, agricultural, and various other industrial applications. HSLA 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. Mechanical properties for selected HSLA steels are given in Table 5.3.

Designations. Three categories compose the system of AISI designations for high-strength sheet steel (Table 5.4). *Structural quality* (S) includes the elements 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 that is approximately four times greater than that of conventional low-carbon steels and contain Si, P, Cu, Ni, and Cr in various combinations. In addition, 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 a 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

TABLE 5.3

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	1,000	700	15	0.13
TRIP 450/800	800	450	29	0.24
TRIP 400/600	600	400	30	0.23
CP 700/800	800	700	12	0.13
MART 950/1200	1,200	950	6	0.07
MART 1250/1520	1,520	1,250	5	0.065

TABLE 5.4

AISI Designations for High-strength Sheet Steel

Yield strength			
psi × 10^3	MPa	Chemical composition	Deoxidation practice
35	240	S = structural quality	F = killed plus sulfide inclusion control
40	275		
45	310		
50	350	X = low alloy	
60	415		K = killed
70	485	W = weathering	
80	550		O = nonkilled
100	690	D = dual-phase	
120	830		
140	970		
Example:			
50		X	F
50 × 10^3 psi (350 MPa) min yield strength		Low alloy	Killed plus sulfide inclusion control

the ultimate tensile strength in MPa. The 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 recently developed HSLA steels provide superior properties and can eliminate the need for heat treatment. They have a ferrite–pearlite microstructure with fine dispersed particles of carbonitride. A number of microalloyed steels have been produced, with a typical microalloyed steel containing 0.5% C, 0.8% Mn, and 0.1% V. When subjected to carefully controlled cooling (usually in air), these materials develop improved and uniform strength. Compared to medium-carbon steels, microalloyed steels also can provide cost savings of as much as 10%, since the manufacturing steps of quenching, tempering, and stress relieving are not required.

Nanoalloyed Steels. Now under development, these steels have extremely small grain sizes (10–100 nm) and are produced using metallic glasses (Section 6.14) as a precursor. The metallic glass is subjected to a carefully controlled vitrification (crystallization) process at a high nucleation rate, resulting in fine nanoscale phases. (See also Section 8.8.)

5.5.7 Ultra-high-strength Steels

Ultra-high-strength 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, tool and die wear, forming loads, and springback.

Dual-phase steels are processed specially to have a mixed ferrite and martensite structure. Developed in the late 1960s, these steels 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 complicated 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 processing. As can be seen in Fig. 5.5, TWIP steels combine high strength with 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 provide ultimate tensile strengths as high as 800 MPa and are therefore of interest for automotive crash applications such as bumpers and roof supports. **Martensitic grades** are also available, consisting of high fractions of martensite to attain tensile strengths as high as 1500 MPa.

EXAMPLE 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 applied advanced high-strength steels in structural elements of automobiles. For example, the application of steel in the Ford 500 automobile is shown in Fig. 5.6. Note that although 60% of the steel in this automobile is mild steel and is associated with body panels and transmission and engine components, structural components are exploiting the higher strength-to-weight ratios of advanced high-strength steels.

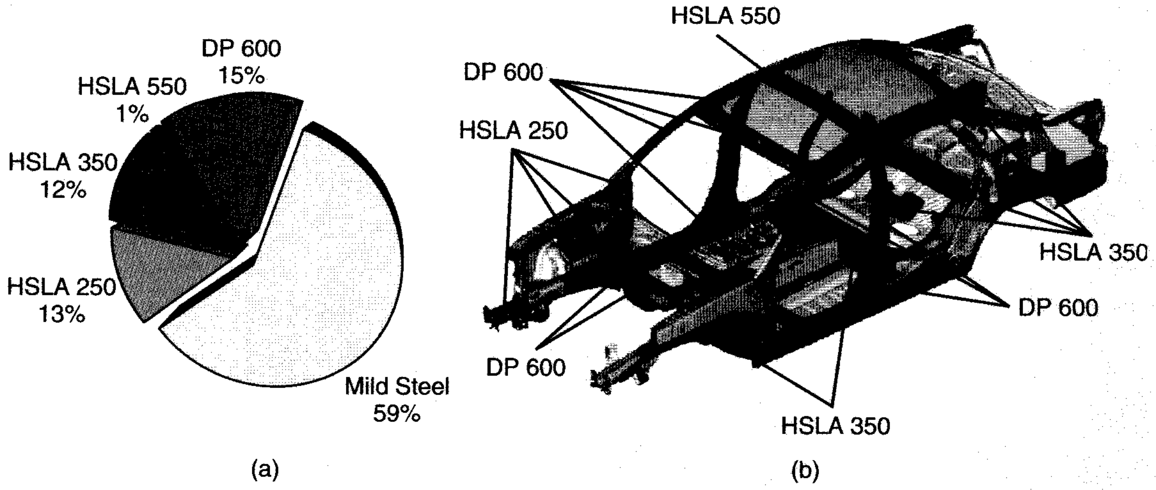


FIGURE 5.6 Advanced high-strength steel applications in the Ford 500. (a) Use of advanced high-strength steels by weight percent; (b) structural components and alloys used. *Source*: Courtesy of the American Iron and Steel Institute—Great Designs in Steel 2005.

5.6 Stainless Steels

Stainless steels are characterized primarily by their corrosion resistance, high strength and ductility, and high chromium content. 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 to 12% by weight. In addition to chromium, other alloying elements in stainless steels typically are nickel, molybdenum, copper, titanium, silicon, manganese, columbium, aluminum, nitrogen, and sulfur.

The letter L is used to identify low-carbon stainless steels. 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 availability of chromium lowers the passivity of the steel. Also, the chromium carbide introduces a second phase and thereby promotes galvanic corrosion.

Developed in the early 1900s, stainless steels are produced in electric furnaces or the basic-oxygen process and by techniques similar to those used in other types of steelmaking. The level of purity is controlled by various refining techniques. Stainless steels are available in a wide variety of shapes, and typical applications include cutlery, kitchen equipment, health care and surgical equipment, and applications in the chemical, food-processing, and petroleum industries. A more recent development is the use of thin, high-strength stainless steels for bus bodies and chassis. Although not commercially successful, the DeLaurean automobile had a stainless-steel body.

Stainless steels generally are divided into five types (see also Table 5.5).

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. Austenitic stainless

TABLE 5.5

Mechanical Properties and Typical Applications of Selected Annealed Stainless Steels at Room Temperature

AISI (UNS)	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

steels are hardened by cold working. They are the most ductile of all stainless steels and can be formed easily. 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 components for severe chemical environments.

Ferritic (400 series). These steels have a high chromium content—up to 27%. They are magnetic and have good corrosion resistance, but they have lower ductility than austenitic stainless steels. Ferritic stainless steels are hardened by cold working and are not heat treatable. 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 by heat treatment. Their chromium content may be as much as 18%. These steels are magnetic, and they have high strength, hardness, and fatigue resistance, good ductility, and moderate corrosion resistance. Martensitic stainless steels typically are used for cutlery, surgical tools, instruments, valves, and springs.

Precipitation-hardening (PH). These steels contain chromium and nickel, along with copper, aluminum, titanium, or molybdenum. They have good corrosion resistance and ductility, and they have high strength at elevated temperatures. Their main application is in aircraft and aerospace structural components.

Duplex Structure. These steels have a mixture of austenite and ferrite. They have good strength and have 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 in heat-exchanger components.

EXAMPLE 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 (by means of strain hardening) and gives the wheel cover a springlike action.

Type 409 is used extensively for catalytic converters. Type 430 had been used for automotive trim, but it is not as resistant as type 434 is to the deicing salts used in colder climates in winter. As a result, its use is now limited. In addition to being more corrosion resistant, type 434 closely resembles the color of chromium plating, so it has become an attractive alternative to 430. Stainless steels also are well suited for use in other automobile components: exhaust manifolds (replacing cast-iron manifolds to reduce weight, increase durability, provide higher thermal conductivity, and reduce emissions), mufflers and tailpipes (to offer better corrosion protection in harsh environments), and brake tubing.

Another example of the use of stainless steel in automobiles is the DeLaurean sports car, manufactured during 1981–1982 in Northern Ireland by the DeLaurean Motor Company. The DMC-12 model of this car (at an original price of $12,000) had a body made of buffed 304 stainless-steel panels attached to a fiberglass underbody; the cars were sold unpainted. For a variety of reasons, however, the car was not commercially successful in spite of its innovative features. It is estimated that about 6,500 are still in existence.

5.7 Tool and Die Steels

Tool and die steels are specially alloyed steels (Tables 5.6 and 5.7) designed for 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).

TABLE 5.6

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 5.7

Processing and Service Characteristics of Common Tool and Die Steels

AISI designation	Resistance to decarburization	Resistance to cracking	Approximate hardness (HRC)	Machinability	Toughness	Resistance to softening	Resistance to wear
M2	Medium	Medium	60–65	Medium	Low	Very high	Very high
T1	High	High	60–65	Medium	Low	Very high	Very high
T5	Low	Medium	60–65	Medium	Low	Highest	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
H26	Medium	High	43–58	Medium	Medium	Very high	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

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 to 18% tungsten with chromium, vanadium, and cobalt as other alloying elements. The M-series steels generally have higher abrasion resistance than T-series steels, undergo less distortion in heat treatment, and are less expensive. The M-series constitutes about 95% of all the high-speed steels produced in the United States. High-speed steel tools can be coated with titanium nitride and titanium carbide for improved wear resistance.

5.7.2 Die Steels

Hot-work steels (H-series) are designed for use at elevated temperatures. They have high toughness, as well as 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. These steels are available as oil-hardening or air-hardening types. **Shock-resisting steels** (S-series) are designed for impact toughness and are used in applications such as header dies, punches, and chisels. Other properties of these steels depend on the particular composition. Various tool and die materials for a variety of manufacturing applications are presented in Table 5.8.

TABLE 5.8

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	
Shearing	
Cold	D2, A2, A9, S2, S5, S7
Hot	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 generally low cost have made them among the most useful of all metallic materials.
- Steelmaking processes have been improved upon continuously, notably by the continuous-casting and secondary-refining techniques. These advances have resulted in higher quality steels and in higher efficiency and productivity.
- Carbon steels generally are classified as low-carbon (mild steel), medium-carbon, and high-carbon steels. Alloy steels contain a variety of alloying elements, particularly chromium, nickel, and molybdenum.
- High-strength low-alloy 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 heat treatment.

- Stainless steels have chromium as a major alloying element; they are called stainless because they form a passivating chromium oxide layer on their surface. Stainless steels generally are classified as austenitic, ferritic, martensitic, and precipitation-hardening steels.
- Tool and die steels are among the most important materials and are used widely in casting, forming, and machining operations for both metallic and nonmetallic materials. They generally consist of high-speed steels (molybdenum and tungsten types), 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-strength low-alloy steels
Ingot
Killed steel
Martensitic steel
Microalloyed steels
Nanoalloyed steels
Open-hearth furnace
TWIP steels
Pig iron
Refining
Rimmed steel
Semikilled steel
Stainless steels
Steel
Strand casting
Tool and die steels
Trace elements
TRIP steels
TWIP steels

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**, 3rd ed., Butterworth-Heinemann, 2006.

Bryson, W.E., **Heat Treatment, Selection and Application of Tool Steels**, 2nd ed., Hanser Gardner, 2005.

Cardelli, F., **Materials Handbook: A Concise Desk Reference**, 2nd ed., Springer, 2008.

Llewellyn, D.T., and Hudd, R.C., **Steels: Metallurgy and Applications**, 3rd ed., Butterworth-Heinemann, 1999.

McGuire, M.F., **Stainless Steels for Design Engineers**, ASM International, 2008.

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.

Schneider, W., **Continuous Casting**, Wiley, 2006.

Totten, G.E., **Steel Heat Treatment: Metallurgy and Technologies**, 2nd ed., CRC Press, 2006.

REVIEW QUESTIONS

5.1. What are the major categories of ferrous alloys?

5.2. List the basic raw materials used in making iron and steel, and explain their functions.

5.3. List the types of furnaces commonly used in steelmaking, and describe their characteristics.

5.4. List and explain the characteristics of the types of steel ingots.

5.5. What does *refining* mean? How is it done?

5.6. What advantages does continuous casting have over casting into ingots?

5.7. Name the four alloying elements that have the greatest effect on the properties of steels.

5.8. What are trace elements?

5.9. What are the percentage carbon contents of low-carbon, medium-carbon, and high-carbon steels?

5.10. How do stainless steels become stainless?

5.11. What are the major alloying elements in tool and die steels and in high-speed steels?

5.12. How does chromium affect the surface characteristics of stainless steels?

5.13. What kinds of furnaces are used to refine steels?

5.14. What is high-speed steel?

5.15. What is TRIP? TWIP?

5.16. What are the applications of advanced high-strength steels?

QUALITATIVE PROBLEMS

5.17. Identify several different products that are made of stainless steel, and explain why they are made of that material.

5.18. 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.

5.19. Why is the control of the structure of an ingot important?

5.20. Explain why continuous casting has been such an important technological advancement.

5.21. Describe applications in which you would not want to use carbon steels.

5.22. 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.

5.23. The cost of mill products of metals increases with decreasing thickness and section size. Explain why.

5.24. Describe your observations regarding the information given in Table 5.8.

5.25. How do trace elements affect the ductility of steels?

5.26. Comment on your observations regarding Table 5.1.

5.27. In Table 5.8, D2 steel is listed as a more common tool and die material for most applications. Why is this so?

5.28. 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?

5.29. 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.

5.30. 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?

5.31. 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.)

5.32. List and explain the advantages and disadvantages of using advanced high-strength steels.

QUANTITATIVE PROBLEMS

5.33. 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.

5.34. 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.

5.35. 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.

▶**5.36.** Using strength and density data, determine the minimum weight of a 600-mm-long tension member that must support a load of 450 kg, 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.

5.37. The endurance limit (fatigue life) of steel is approximately one-half the ultimate tensile strength (see Fig. 2.17), 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.)

5.38. Using the data given in Table 5.3, obtain the power-law curves for the advanced high-strength steels shown and plot the curves. Compare these materials with those given in Table 2.6.

SYNTHESIS, DESIGN, AND PROJECTS

5.39. 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.

5.40. 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.

5.41. 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.

5.42. 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.

5.43. 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.

5.44. Aluminum is being used as a substitute material for steel in automobiles. Describe your concerns, if any, in purchasing an aluminum automobile.

5.45. In the 1940s (The Second World War), the *Yamato* was the largest battleship ever built. Find out the weight of this ship, and estimate the number of automobiles that could have been built from the steel used in this one ship alone. Estimate the time it would take to cast that much steel by continuous casting.

5.46. Search the technical literature, and add more parts and materials to those shown in Table 5.1.

Nonferrous Metals and Alloys: Production, General Properties, and Applications

CHAPTER 6

- Nonferrous metals include a wide variety of materials, ranging from aluminum to zinc, with special properties that are indispensable in many consumer and commercial 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, together with several examples
- Shape-memory alloys, amorphous alloys, and metal foams are also described, with examples of their unique applications.

6.1 Introduction

Nonferrous metals and alloys cover a 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 and alloys have numerous important applications because of properties such as good corrosion resistance, high thermal and electrical conductivity, low density, and ease of fabrication (Table 6.2). Typical examples of nonferrous metal and alloy applications are aluminum for cooking utensils and aircraft bodies, 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 engines.

As an example, a turbofan jet engine 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. Without these materials, a jet engine (Fig. 6.1) could not be designed, manufactured, and operated at the power and efficiency levels required.

This chapter introduces the general properties, the production methods, and the important engineering applications for nonferrous metals and alloys. The manufacturing properties of these materials (such as formability, machinability, and weldability) are described in various chapters throughout this text.

TABLE 6.1

Approximate Cost-per-unit-volume for Wrought Metals and Plastics Relative to the Cost of Carbon Steel

Gold	30,000	Magnesium alloys	4–6
Silver	600	Aluminum alloys	2–3
Molybdenum alloys	75–100	High-strength low-alloy steels	1.4
Nickel	20	Gray cast iron	1.2
Titanium alloys	20–40	Carbon steel	1
Copper alloys	8–10	Nylons, acetals, and silicon rubber*	1.1–2
Zinc alloys	1.5–3.5	Other plastics and elastomers*	0.2–1
Stainless steels	2–9		

*As molding compounds.

Note: Costs vary significantly with quantity of purchase, supply and demand, size and shape, and 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
Titanium	Highest strength-to-weight ratio of all metals; good strength and corrosion resistance at high temperatures
Refractory metals	Molybdenum, niobium (columbium), tungsten, and tantalum; high strength at elevated temperatures
Precious metals	Gold, silver, and platinum; generally good corrosion resistance

6.2 Aluminum and Aluminum Alloys

The important advantages of **aluminum** (Al) and its alloys are their high strength-to-weight ratios, resistance to corrosion by many chemicals, high thermal and electrical conductivities, nontoxicity, reflectivity, appearance, and ease of formability and machinability; they are also nonmagnetic. The principal uses of aluminum and its alloys, in decreasing order of consumption, are in containers and packaging (aluminum 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), 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. The frame and the body panels of the new Rolls Royce

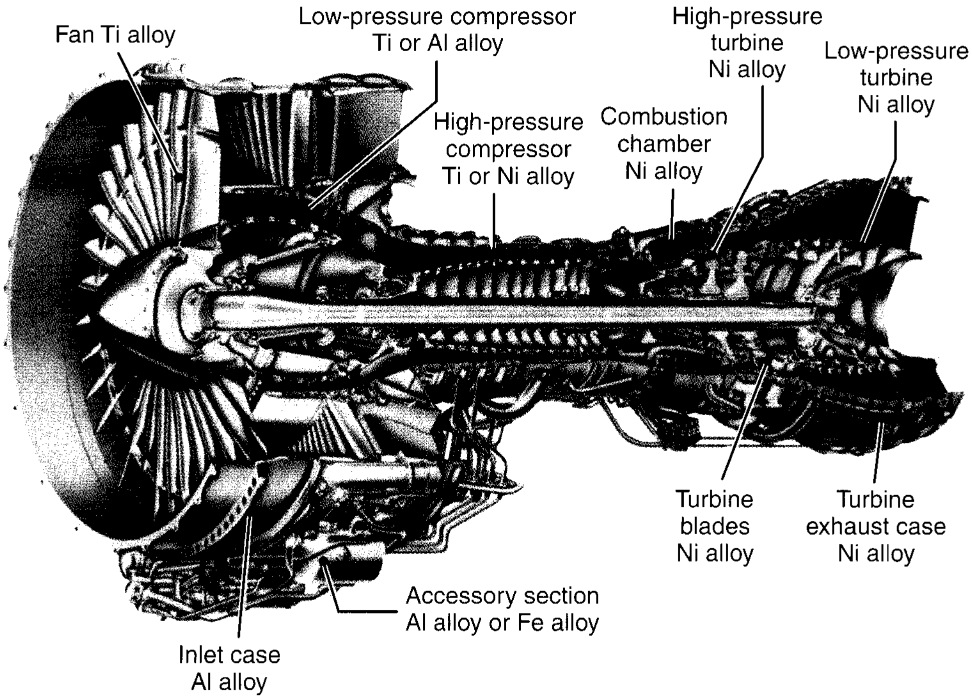


FIGURE 6.1 Cross section of a jet engine (PW2037), showing various components and the alloys used in manufacturing them. *Source*: Courtesy of United Aircraft Pratt & Whitney.

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
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

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.

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
2024	C	B–C	B–C	Truck wheels, screw machine products, aircraft structures
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; 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
7075	C	B–D	D	Aircraft and other structures, keys, hydraulic fittings

*A, excellent; D, poor.

There are two types of wrought alloys of aluminum:

1. Alloys that can be *hardened by cold working* and are not heat treatable.
2. Alloys that can be *hardened by heat treatment.*

Designation of Wrought Aluminum Alloys. Wrought aluminum alloys are identified by four digits and by a **temper designation** that shows the condition of the material. (See also Unified Numbering System later in this section.) The major alloying element is identified by the first digit:

1xxx—Commercially pure aluminum: excellent corrosion resistance, high electrical and thermal conductivity, good workability, low strength, not heat treatable
2xxx—Copper: high strength-to-weight ratio, low resistance to corrosion, heat treatable
3xxx—Manganese: good workability, moderate strength, generally not heat treatable
4xxx—Silicon: lower melting point, forms an oxide film of a dark gray to charcoal color, generally not heat treatable
5xxx—Magnesium: good corrosion resistance and weldability, moderate to high strength, not heat treatable
6xxx—Magnesium and silicon: medium strength; good formability, machinability, weldability, and corrosion resistance; heat treatable
7xxx—Zinc: moderate to very high strength, heat treatable
8xxx—Other element

The second digit in these designations indicates modifications of the alloy. For the 1xxx series, the third and fourth digits stand for the minimum amount of aluminum in the alloy. For example, 1050 indicates a minimum of 99.50% Al, and 1090 indicates a minimum of 99.90% Al. In other series, the third and fourth digits identify the different alloys in the group and have no numerical significance. For instance, a typical aluminum beverage can may consist of the following aluminum alloys, all in the H19 condition (which is the highest cold-worked state): 3004 or 3104 for the can body, 5182 for the lid, and 5042 for the tab. These alloys are selected for their manufacturing characteristics as well as for economic reasons.

Designation of Cast Aluminum Alloys. Designations for cast aluminum alloys also consist of four digits. The first digit indicates the major alloy group, as follows:

1xx.x—Aluminum (99.00% minimum)
2xx.x—Aluminum–copper
3xx.x—Aluminum–silicon (with copper and/or magnesium)
4xx.x—Aluminum–silicon
5xx.x—Aluminum–magnesium
6xx.x—Unused series
7xx.x—Aluminum–zinc
8xx.x—Aluminum–tin

In the 1xx.x series, the second and third digits indicate the minimum aluminum content, as do the third and fourth in wrought aluminum. For the other series, the second and third digits have no numerical significance. The fourth digit (to the right of the decimal point) indicates the product form.

Temper Designations. The temper designations for both wrought and cast aluminum are as follows:

- F—As fabricated (by cold or hot working or by casting)
- O—Annealed (from the cold-worked or the cast state)
- H—Strain hardened by cold working (for wrought products only)
- T—Heat treated
- W—Solution treated only (unstable temper)

Unified Numbering System. As is the case with steels, aluminum and other nonferrous metals and alloys now 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 is for aluminum, C for copper, N for nickel alloys, P for precious metals, and Z for zinc. In the UNS designation, 2024 wrought aluminum alloy is A92024.

Production. Aluminum was first produced in 1825. It is the most abundant metallic element, making up about 8% of the earth's crust, and is produced in a quantity second only to that of iron. The principal ore for aluminum is *bauxite*, which is a hydrous (water-containing) aluminum oxide and includes various other oxides. After the clay and dirt are washed off, the ore is crushed into powder and treated with hot caustic soda (sodium hydroxide) to remove impurities. Next, Alumina (aluminum oxide) is extracted from this solution and then dissolved in a molten sodium-fluoride and aluminum-fluoride bath at 940° to 980°C. This mixture is then subjected to direct-current electrolysis. Aluminum metal forms at the cathode (negative pole), while oxygen is released at the anode (positive pole). *Commercially pure aluminum* is up to 99.99% Al, also referred to in industry as "four nines" aluminum. The production process consumes a great deal of electricity, which contributes significantly to the cost of aluminum.

Porous Aluminum. Blocks of aluminum have been produced that are 37% lighter than solid aluminum and have uniform permeability (*microporosity*). This characteristic allows their use in applications where a vacuum or differential pressure has to be maintained. Examples are the vacuum holding of fixtures for assembly and automation, and the vacuum forming or thermoforming of plastics (Section 19.6). These blocks are 70 to 90% aluminum powder; the rest is epoxy resin. They can be machined with relative ease and can be joined together using adhesives.

EXAMPLE 6.1 An All-aluminum Automobile

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 anywhere in the world, but in 1997 there were seven, including the Plymouth Prowler and the Audi A8 (Fig. 6.2). With weight savings of up to 47% over steel vehicles, such cars use less fuel, create less pollution, and are recyclable.

New alloys and new design and manufacturing methodologies had to be developed. For example, welding and adhesive bonding procedures had to be refined, the structural frame design had to be optimized, and new tooling designs (to allow forming of aluminum) had to be developed. Because of these new technologies, the desired environmental savings were able to be realized without an accompanying drop in performance or safety. In fact, the Audi A8 is the first luxury-class car to earn a dual five-star (highest safety) rating for both driver and front-seat passenger in the National Highway Transportation Safety Administration (NHSTA) New Car Assessment Program.

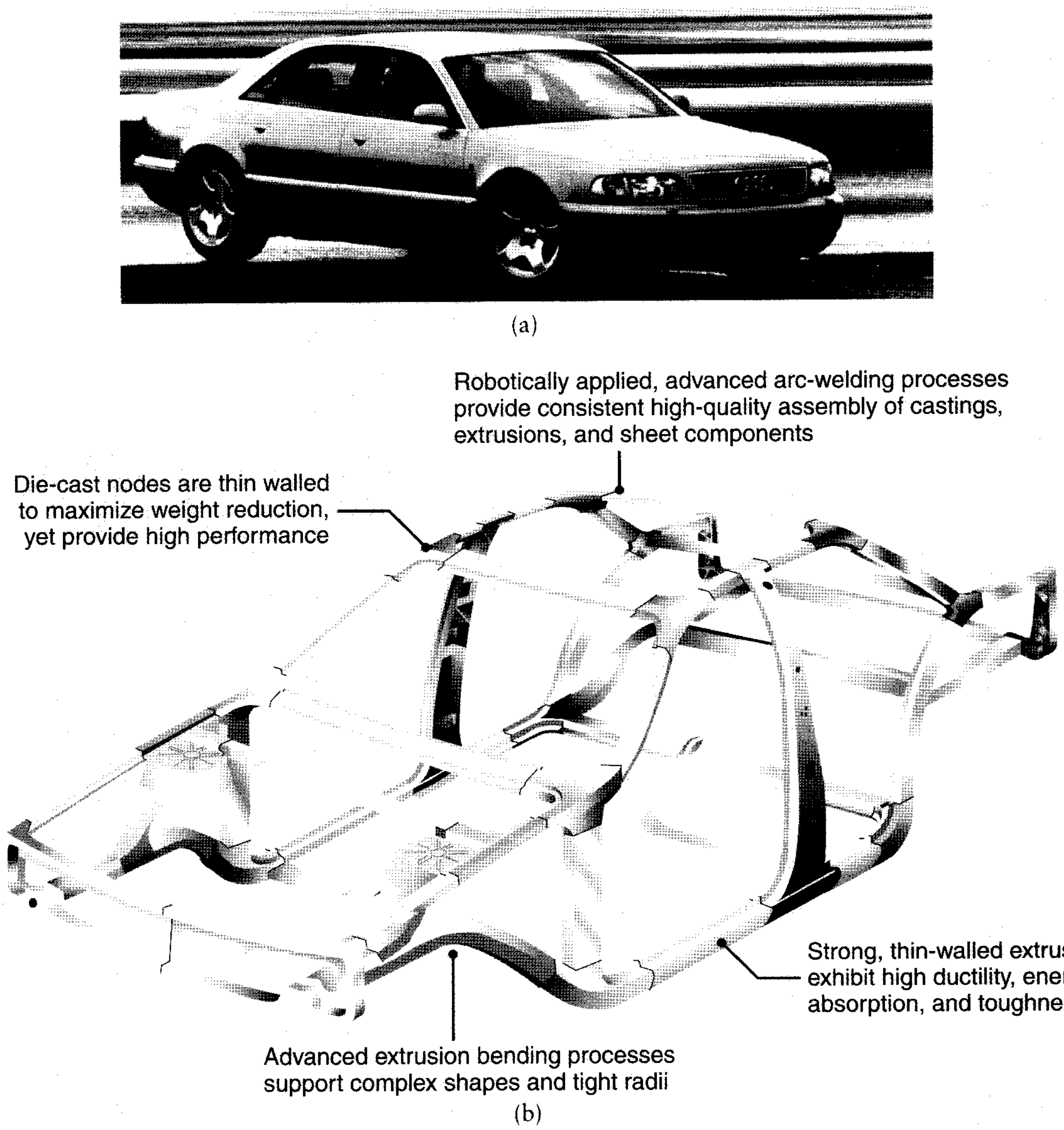


FIGURE 6.2 (a) The Audi A8 automobile, which has an all-aluminum body structure. (b) The aluminum body structure, showing various components made by extrusion, sheet forming, and casting processes. *Source*: Courtesy of ALCOA, Inc.

6.3 Magnesium and Magnesium Alloys

Magnesium (Mg) is the lightest engineering metal available, and it has good vibration-damping characteristics. Its alloys are used in structural and nonstructural applications wherever weight is of primary importance. Magnesium is also an alloying element in various nonferrous metals.

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. Like aluminum, magnesium is finding increased use in the automotive sector, mainly in order to achieve 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 sheets). Magnesium alloys are also used in printing and textile machinery to minimize inertial forces in high-speed components (Section 3.2).

Because it is not sufficiently strong in its pure form, magnesium is alloyed with various elements (Table 6.5) in order to gain certain specific properties, particularly a high strength-to-weight ratio. A variety of magnesium alloys have good casting, forming, and machining characteristics. Because they oxidize rapidly (i.e., they are *pyrophoric*), a fire hazard exists, and precautions must be taken when machining, grinding, or sand-casting magnesium alloys. Products made of magnesium and its alloys are, nonetheless, not a fire hazard during normal use.

Designation of Magnesium Alloys. Magnesium alloys are designated with the following:

a. One or two prefix letters, indicating the principal alloying elements.
b. Two or three numerals, indicating the percentage of the principal alloying elements and rounded off to the nearest decimal.
c. A letter of the alphabet (except the letters I and O) indicating the standardized alloy with minor variations in composition.
d. A symbol for the temper of the material, following the system used for aluminum alloys.

For example, consider the alloy AZ91C-T6:

- The principal alloying elements are aluminum (A, at 9%, rounded off) and zinc (Z, at 1%).
- The letter C, the third letter of the alphabet, indicates that this alloy was the third one standardized (later than A and B, which were the first and second

TABLE 6.5

Properties and Typical Forms of Selected Wrought Magnesium Alloys

	Composition (%)						Ultimate tensile strength (MPa)	Yield strength (MPa)	Elongation in 50 mm (%)	
Alloy	Al	Zn	Mn	Zr	Th	Condition				Typical forms
AZ31B	3.0	1.0	0.2	—	—	F	260	200	15	Extrusions
						H24	290	220	15	Sheet and plates
AZ80A	8.5	0.5	0.2	—	—	T5	380	275	7	Extrusions and forgings
HK31A	—	—	—	0.7	3	H24	255	200	8	Sheet and plates
ZK60A	—	5.7	—	0.55	—	T5	365	300	11	Extrusions and forgings

alloys that were standardized respectively). Specific information related to this standardized alloy can then be obtained.

- T6 indicates that this alloy has been solution treated and artificially aged.

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 in the form of magnesium chloride. First produced in 1808, magnesium metal can be obtained electrolytically or by thermal reduction. In the *electrolytic method*, seawater is mixed with lime (calcium hydroxide) in settling tanks. Magnesium hydroxide precipitates to the bottom, is filtered and mixed with hydrochloric acid. The resulting solution is subjected to electrolysis (as is done with aluminum), producing magnesium metal, which is then cast into ingots for further processing into various shapes. In the *thermal-reduction method*, magnesium ores (dolomite, magnesite, and other rocks) 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, which 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 are among the best conductors of electricity and heat (Tables 3.1 and 3.2), and they have good corrosion resistance. Copper and its alloys can be processed easily by various forming, machining, casting, and joining techniques.

Copper alloys often are attractive for applications in which a combination of electrical, mechanical, nonmagnetic, corrosion-resistant, thermally conductive, and wear-resistant qualities 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 other decorative objects). Although aluminum is the most common material for dies in polymer injection molding (Section 19.3), copper often is used because of its better thermal properties. Pure copper also can be used as a solid lubricant in hot-metal-forming operations.

Copper alloys can acquire a wide variety of properties by the addition of alloying elements and by heat treatment, to improve their manufacturing characteristics. 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 tin bronzes. **Beryllium copper** (or **beryllium bronze**) and **phosphor bronze** have good strength and hardness for applications such as springs and bearings. Other major copper alloys are *copper nickels* and *nickel silvers*.

Designation of Copper Alloys. In the Unified Numbering System, copper is identified with the letter C, such as C26200 for cartridge brass. In addition to being identified by their composition, copper and copper alloys are known by various names (Tables 6.6 and 6.7). The temper designations (such as *1/2 hard*, *extra hard*, *extra spring*, and so on) are based on degree of cold work (such as by rolling or drawing).

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, B (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

Production. Copper is found in several types of ores, the most common being sulfide ores. The ores are generally of low grade (although some contain up to 15% Cu) and usually are obtained from open-pit mines. The slurry is ground into fine particles in ball mills (rotating cylinders with metal balls inside to crush the ore, as shown in Fig. 17.6b, the resulting particles are then suspended in water to form a slurry. Chemicals and oil are then 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; this process is known as **pyrometallurgy**, because heat is used to refine the metal. For applications such as electrical conductors, the copper is refined further electrolytically to a purity of at least 99.95% (*oxygen-free electrolytic copper*). A more recent technique for processing copper is **hydrometallurgy**, a process involving both chemical and electrolytic reactions.

6.5 Nickel and Nickel Alloys

Nickel (Ni) is a silver-white metal discovered in 1751 and a major alloying element that imparts 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), in food-handling and chemical-processing equipment, in coins, and in marine applications. Because nickel is magnetic, nickel alloys also 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 the improvement of their corrosion and wear resistance. Nickel alloys have high strength and corrosion resistance at elevated temperatures. 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.

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 now identified in the UNS system with the letter N. Thus, Hastelloy G is now N06007. Other common trade names are:

- **Monel** is a nickel–copper alloy.
- **Inconel** is a nickel–chromium alloy with a tensile strength of up to 1400 MPa. **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 a 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 (Section 3.6).

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)	None	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

Production. The main sources of nickel are sulfide and oxide ores, all of which have low concentrations of nickel. Nickel metal is produced by preliminary sedimentary and thermal processes, followed by electrolysis; this sequence yields 99.95% pure nickel. Although nickel also is present in the ocean bed in significant amounts, undersea mining is not yet economical.

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, 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 of metals, and the nuclear, chemical, and petrochemical industries. Generally, superalloys are identified by trade names or by special numbering systems, and they are available in a variety of shapes. Most superalloys have a maximum service temperature of about 1000°C in structural applications. The temperatures can be as high as 1200°C for non-load-bearing components.

Superalloys are referred to as *iron-based*, *cobalt-based*, or *nickel-based.*

- **Iron-based superalloys** generally contain from 32 to 67% Fe, from 15 to 22% Cr, and from 9 to 38% Ni. Common alloys in this group are the *Incoloy* series.
- **Cobalt-based superalloys** generally contain from 35 to 65% Co, from 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%. These superalloys 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.

TABLE 6.9

Properties and Typical Applications of Selected Nickel-based 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
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

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 have been developed for service at 550°C for long periods of time and 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 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. Therefore, control of composition and processing are important, especially the prevention of surface contamination by hydrogen, oxygen, or nitrogen during processing; these elements cause embrittlement 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 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.

Production. Ores containing titanium first are reduced to titanium tetrachloride in an arc furnace, then converted to titanium chloride in a chlorine atmosphere. This 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 processed later into various shapes. The complexity of these multistep thermochemical operations (the *Kroll process* developed in the 1940–1950s) adds considerably to the cost of titanium. New developments in electrochemical extraction processes are taking place to reduce the number of steps involved and the energy consumption, thereby reducing the cost of producing titanium.

TABLE 6.10

Properties and Typical Applications of Selected Wrought Titanium Alloys at Various Temperatures

Nominal composition (%)	UNS	Condition	Ultimate tensile strength (MPa)	Yield strength (MPa)	Elongation (%)	Reduction of area (%)	Temp. (°C)	Ultimate tensile strength (MPa)	Yield strength (MPa)
99.5 Ti	R50250	Annealed	330	240	30	55	300	150	95
5 Al, 2.5 Sn	R54520	Annealed	860	810	16	40	300	565	450
6 Al, 4 V	R56400	Annealed	1000	925	14	30	300	725	650
		Solution + age	1175	1100	10	20	300	980	900
13 V, 11 Cr, 3 Al	R58010	Solution + age	1275	1210	8	—	425	1100	830

6.8 Refractory Metals and Alloys

There are four **refractory metals**: molybdenum, niobium, tungsten, and tantalum. These metals are called *refractory* because of their high melting points. Although they were discovered about 200 years ago and have been used as important alloying elements in steels and superalloys, their use as engineering metals and alloys did not begin until about the 1940s. More than most other metals and alloys, the refractory metals maintain their strength at elevated temperatures. Therefore, they 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 discovered in the 18th century and has a high melting point, high modulus of elasticity, good resistance to thermal shock, and good electrical and thermal conductivity. Molybdenum is used in greater 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, which necessitates the use of protective coatings.

Production. The main source of molybdenum is the mineral *molybdenite* (molybdenum disulfide). The ore first is processed and the molybdenum is concentrated; later, it is chemically reduced—first with oxygen and then with hydrogen. Powder-metallurgy techniques 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 referred to as **columbium** (after its source mineral, *columbite*). Niobium possesses good ductility and formability and has greater 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 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 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") was first identified in 1781; it is the most abundant of all the refractory metals. Tungsten has the highest melting point of any metal (3410°C). As a result, it is notable for its high strength at elevated temperatures. However, it has high density (hence it is used for balancing weights and counterbalances in mechanical systems, including self-winding watches), is brittle at low temperatures, and offers poor resistance to oxidation. As an alloying element, tungsten imparts strength and hardness to steels at elevated temperatures.

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. The filament wire in incandescent light bulbs (Section I.1) is made of pure tungsten and is produced by the use of powder-metallurgy and wire-drawing techniques. Tungsten carbide, with cobalt as a binder for the carbide particles, is one of the most important tool and die materials. Tungsten is processed from ore concentrates by chemical decomposition and is then reduced. It is further processed by powder-metallurgy techniques in a hydrogen atmosphere.

6.8.4 Tantalum

Identified in 1802, **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, and chemical industries; it also is used for thermal applications, such as in furnaces and in acid-resistant heat exchangers. A variety of tantalum-based alloys are available in many forms for use in missiles and aircraft. Tantalum also is used as an alloying element. 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 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 employed in various applications, including springs (*beryllium copper*), electrical contacts, and nonsparking tools for use in such explosive environments as mines and metal-powder production (Section 17.2.3). Beryllium and its oxide are toxic; their associated dust and fumes should not be inhaled.

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 Low-melting Alloys

Low-melting alloys are so named because of their relatively low melting points. The major metals in this category are lead, zinc, tin, and their alloys.

6.11.1 Lead

Lead (Pb, after *plumbum,* the root of the word "plumber") has characteristic 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 sound and vibrations, radiation shielding against X-rays, ammunition, as weights, and in the chemical industry. Because it creeps even at room temperature, the use of lead for load-bearing applications is very limited.

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. Lead is also an alloying element in *solders,* steels, and copper alloys; it promotes corrosion resistance and machinability. An additional use of lead is as a solid lubricant for hot-metal-forming operations. Because of its toxicity, however, environmental contamination by lead (causing lead poisoning) is a major concern; major efforts are currently being 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.

6.11.2 Zinc

Zinc (Zn), is bluish white in color and is the metal that is fourth most utilized industrially, after iron, aluminum, and copper. Although its existence was known for many centuries, zinc was not developed until the 18th century. It has three major uses: (1) for galvanizing iron, steel sheet, and wire, (2) as an alloy in other metals, and (3) as a material in castings.

In **galvanizing,** zinc serves as an anode and protects the steel (cathode) from corrosive attack should the coating be scratched or punctured. Zinc also is used as an alloying element; brass, for example, is an alloy of copper and zinc. Major alloying elements in zinc-based alloys 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 for making such products as fuel pumps and grills for automobiles, components for household appliances such as vacuum cleaners and washing machines, kitchen equipment, various machinery parts, and photoengraving equipment. Another use for zinc is in superplastic alloys 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.

Production. A number of minerals containing zinc are found in nature. The principal mineral source is zinc sulfide, also called *zincblende.* The ore first is roasted in air and converted to zinc oxide. It then is reduced to zinc either electrolytically (with the use of sulfuric acid) or by heating it in a furnace with coal (which causes the molten zinc to separate).

6.11.3 Tin

Although used in small amounts compared with iron, aluminum, or copper, **tin** (Sn, from the Latin *stannum*) is an important metal. The most extensive use of tin (a silver-white, lustrous metal) 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 sheets improves deep drawability (Section 16.7.1) and performance in general pressworking. 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 a lining material 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. The alloying elements impart hardness, strength, and corrosion resistance. Tin is an alloying element for dental alloys and for bronze (copper–tin alloy), titanium, and zirconium alloys. Tin–lead alloys are common soldering materials, 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. These alloys are known as **babbitts** (after I. Babbitt, 1799–1862) and contain tin, copper, and antimony. **Pewter**, an alloy of tin, copper, and antimony, was developed in the 15th century and has been used for tableware, hollowware, and decorative artifacts. Organ pipes are made of tin alloys. The most important tin mineral is *cassiterite* (tin oxide), which is of low grade. The ore is mined, concentrated by various techniques, smelted, refined, and cast into ingots for further processing.

6.12 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 (Table 3.2). However, it develops an oxide film that affects its surface characteristics and appearance. Typical applications for silver include tableware, jewelry, coinage, electroplating, electrical contacts, 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.

6.13 Shape-memory Alloys (Smart Materials)

Shape-memory alloys are unique in that, after being plastically deformed into various shapes at room temperature, they return to their original shape upon heating. For example, a piece of straight wire made of such material can be wound into the shape of a helical spring; when heated with a match, the spring uncoils and returns to the 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 such alloys are copper–aluminum–nickel, copper–zinc–aluminum, iron–manganese–silicon, and titanium–nickel–hafnium. Shape-memory alloys generally have such properties as good ductility, corrosion resistance, and high electrical conductivity.

Applications of shape-memory alloys include various sensors, eyeglass frames, stents for blocked arteries, relays, pumps, switches, connectors, clamps, fasteners, and seals. 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 3 seconds after the water temperature reaches 47°C. New developments include thin-film shape-memory alloys deposited on polished silicon substrates for use in microelectromechanical (MEMS) devices (see Chapter 29).

6.14 Amorphous Alloys (Metallic Glasses)

A class of metal alloys that, unlike metals, do not have a long-range crystalline structure is called *amorphous alloys*; these metals have no grain boundaries, and their atoms are packed randomly and tightly. The amorphous structure first was obtained in the late 1960s by the extremely **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, which are 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; this 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. The latter property is utilized in the production of magnetic steel cores for transformers, generators, motors, lamp ballasts, magnetic amplifiers, and linear accelerators. The low magnetic hysteresis loss provides greatly improved efficiency; however, fabrication costs are significant. Amorphous steels are being developed with strengths twice those of high-strength steels, with 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.

6.15 Metal Foams

Metal foams are material structures where the metal consists of only 5 to 20% of the structure's volume, as shown in Fig. 6.3. 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; this 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.

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 for aerospace applications. Because of their porosity, other applications of metal foams are filters and orthopedic implants. Recent developments include nickel–manganese–gallium metal foams with shape-memory characteristics (Section 6.13).

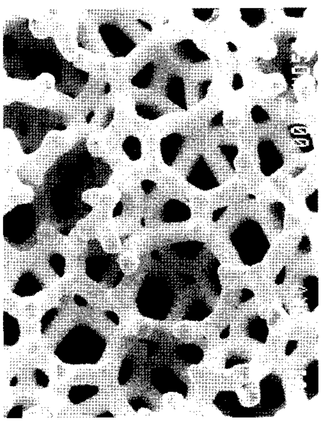


FIGURE 6.3 Microstructure of a metal foam used in orthopedic implants to encourage bone ingrowth. *Source*: Courtesy of Zimmer, Inc.

SUMMARY

- Nonferrous metals and alloys include a very broad range of materials. The most common are aluminum, magnesium, and copper and their alloys, which have a wide range of applications. For higher 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 desired properties.
- Shape-memory alloys (smart materials) have unique properties, with numerous applications in a variety of products and manufacturing operations.
- Amorphous alloys (metallic glasses) have several 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.
- As with all materials, the selection of a nonferrous material for a particular application requires a careful consideration of many 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

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.

Cardelli, F., **Materials Handbook: A Concise Desk Reference**, 2nd ed., Springer, 2008.

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.

Farag, M.M., **Materials Selection for Engineering Design**, Prentice Hall, 1997.

Fremond, M., and Miyazaki, S., **Shape-Memory Alloys**, Springer Verlag, 1996.

Kaufman, J.G., **Introduction to Aluminum Alloys and Tempers**, ASM International, 2000.

Lagoudas, D.C. (ed.), **Shape Memory Alloys: Modeling and Engineering Applications**, Springer, 2008.

Leo, D.J., **Engineering Analysis of Smart Material Systems**, Wiley, 2007.

Lutjering, G., and Williams, J.C., **Titanium**, 2nd ed., Springer, 2007.

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 metallic glasses? Why is the word "glass" used for these materials?

6.8. What is the composition of (a) babbitts, (b) pewter, and (c) sterling silver?

6.9. Name the materials described in this chapter that have the highest (a) density, (b) electrical conductivity, (c) thermal conductivity, (d) strength, and (e) cost.

6.10. What are the major uses of gold, other than in jewelry?

6.11. Describe the advantages to using zinc as a coating for steel.

6.12. What are nanomaterials? Why are they being developed?

6.13. Why are aircraft fuselages made of aluminum alloys, even though magnesium is the lightest metal?

QUALITATIVE PROBLEMS

6.14. Explain why cooking utensils generally are made of stainless steels, aluminum, or copper.

6.15. 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.

6.16. Compare the contents of Table 6.3 with those in various other tables and data on materials in this book, then comment on which of the two hardening processes (heat treating and work hardening) is more effective in improving the strength of aluminum alloys.

6.17. What factors other than mechanical strength should be considered in selecting metals and alloys for high-temperature applications? Explain.

6.18. 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.

6.19. If aircraft, such as a Boeing 757, are made of 79% aluminum, why are automobiles made predominantly of steel?

6.20. Portable (notebook) computers and digital cameras can have their housing made of magnesium. Why?

6.21. 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.

6.22. 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.

QUANTITATIVE PROBLEMS

▶**6.23.** A simply supported rectangular beam is 30 mm wide and 1 m long, and it is subjected to a vertical load of 40 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.

6.24. 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.)

6.26. 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.

▶**6.27.** Using strength and density data, determine the minimum weight of a 900-mm long tension member that must support 340 kg 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.

▶**6.28.** 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

6.29. 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.

6.30. 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.

6.31. 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.

6.32. 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.

6.33. Give some applications for (a) amorphous metals, (b) precious metals, (c) low-melting alloys, and (d) nanomaterials.

6.34. Describe the advantages of making products with multilayer materials. (For example, aluminum bonded to the bottom of stainless-steel pots.)

Polymers: Structure, General Properties, and Applications

CHAPTER 7

- Polymers display a wide range of properties and have several advantages over metallic materials, including low cost, good performance, and ease of manufacturing; for these reasons, polymers 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 discussed.
- The chapter describes the two basic types of polymers: thermoplastics and thermosets.
- Thermoplastics allow a basic manufacturing process of heating them until they soften or melt, and then shaping them into the desired product.
- The process for thermosets is to form the precursors to a desired shape and then set it through polymerization or cross-linking between polymer chains.
- The chapter also describes the properties and uses of elastomers, or rubbers.
- The general properties, typical applications, advantages, and limitations of polymers are discussed throughout the chapter, with several specific examples given.

EXAMPLES:

7.1 Introduction

The word **plastics** first was used as a noun around 1909 and commonly is employed 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 bodies and components.

Because of their many unique and diverse properties, polymers increasingly have replaced metallic components in applications such as automobiles, civilian and military aircraft, sporting goods, toys, appliances, and office equipment. These substitutions reflect the advantages of polymers in terms of the following characteristics:

- Relatively low cost (see Table 6.1) and ease of manufacture
- Corrosion resistance and 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
- Ease of manufacturing and complexity of design possibilities
- Other characteristics that may or may not be desirable (depending on the application), such as low strength and stiffness (Table 7.1), high coefficient of thermal expansion, low useful-temperature range—up to about 350°C—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, machined, cast, and joined into various shapes with relative ease. Minimal additional surface-finishing operations, if any at all, are required; this characteristic provides an important advantage over metals. Plastics are available commercially as film, sheet, plate, rods, and tubing of various cross-sections.

The earliest polymers were made of **natural organic materials** from animal and vegetable products; *cellulose* is the most common example. By means of various chemical reactions, cellulose is modified into *cellulose acetate*, used in making sheets

TABLE 7.1

Range of Mechanical Properties for Various Engineering Plastics at Room Temperature

Material	UTS (MPa)	Young's modulus (E) (GPa)	Elongation (%)	Poisson's ratio, ν
Acrylonitrile-butadiene-styrene (ABS)	28–55	1.4–2.8	75–5	—
ABS, reinforced	100	7.5	—	0.35
Acetal	55–70	1.4–3.5	75–25	—
Acetal, reinforced	135	10	—	0.35–0.40
Acrylic	40–75	1.4–3.5	50–5	—
Cellulosic	10–48	0.4–1.4	100–5	—
Epoxy	35–140	3.5–17	10–1	—
Epoxy, reinforced	70–1400	21–52	4–2	—
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	—
Phenolic	28–70	2.8–21	2–0	—
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	—
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	—

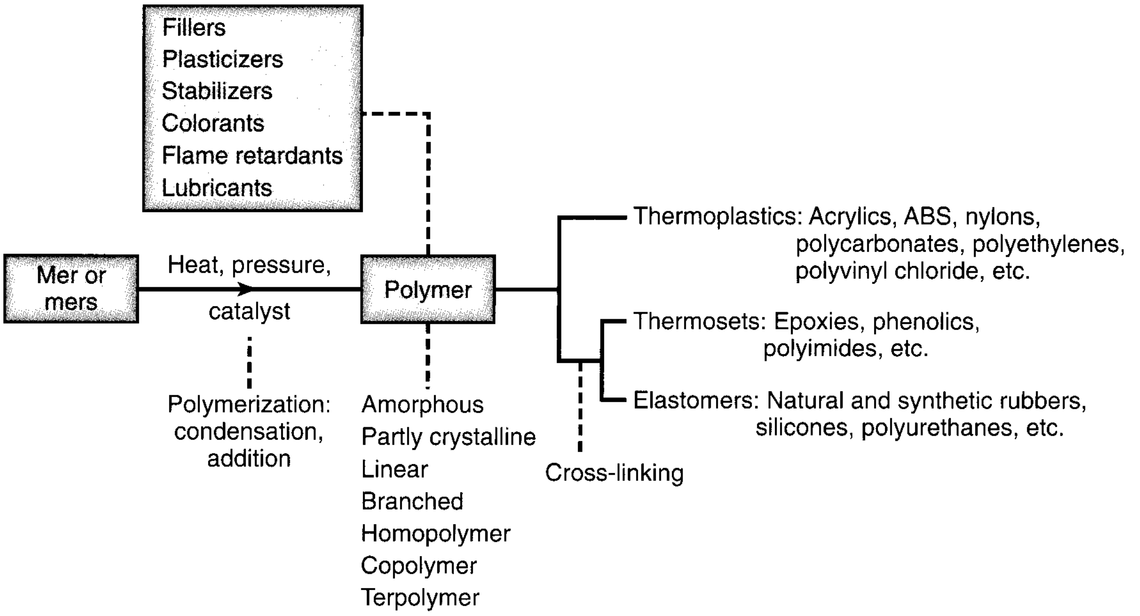


FIGURE 7.1 Outline of the topics described in Chapter 7.

for packaging and textile fibers such as rayon; *cellulose nitrate*, for plastics and explosives; and varnishes. The earliest **synthetic** (manmade) polymer was phenol formaldehyde, a thermoset developed in 1906 and called *Bakelite* (a trade name, after L.H. Baekeland, 1863–1944).

The development of modern plastics technology began in the 1920s when the raw materials necessary for making polymers were extracted from coal and petroleum products. Ethylene was the first example of such a raw material; it became the building block for polyethylene. Ethylene is the product of the reaction between acetylene and hydrogen, and acetylene is the product of the reaction between coke and methane. Commercial polymers, such as polypropylene, polyvinyl chloride, polymethylmethacrylate, polycarbonate, and others, are all made in a similar manner; these materials are known as **synthetic organic polymers**.

An outline of the basic process for making various synthetic polymers is given in Fig. 7.1. In polyethylene, only carbon and hydrogen atoms are involved, but other polymer compounds can be obtained by including chlorine, fluorine, sulfur, silicon, nitrogen, and oxygen. As a result, an extremely wide range of polymers—having among them an equally wide range of properties—has been developed.

This chapter describes the relationship of the structure of a polymer to its properties and behavior, during both manufacturing and its service life under various physical and environmental conditions. This chapter also describes the properties and engineering applications of plastics, rubbers, and elastomers. Reinforced plastics and composite materials are described in Chapter 9, and processing methods for plastics and reinforced plastics in Chapter 19.

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

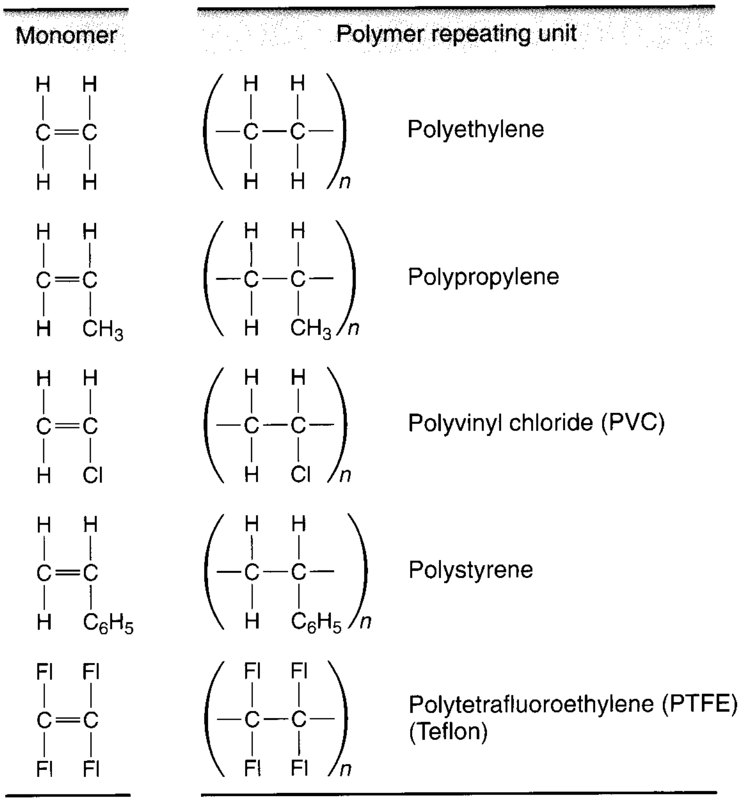


FIGURE 7.2 Molecular structure of various polymers. These are examples of the basic building blocks for plastics.

feature that distinguishes them from most other organic chemical compositions. Polymers are **long-chain molecules** that are 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** (from the Greek *meros*, meaning "part") indicates the smallest repetitive unit; the use of the term is similar to that of *unit cell* in crystal structures of metals (Section 1.3).

The term **polymer** means "many mers (or units)," generally repeated hundreds or thousands of times in a chainlike structure. Most monomers are *organic materials* in which carbon atoms are joined in *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 into polymers in repeating units to make longer and larger molecules by a chemical process called a **polymerization reaction.** Polymerization processes are complex; they will be described only briefly here. 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 reaction by-products (such as water) are condensed out (hence the name). 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. This rate is much higher than that in condensation polymerization. In addition polymerization, an *initiator* is added to open the double bond between two carbon atoms, which begins the linking process by adding many more monomers to a growing chain. For example, ethylene monomers (Fig. 7.3b) link to produce the polymer *polyethylene*; other examples of addition-formed polymers are shown 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 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, but the chain lengths produced fall into a traditional distribution curve. The

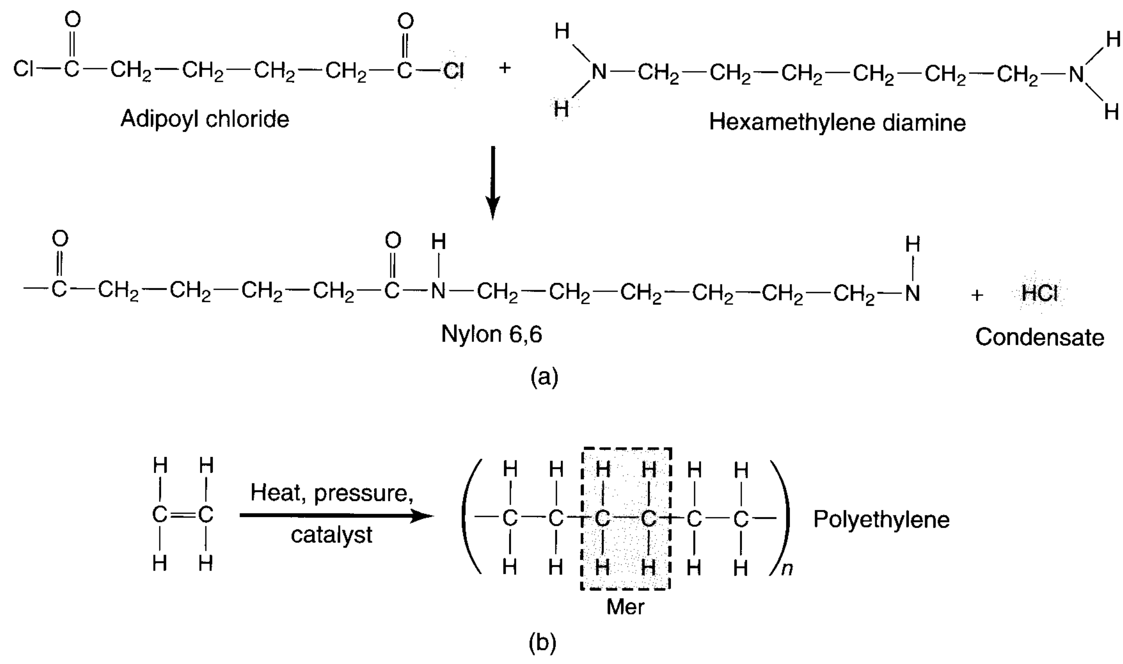


FIGURE 7.3 Examples of polymerization. (a) Condensation polymerization of nylon 6,6 and (b) addition polymerization of polyethylene molecules from ethylene mers.

molecular weight of a polymer is determined on a statistical basis by averaging. The spread of the molecular weights in a chain is referred to as the **molecular weight distribution** (MWD). A polymer's molecular weight and its MWD have a strong 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), which is 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, the DP of PVC with a molecular weight of 50,000 is 50,000/62.5 = 800. In terms of polymer processing (Chapter 19), the higher the DP, the higher is the polymer's viscosity, or its resistance to flow (Fig. 7.4). On the one hand, high viscosity adversely affects the ease of shaping and thus raises the overall cost of processing. On the other hand, 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 also are called **primary bonds**. The polymer chains are, in turn, 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.

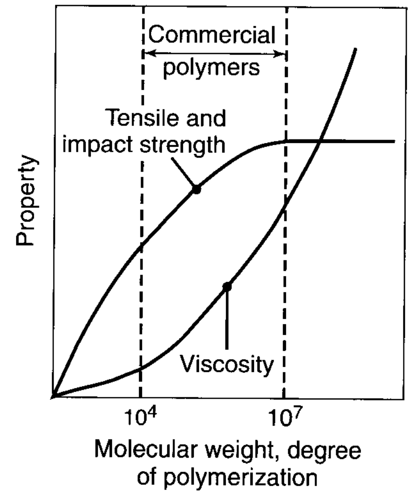
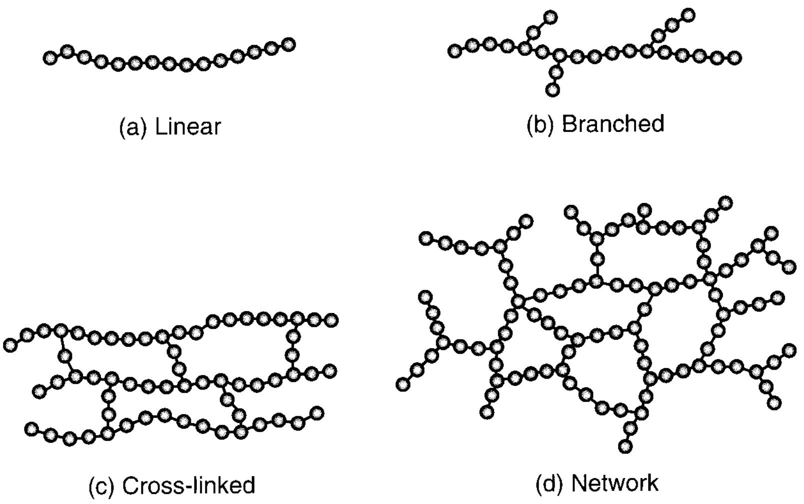


FIGURE 7.4 Effect of molecular weight and degree of polymerization on the strength and viscosity of polymers.



(a) Linear

(b) Branched

(c) Cross-linked

(d) Network

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.

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 shape. In addition to those shown in the figure, other linear polymers are 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 cross-linked chains. As a result of branching and cross-linking, 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 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.

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**; examples are epoxies, phenolics, and silicones. Cross-linking has a major influence on the properties of polymers (generally imparting hardness, strength, stiffness, brittleness, and better dimensional stability; see Fig. 7.6), as well as in the **vulcanization** of rubber (Section 7.9).

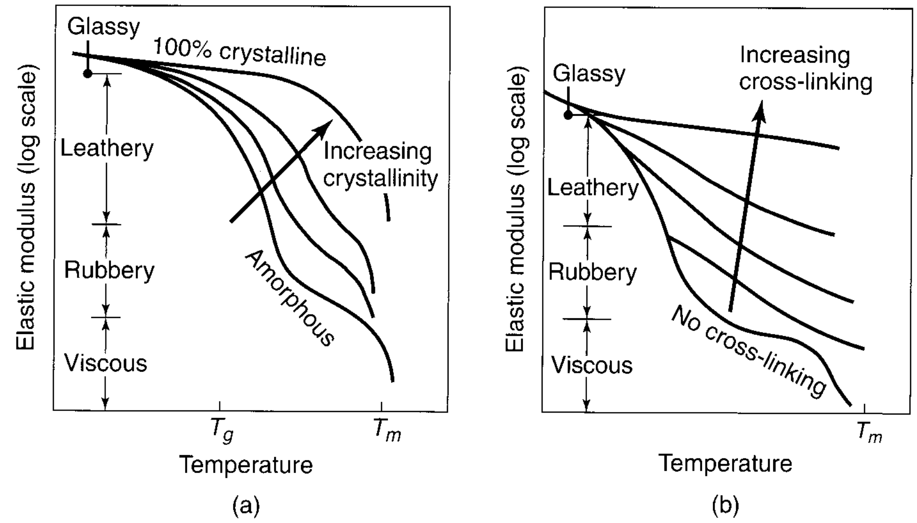


FIGURE 7.6 Behavior of polymers as a function of temperature and (a) degree of crystallinity and (b) cross-linking. The combined elastic and viscous behavior of polymers is known as viscoelasticity.

Network Polymers. These polymers consist of spatial (three-dimensional) networks 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 obtain higher strength by subjecting them to high-energy radiation, such as ultraviolet light, X-rays, or electron beams. However, excessive radiation can 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*. However, as with solid–solution metal alloys (Section 4.2), two or three different types of monomers can be combined to develop certain special properties and characteristics, such as improved strength, toughness, and formability of the polymer. *Copolymers* contain two types of polymers (for example, styrene-butadiene, which is used widely for automobile tires). *Terpolymers* contain three types (for example, acrylonitrile-butadiene-styrene (ABS), which is used for helmets, telephones, and refrigerator liners).

EXAMPLE 7.1 Dental and Medical Bone Cement

Polymethylmethacrylate (PMMA) is an acrylic polymer commonly used in dental and medical applications as an adhesive and is often referred to as bone cement. There are a number of forms of PMMA, but this example describes one common form involving an addition-polymerization reaction. PMMA is delivered in two parts: a powder and a liquid, which are mixed by hand. The liquid wets and partially dissolves the powder, resulting in a liquid with viscosity on the order of 0.1 Ns/m^2, similar to that of vegetable oil. The viscosity increases markedly until a "doughy" state is reached in about five minutes. The "dough" fully hardens in an additional five minutes.

The powder consists of high-molecular-weight poly[(methylmethacrylate)-costyrene] particles about 50 μm in diameter and containing a small volume fraction of benzoyl peroxide. The liquid consists of a methyl methacrylate (MMA) monomer, with a small amount of dissolved *n,n* dimethyl-p-toluidine (DMPT). When the liquid and powder are mixed, the MMA wets the particles (dissolving a surface layer of the PMMA particles) and the DMPT cleaves the benzoyl peroxide molecule into two parts to form a catalyst with a free electron (sometimes 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 fully set bone cement is shown in Fig. 7.7.

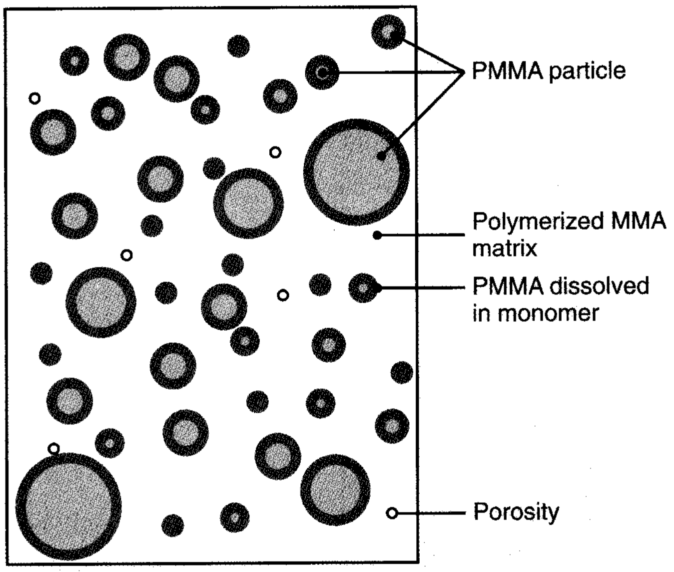


FIGURE 7.7 Schematic illustration of the microstructure of polymethylmethacrylate cement used in dental and medical applications.

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 6.14). The amorphous arrangement of polymer chains often is 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). These crystals 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 rate of solidification during cooling and the chain structure, 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

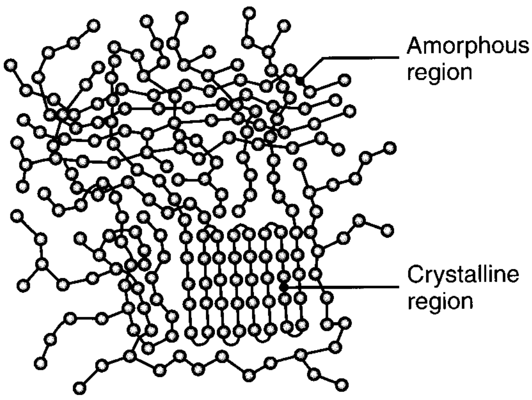


FIGURE 7.8 Amorphous and crystalline regions in a polymer. The crystalline region (crystallite) has an orderly arrangement of molecules. The higher the crystallinity, the harder, stiffer, and less ductile the polymer.

(mostly amorphous) polymers. The degree of crystallinity also is affected by branching. A linear polymer can become highly crystalline, but a highly branched polymer cannot, although it may develop some low level of crystallinity. It will never achieve a high crystallite content, because the branches interfere with the alignment of the chains into a regular crystal array.

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, more dense, 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 a 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 to 0.970 (80 to 95% crystalline). It is stronger, stiffer, tougher, and less ductile than low-density polyethylene (LDPE), which is about 60 to 70% crystalline and has a specific gravity of about 0.910 to 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 causes opaqueness. Furthermore, because the index of refraction is proportional to density, the greater the density difference between the amorphous and 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 6.14.) 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). In the case of highly cross-linked polymers, the slope of the curve changes gradually near T_g; hence, it can be difficult to determine T_g for these polymers. The glass-transition temperature varies with different polymers (Table 7.2) and 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 (first-order changes) that occur, the specific volume of the polymer drops suddenly as its temperature is reduced.

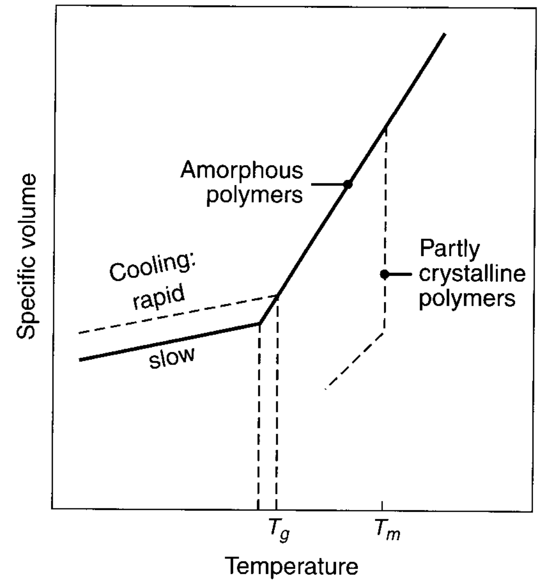


FIGURE 7.9 Specific volume of polymers as a function of temperature. Amorphous polymers, such as acrylic and polycarbonate, have a glass-transition temperature, T_g, but do not have a specific melting point, T_m. Partly crystalline polymers, such as polyethylene and nylons, contract sharply while passing through their melting temperatures during cooling.

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	—

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 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.**

Advances in blending involve several components, creating **polyblends** that utilize the favorable properties of different polymers. **Miscible blends** (mixing without separation of two phases) are created 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 earlier 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, certain polymers become easier to form or mold into desired shapes. The increased temperature weakens the secondary bonds (through thermal vibration of the long molecules), and the adjacent chains can then move more easily when subjected to external shaping forces. 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 depends on other variables as well as their structure and composition. Among the most important are temperature and rate of deformation. Below the glass-transition temperature, most polymers are *glassy* (brittle) and behave like an elastic solid. (That is, the relationship between stress and strain is linear, as shown in Fig. 2.2.) The particular behavior depends on the polymer. For example, PMMA is glassy below its T_g, whereas polycarbonate is not glassy below its T_g: When the applied stress is increased further, polycarbonate eventually fractures, just as a piece of glass does at ambient temperature.

Typical stress–strain curves for some thermoplastics and thermosets at room temperature are shown in Fig. 7.10. Note that these plastics exhibit various behaviors, which may be described as rigid, soft, brittle, flexible, and so on. The mechanical properties of several polymers listed in Table 7.1 indicate that thermoplastics are about two orders of magnitude less stiff than metals. Their ultimate tensile strength is about one order of magnitude lower than that of metals (see Table 2.2).

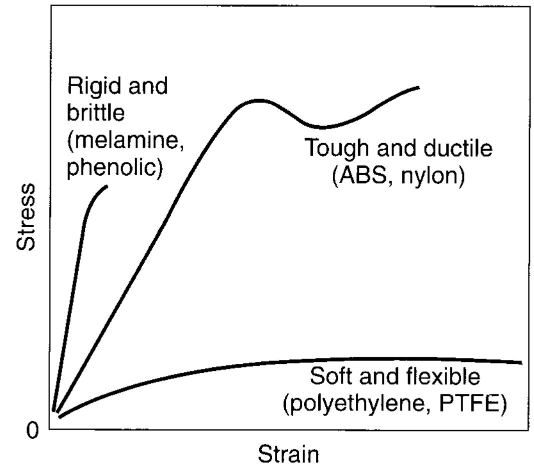


FIGURE 7.10 General terminology describing the behavior of three types of plastics. PTFE (polytetrafluoroethylene) has *Teflon* as its trade name. *Source*: After R.L.E. Brown.

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 (e.g., above T_m for crystalline thermoplastics), it becomes a *viscous fluid*, and its viscosity decreases with increasing temperature. As a viscous fluid, it can be softened, molded into shapes, resolidified, remelted, and remolded a number of times. In practice, however, repeated heating and cooling causes **degradation**, or **thermal aging**, of thermoplastics.

The typical effect of temperature on the strength and elastic modulus of thermoplastics is similar to that of metals: With increasing temperature, the strength and the modulus of elasticity decrease 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 behaviors of various polymers.

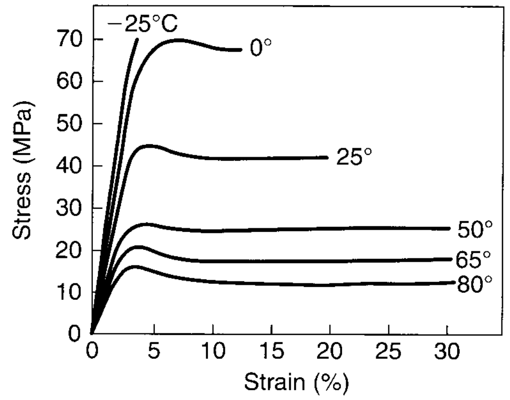


FIGURE 7.11 Effect of temperature on the stress–strain curve for cellulose acetate, a thermoplastic. Note the large drop in strength and the large increase in ductility with a relatively small increase in temperature. *Source*: After T.S. Carswell and H.K. Nason.

Effect of Rate of Deformation. When deformed rapidly, the behavior of thermoplastics is similar to metals, as was indicated 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 fracture. Note in Fig. 7.13 how (unlike in ordinary metals) the necked region elongates considerably. This phenomenon can be easily demonstrated by stretching a piece of the plastic holder for a 6-pack of beverage cans. Observe the sequence of necking and stretching behavior shown in Fig. 7.13a. This characteristic (which is the same in the superplastic metals, Section 2.2.7) enables the thermoforming of thermoplastics (Section 19.6) into such complex shapes as meat trays, lighted signs, and bottles for soft drinks.

Orientation. When thermoplastics are deformed (say, by stretching), the long-chain molecules tend to align in the general direction of the elongation; this process is called *orientation*. As in metals, the polymer becomes anisotropic (see also Section 1.6), so 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 composite materials, as discussed in Chapter 9.

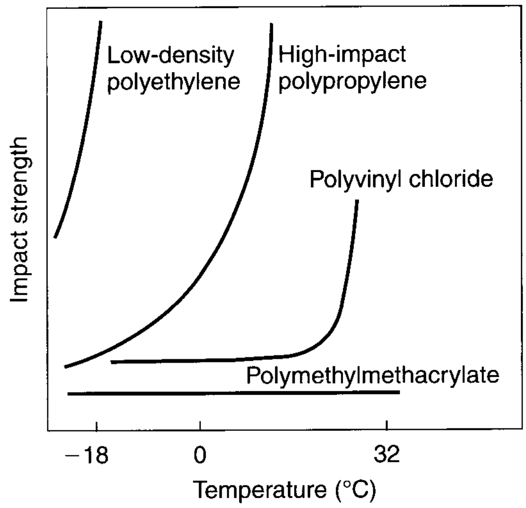


FIGURE 7.12 Effect of temperature on the impact strength of various plastics. Small changes in temperature can have a significant effect on impact strength. *Source*: After P.C. Powell.

Creep and Stress Relaxation. Because of their viscoelastic behavior, thermoplastics are particularly susceptible to creep and stress relaxation, and to a larger extent than metals. The extent of these phenomena depends on the polymer, stress level, temperature, and time. Thermoplastics exhibit creep and stress relaxation at room temperature; 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, a process called *crazing*. Although they may appear to be like cracks, crazes are spongy material, typically containing about 50% voids. With increasing tensile load on the specimen, these voids coalesce to form a crack, which eventually can lead to a 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 (**environmental-stress cracking** and **solvent crazing**). Residual stresses in the material also contribute to crazing and cracking of the polymer; radiation (especially ultraviolet radiation) can increase the crazing behavior in certain polymers.

A phenomenon related to crazing is **stress whitening**. When subjected to tensile stresses (such as those caused by folding or bending), the plastic becomes lighter in color, a phenomenon usually attributed to the formation of microvoids in the material. As a result, the material becomes less translucent (transmits less light), or more opaque. This behavior easily can be demonstrated by bending plastic components commonly found in colored binder strips for report covers, household products, and toys.

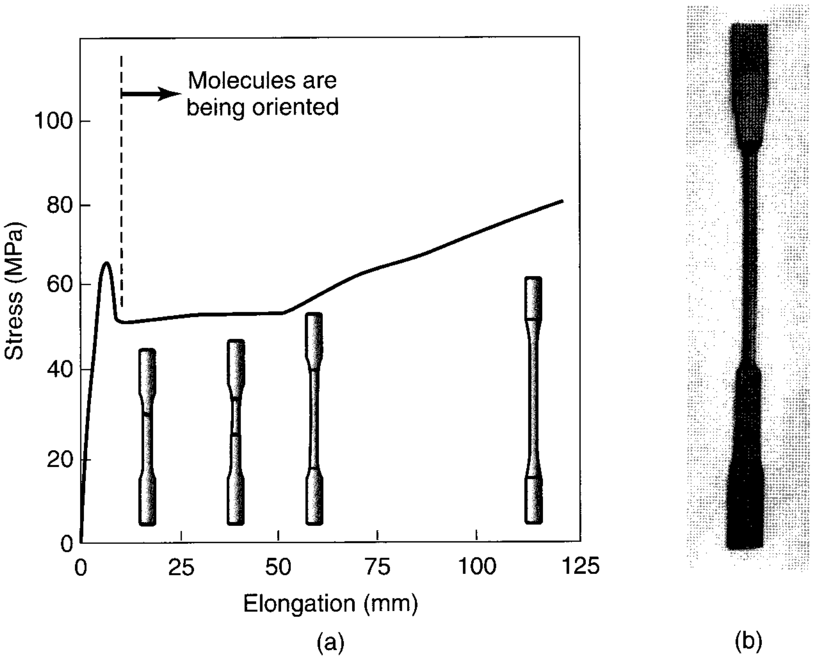


FIGURE 7.13 (a) Stress–elongation curve for polycarbonate, a thermoplastic. *Source*: Courtesy of R.P. Kambour and R.E. Robertson. (b) High-density polyethylene tensile-test specimen, showing uniform elongation (the long, narrow region in the specimen).

Water Absorption. An important characteristic of some polymers, such as nylons, is their ability to absorb water (*hygroscopic*). Water acts as a plasticizing agent: It makes the polymer more plastic (see Section 7.5). In a sense, it lubricates the chains in the amorphous region. With increasing moisture absorption, the glass-transition temperature, the yield stress, and the elastic modulus of the polymer typically are lowered severely. 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 a high coefficient of thermal expansion (about an order of magnitude higher; see 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, and polythiophene. 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).

Thermally conducting polymers also are being developed for applications requiring dimensional stability and heat transfer (such as heat sinks), as well as for reducing cycle times in molding and processing of thermoplastics. These polymers are typically thermoplastics (such as polypropylene, polycarbonate, and nylon) and are embedded with nonmetallic thermally conducting particles; their conductivity can be as much as 100 times that of conventional plastics.

Shape-memory Polymers. Recent investigations have shown that polymers also can behave in a manner similar to the shape-memory alloys described in Section 6.13. The polymers can be stretched or compressed to very large strains, and then, when subjected to heat, light, or a chemical environment, they recover their shape. The potential applications for these polymers are similar to those for shape-memory metals, such as in opening blocked arteries, as well as probing neurons in the brain and improving the toughness of spines.

EXAMPLE 7.2 Use of Electrically Conducting Polymers in Rechargeable Batteries

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. Lithium is used because it is the lightest of all metals and has a high electrochemical potential, so that its energy per volume is highest.

The polymer, usually polyethylene oxide (PEO), with a dissolved lithium salt, is placed between the cathode and anode. During discharge of the battery, Li_yC_6 is oxidized, emitting free electrons and lithium ions. The electrons drive external electronics, and the Li^+ ions are stored in the polymer. When the cathode is depleted, the battery must be recharged to restore the cathode. During charging, Li^+ is transferred through the polymer electrolytes to the cathode. Lithium-ion batteries have good capacity, can generate up to 4.5 V, and can be placed in series to obtain higher voltages. Battery cells are now being developed in which both electrodes are made of conducting polymers; one has been constructed with a capacity of 3.5 V.

7.4 Thermosetting Plastics

When the long-chain molecules in 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. This **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 in the baking of a cake or in the boiling of 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.

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 the 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 (unlike those of thermoplastics) are not affected by temperature or by rate of deformation. If the temperature is increased sufficiently, the thermosetting polymer instead begins to burn up, degrade, and char. Thermosets generally possess better mechanical, thermal, and chemical properties; electrical resistance; and dimensional stability than do thermoplastics. A typical and common thermoset is **phenolic**, which is a product of the reaction between phenol and formaldehyde. Common products made from this polymer are the handles and knobs on cooking pots and pans and components of light switches and outlets.

7.5 Additives in Plastics

In order to impart certain specific properties, polymers usually are compounded with *additives*. These additives modify and improve certain characteristics of the polymer, such as stiffness, strength, color, weatherability, flammability, arc resistance (for electrical applications), and ease of subsequent processing.

- **Plasticizers** are added to polymers to impart *flexibility* and *softness* by lowering their glass-transition temperature. Plasticizers 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 many uses; other applications are in thin sheets, films, tubing, shower curtains, and clothing materials.
- 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. On the other hand, degradation may be beneficial, as in the disposal of plastic objects by subjecting them to environmental attack (see also Section 7.8). A typical example of protection against ultraviolet radiation is the compounding of certain plastics and rubber with **carbon black** (soot). The carbon black absorbs a high percentage of the ultraviolet radiation. Protection

against degradation caused by oxidation, particularly at elevated temperatures, is achieved by adding *antioxidants* to the polymer. Various *coatings* are another means of protecting polymers.

- **Fillers** used in plastics are generally wood flour (fine sawdust), silica flour (fine silica powder), clay, powdered mica, talc, calcium carbonate, and short fibers of cellulose, glass, or asbestos. Because of their low cost, fillers are important in reducing the overall cost of polymers. Depending on their type, fillers also may improve the strength, hardness, toughness, abrasion resistance, dimensional stability, or stiffness of plastics. These properties are greatest at specific percentages of different types of polymer–filler combinations. As with reinforced plastics (Section 9.2), a filler's effectiveness depends on the nature of the bond between the filler material and the polymer chains.
- The wide variety of colors available in plastics is obtained by adding **colorants**—either organic (*dyes*) or inorganic (*pigments*). The selection of a colorant depends on the service temperature and the expected amount of exposure to light. Pigments are dispersed particles; they generally have greater resistance than do dyes to temperature and light.
- If the temperature is sufficiently high, most polymers will ignite and burn; flame colors are usually yellow or blue. The **flammability** (ability to support combustion) of polymers varies considerably, depending on their composition (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.

  The following list gives several common polymers with different burning characteristics:

  1. Plastics that do not burn: *fluorocarbons* (e.g., *Teflon*)
  2. Plastics that do burn but are self-extinguishing: *carbonate, nylon, vinyl chloride*
  3. Plastics that burn and are not self-extinguishing: *acetal, acrylic, acrylonitrile–butadiene–styrene, cellulose, polyester, propylene, styrene.*

- **Lubricants** may be added to polymers to reduce friction during their subsequent processing into useful products and to prevent parts from sticking to the molds. Typical lubricants are linseed oil, mineral oil, and waxes (natural and synthetic); metallic soaps, such as calcium stearate and zinc stearate also are used. Lubrication also is important in preventing thin polymer films from sticking to each other.

7.6 General Properties and Applications of Thermoplastics

The general characteristics and typical applications of major thermoplastics, particularly as they related 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 requiring high performance over a long

TABLE 7.3

General Recommendations for Plastic Products

Design requirement	Typical applications	Plastics
Mechanical strength	Gears, cams, rollers, 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

Trade name	Type	Trade name	Type
Alathon	Ethylene	Noryl	Polyphenylene oxide
Cycolac	Acrylonitrile-butadiene-styrene	Nylon	Polyamide
Dacron	Polyester	Orlon	Acrylic
Delrin	Acetal	Plexiglas	Acrylic
Dylene	Styrene	Royalite	Acrylonitrile-butadiene-styrene
Envex	Polyimide	Saran	Polyvinyl chloride
Hyzod	Polycarbonate	Sintra	Polyvinyl chloride
Implex	Acrylic (rubber-modified)	Styrofoam	Polystyrene
Kapton	Polyimide	Teflon	Fluorocarbon
Kevlar	Aramid	Torlon	Polyimide
Kodel	Polyester	Tygon	Polyvinyl chloride
Kydex	Acrylic-polyvinyl chloride	Ultem	Polyetherimide
Kynar	Polyvinylidene fluoride	Vespel	Polyimide
Lexan	Polycarbonate	Zerlon	Styrene-methylmethacrylate
Lucite	Acrylic	Zytel	Polyamide
Mylar	Polyester		

period (e.g., bearings, cams, gears, bushings, and rollers), 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 [for example, a melting point of 327°C for Teflon], chemicals, weather, and electricity. They also have unique nonadhesive 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 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, rollers, 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 are versatile. They 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, rollers, 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 (PVC) has a wide range of properties, is inexpensive and water resistant, 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 (for example, in pipes and conduits). *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 electrical 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 7.6) 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 composition. Generally, they weather well, possess excellent electrical properties over a wide range of humidity and temperature, and resist chemicals and heat (see also Section 7.9). Typical applications include electrical components requiring strength at elevated temperatures, oven gaskets, heat seals, and waterproof materials.

Health Hazards. There is increasing concern that 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, or BPA, a widely produced and used chemical. Numerous worldwide investigations are being conducted to determine whether there is any link to human diseases, specifically heart disease and diabetes.

EXAMPLE 7.3 Materials for a Refrigerator Door Liner

In the selection of 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, 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.

An extensive study, considering all of the factors involved, identified two candidate materials for door liners: ABS (acrylonitrile-butadiene-styrene) and HIPS (high-impact polystyrene). One aspect of the study involved the effect of vegetable oils, such as from salad dressing stored in the door shelf, on the strength of those plastics. Experiments showed that the presence of vegetable oils significantly reduced the load-bearing capacity of HIPS. It was found that HIPS becomes brittle in the presence of oils (solvent–stress cracking), whereas ABS is not affected to any significant extent.

7.8 Biodegradable Plastics

Plastic wastes contribute about 10% of municipal solid waste by weight; 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 great concern over environmental issues regarding the disposal of plastic products and the shortage of landfills, major efforts are underway to develop completely biodegradable plastics. The first attempts were made in the 1980s as a possible solution to roadside litter.

Traditionally, most plastic products have been made from synthetic polymers that are derived from nonrenewable natural resources, are not biodegradable, and are difficult to recycle. **Biodegradability** means that microbial species in the environment (e.g., microorganisms in soil and water) will degrade a portion of (or even the entire) polymeric material under the proper environmental conditions and 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. Consequently, only a portion of them may be truly biodegradable.

Three different *biodegradable plastics* have thus far been developed. They have different degradability characteristics, and they degrade over different periods of time (anywhere 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. Various additives and binders are blended with the starch to impart special 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 the **fermentation of sugar** (the third system), organic acids are added to a sugar feedstock. With the use of a specially developed process, 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.

Numerous attempts continue to be made to produce fully biodegradable plastics by the use of various agricultural waste (*agrowastes*), plant carbohydrates, plant proteins, and vegetable oils. Typical applications 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 applications
- Food and beverage containers (made from potato starch, limestone, cellulose, and water) that can dissolve in storm sewers and oceans without affecting marine life or wildlife.

The long-term performance of biodegradable plastics (both during their useful life cycle as products and in landfills) has not been assessed fully. There also is concern that the emphasis on biodegradability will divert attention from the issue of the *recyclability* of plastics and the efforts for *conservation* of materials and energy. A major consideration is the fact that the cost of today's biodegradable polymers is substantially higher than that of synthetic polymers. Consequently, a mixture of agricultural waste—such as hulls from corn, wheat, rice, and soy (as the major component)—and biodegradable polymers (as the minor component) is an attractive alternative.

Recycling of Plastics. Much effort continues to be expended globally on the collecting and recycling of used plastic products. Thermoplastics are recycled by remelting them and then reforming them into other products. They carry *recycling symbols*, in the shape of a triangle outlined by three clockwise arrows and having a number in the middle. These numbers correspond to 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

Recycled plastics increasingly are being used for a variety of products. For example, a recycled polyester (filled with glass fibers and minerals) is used for the engine cover for an F-series Ford pickup truck, as it has the appropriate stiffness, chemical resistance, and shape retention up to 180°C.

7.9 Elastomers (Rubbers)

Elastomers consist of a large family of amorphous polymers having a low glass-transition temperature. They have a characteristic ability to undergo large elastic deformations without rupture; also, they are soft and have a low elastic modulus. The term **elastomer** is derived from the words *elastic* and *mer*.

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). They can also be cross-linked, the best example of this being the elevated-temperature **vulcanization** of rubber with sulfur, discovered by Charles Goodyear in 1839 and named for Vulcan, the Roman god of fire. Once the elastomer is cross-linked, it cannot be reshaped (for example, an automobile tire, which is one giant molecule, cannot be softened and reshaped).

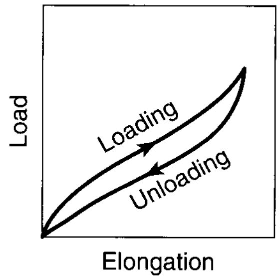


FIGURE 7.14 Typical load–elongation curve for rubbers. The clockwise loop, indicating the loading and the unloading paths, displays the hysteresis loss. Hysteresis gives rubbers the capacity to dissipate energy, damp vibration, and absorb shock loading, as is necessary in automobile tires and in vibration dampers placed under machinery.

The terms *elastomer* and *rubber* often are used interchangeably. Generally, an **elastomer** is defined as being capable of recovering substantially in shape and size after the load has been removed. A **rubber** is defined as being capable of recovering from large deformations quickly.

The hardness of elastomers, which is measured with a durometer (Section 2.6), 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 in high-friction and nonskid surfaces, protection against corrosion and abrasion, electrical insulation, and shock and vibration insulation. Examples include tires, hoses, weatherstripping, footwear, linings, gaskets, seals, printing rolls, and flooring.

One 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 desirable for absorbing vibrational energy (damping) and sound deadening.

Natural Rubber. The base 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 characteristic high friction, but low resistance, to oil, heat, ozone, and sunlight. Typical applications are tires, seals, shoe heels, 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 range of useful temperatures. Synthetic rubbers that are resistant to oil are neoprene, nitrile, urethane, and silicone. Typical applications of synthetic rubbers are tires, shock absorbers, seals, and belts.

Silicones. Silicones (see also Section 7.7) have the highest useful temperature range of elastomers (up to 315°C), but other 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 has exceptional resistance to abrasion, cutting, and tearing. Typical applications are seals, gaskets, cushioning, diaphragms for the rubber forming of sheet metals (Section 16.8), and auto body parts.

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 generally are characterized by a lower density, strength, elastic modulus, thermal and electrical conductivity, and cost; by a higher strength-to-weight ratio, higher resistance to corrosion, higher thermal expansion, and wider choice of colors and transparencies; and by a 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 and 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; they return to their original properties when cooled. Their mechanical behavior can be characterized by

various spring and damping models. Their behavior includes such phenomena as creep and stress relaxation, crazing, and water absorption. Thermosets, which are obtained by cross-linking polymer chains, do not become soft to any significant extent with increasing temperature. They are much more rigid and harder than thermoplastics, and they offer many fewer choices of color.

- 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 recyclability and biodegradability. Several formulations of biodegradable plastics are available, and others are under continued development.

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
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.

Buckley, C.P., Bucknall, C.B., and McCrum, N.G., **Principles of Polymer Engineering**, 2nd ed., Oxford University Press, 1997.

Campbell, P., **Plastics Component Design**, Industrial Press, 1996.

Campo, E.A., **Selection of Polymeric Materials**, William Andrew, Inc., 2008.

Cardelli, F., **Materials Handbook: A Concise Desk Reference**, 2nd ed., Springer, 2008.

Chanda, M., and Roy, S.K., **Plastics Technology Handbook**, 4th ed., Marcel Dekker, 2006.

Characterization and Failure Analysis of Plastics, ASM International, 2003.

Engineering Plastics and Composites, 2nd ed., ASM International, 1993.

Feldman, D., and Barbalata, A., **Synthetic Polymers: Technology, Properties, Applications**, Chapman & Hall, 1996.

Fink, J.K, **High Performance Polymers**, William Andrew, Inc., 2008.

Harper, C., **Handbook of Plastics, Elastomers, and Composites**, 4th ed., McGraw-Hill, 2003.

Goodship, V., **Introduction to Plastics Recycling**, 2nd ed., Smithers Rapra Press, 2008.

Griskey, R.G., **Polymer Process Engineering**, Chapman & Hall, 1995.

Harper, C.A., **Modern Plastics Handbook**, McGraw-Hill, 2000.

MacDermott, C.P., and Shenoy, A.V., **Selecting Thermoplastics for Engineering Applications**, 2nd ed., Marcel Dekker, 1997.

Margolis, J., **Engineering Plastics Handbook**, McGraw-Hill, 2006.

Mark, J.E. (ed.), **Physical Properties of Polymers Handbook**, 2nd ed., Springer, 2006.

Michler, G.H., and Balta-Calleja, F.J., **Mechanical Properties of Polymers Based on Nanostructure and Morphology**, CRC Press, 2006.

Mills, N., **Plastics: Microstructure and Engineering Applications**, 3rd ed., Butterworth-Heinemann, 2006.

Rudin, A., **Elements of Polymer Science and Engineering**, 2nd ed., Academic Press, 1999.

Salamone, J.C. (ed.), **Concise Polymeric Materials Encyclopedia**, CRC Press, 1999.

Sperling, L.H., **Polymeric Multicomponent Materials: An Introduction**, Wiley, 1997.

Strong, A.B., **Plastics: Materials and Processing**, 3rd ed., Prentice Hall, 2006.

Ulrich, H., **Introduction to Industrial Polymers**, Hanser, 1994.

Ward, I.M., and Sweeny, J., **An Introduction to the Mechanical Properties of Solid Polymers**, 2nd ed., Wiley, 2004.

Young, R.J., and Lovell, P., **Introduction to Polymers**, 3rd ed., CRC Press, 2008.

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.

QUALITATIVE PROBLEMS

7.14. Inspect various plastic components in an automobile, and state whether they are made of thermoplastic materials or of thermosetting plastics.

7.15. Give applications for which flammability of plastics would be of major importance.

7.16. What characteristics make polymers advantageous for applications such as gears? What characteristics are drawbacks in such applications?

7.17. What properties do elastomers have that thermoplastics in general do not have?

7.18. 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?

7.19. 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.

7.20. 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.)

7.21. Add more to the applications column in Table 7.3.

7.22. Discuss the significance of the glass-transition temperature, T_g, in engineering applications.

7.23. Describe how a rechargeable lithium battery works.

7.24. Explain how cross-linking improves the strength of polymers.

7.25. Describe the methods by which the optical properties of polymers can be altered.

7.26. How can polymers be made to conduct electricity? Explain.

7.27. Explain the reasons for which elastomers were developed.

7.28. Give several examples of plastic products or components in which creep and stress relaxation would be important considerations.

7.29. Describe your opinions regarding the recycling of plastics versus the development of plastics that are biodegradable.

7.30. Explain how you would go about determining the hardness of plastics.

7.31. Compare the values of the elastic modulus given in Table 7.1 to the values for metals given in Chapters 2, 5, and 6.

7.32. Why is there so much variation in the stiffness of products made of polymers? Explain.

7.33. Explain why thermoplastics are easier to recycle than thermosets.

7.34. Give an example of a process where crazing is desirable.

7.35. Describe the principle behind shrink wrapping.

7.36. List and explain some environmental pros and cons of using plastic shopping bags vs. paper bags.

7.37. 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.

7.38. How can you tell whether a part is made of a thermoplastic or a thermoset?

7.39. As you know, there are plastic paper clips available in various colors. Why are there no plastic staples?

7.40. 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.

7.41. 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.

QUANTITATIVE PROBLEMS

▶**7.42.** 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.

7.43. 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.

▶**7.44.** A rectangular cantilever beam 100 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.

7.45. Estimate the number of molecules in a typical automobile tire. Estimate the number of atoms in the tire.

▶**7.46.** Using strength and density data, determine the minimum weight of a 2-m-long tension member that must support a load of 5000 N if it is manufactured from (a) high-molecular-weight polyethylene, (b) polyester, (c) rigid PVC, (d) ABS, (e) polystyrene, and (e) reinforced nylon.

▶**7.47.** Plot the following for any five polymers described in this chapter: (a) UTS vs. density and (b) elastic modulus vs. UTS. Where appropriate, plot a range of values.

SYNTHESIS, DESIGN, AND PROJECTS

7.48. Describe the design considerations involved in replacing a metal beverage container with one made of plastic.

7.49. 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.

7.50. 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.) which shows the types of plastic that can be used to make these products.

7.51. Make a list of products or parts that currently are not made of plastics and offer possible reasons why they are not.

7.52. Review the three curves shown in Fig. 7.10 and give some applications for each type of behavior. Explain your choices.

7.53. Repeat Problem 7.52 for the curves shown in Fig. 7.12.

7.54. 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.

7.55. 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.)

7.56. Add a column to Table 7.1 which describes the appearance of these plastics, including available colors and opaqueness.

7.57. 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, Graphite, Diamond, and Nanomaterials: Structure, General Properties, and Applications

EXAMPLES:

- Ceramics, glass, and various forms of carbon present combinations of physical and mechanical properties that cannot be obtained with other metallic or nonmetallic materials and thus have major industrial applications.
- 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; this chapter describes the general properties and typical uses of glass.
- Various forms of carbon are commercially important; graphite is the most common, with numerous uses, including as reinforcement in composite materials, as electrodes for electrical discharge machining, and as a 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 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:

a. An electrical insulator to be used at high temperatures
b. Floor tiles to resist spills, scuffing, and abrasion
c. A transparent baking dish
d. Small ball bearings that are light, rigid, hard, and resist high temperatures
e. Automobile windshields that are hard, abrasion resistant, and transparent.

From these few examples, it is apparent that the properties required include high-temperature strength; hardness; inertness to chemicals, foods, and the environment; resistance to wear and corrosion; and low electrical and thermal conductivity.

This chapter describes the general characteristics and applications of those ceramics, glasses, and glass ceramics that are of importance in engineering applications and in manufacturing. Because of their unique properties and uses, the various forms of carbon—namely, graphite, diamond, and carbon nanotubes—are also discussed here. The manufacturing of ceramic and of glass components and various shaping and finishing operations are detailed in Chapter 18. Composites, which are another important group of materials, 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 words *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 elements, a wide variety of ceramics now is 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. Ceramics have been used for many years in automotive spark plugs, both as an electrical insulator and for their high-temperature strength. They have become increasingly important in tool and die materials, heat engines, and automotive components (such as exhaust-port liners, coated pistons, and cylinder liners).

Ceramics can be divided into two general categories:

1. **Traditional ceramics**, such as whiteware, tiles, brick, sewer pipe, pottery, and abrasive wheels.
2. **Industrial ceramics** (also called **engineering, high-tech,** or **fine ceramics**), such as turbine, automotive, and aerospace components (Fig. 8.1); heat exchangers; semiconductors; seals; prosthetics; and cutting tools.

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** (see Section 1.2) and as such are much stronger

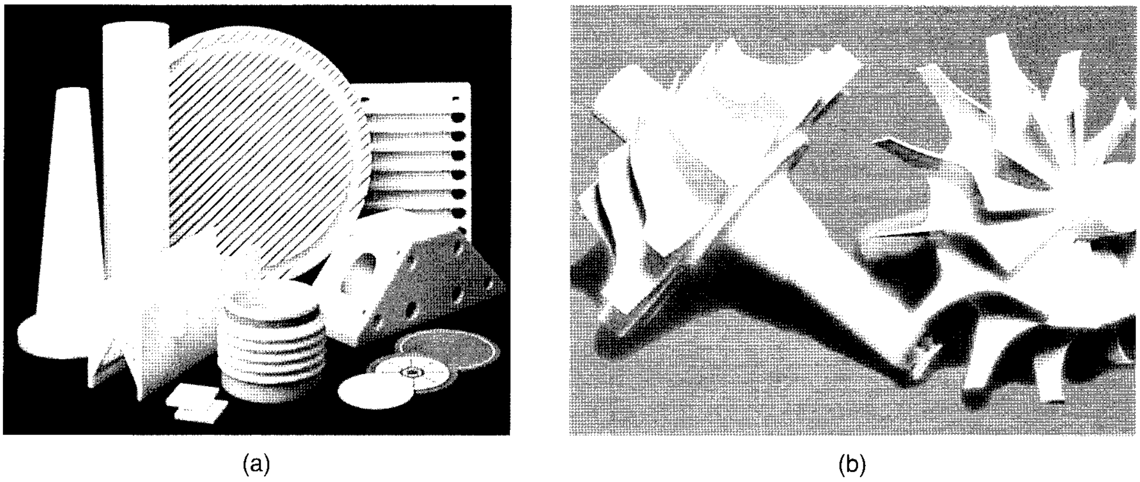

(a) (b)

FIGURE 8.1 A variety of ceramic components. (a) High-strength alumina for high-temperature applications. (b) Gas-turbine rotors made of silicon nitride. *Source*: Courtesy of Wesgo Div., GTE.

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 in *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), a white clay consisting of silicate of aluminum with alternating weakly bonded layers of silicon and aluminum ions. When added to kaolinite, water attaches itself to the layers (*adsorption*). This makes the layers slippery and gives wet clay both its well-known softness and the plastic properties (*hydroplasticity*) that make it formable.

Other major raw materials for ceramics that are found in nature are *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 kinds, which have to be removed prior to further processing of the materials into useful products with reliable performance.

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 possesses nonuniform properties; as a result, its performance also varies. Aluminum oxide, silicon carbide, and many other ceramics are now manufactured almost totally synthetically, so that their quality can be controlled. First made in 1893, synthetic aluminum oxide is obtained from the fusion of molten bauxite (an aluminum-oxide ore that is the principal source of aluminum), iron filings, and coke in electric furnaces. The cooled product is crushed and graded by size by passing the resulting particles through standard screens. Aluminum oxide can 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;* they are 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 good toughness; good 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 does zirconia. It is obtained by doping 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 dispersion media for aerospace coatings, for automotive primers and topcoats, and for fine glossy print on flexible food packaging.

TABLE 8.1

Types and General Characteristics of Ceramics

Type	General characteristics
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.
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.
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.
Sialon	Consists of silicon nitrides and other oxides and carbides; used as cutting tools.
Cermets	Consist of oxides, carbides, and nitrides; used in high-temperature applications.
Silica	High-temperature resistance; quartz exhibits piezoelectric effect; silicates containing various oxides are used in high-temperature nonstructural applications.
Glasses	Contain at least 50 percent 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.
Graphite	Crystalline form of carbon; high electrical and thermal conductivity; good thermal-shock resistance.
Diamond	Hardest substance known; available as single crystal or in polycrystalline form; used as cutting tools and abrasives and as dies for fine wire drawing.
Carbon nanotubes	Unique crystalline form of graphite, with high electrical and thermal conductivity; under investigation for MEMS and microelectronics applications and in composite materials.

Two important characteristics of PSZ are its coefficient of thermal expansion (which is only about 20% lower than that of cast iron) and its thermal conductivity (which is about one-third that of other ceramics). Consequently, PSZ is very 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.

EXAMPLE 8.1 Ceramic Knives

The use of ceramics now is being extended to knives, generally made of zirconium oxide. Ceramic knives are produced by a process (described in Section 18.2) that starts with a ceramic powder mixed with various binders and compacted (molded) into blanks under high pressure. The blanks are then fired (sintered) at temperatures above 1000°C for several days. 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 with 6 for hardened steel and a maximum of 10 for diamond.

Among the advantages ceramic knives have over steel knives are the following: (a) Because of their very high hardness and wear resistance, ceramic knives can last months and even years before sharpening. (b) They are chemically inert; consequently, they do not stain, food does not stick to them (hence, they are easy to clean), and they 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 the sharp edges of the knife can chip. Also, they should be used only for cutting (not for prying), and in cutting meat, contact with bones is not advisable. Knives have to be sharpened at the factory to a precise edge shape, using diamond grinding wheels. Ceramic knives are more expensive than steel knives, typically ranging from $60 for a 75-mm paring knife to $250 for a 150-mm serrated knife.

Source: Courtesy of Kyocera Corporation.

8.2.3 Other Ceramics

The major other types of ceramics may be classified as follows:

Carbides. Typical *carbides* are those made of tungsten and titanium (used as cutting tools and die materials) and of silicon (used as an abrasive, especially in grinding wheels). The following are some examples of carbides:

- **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 the binder and is not as tough as tungsten carbide.
- **Silicon carbide** (SiC) has good resistance to wear, thermal shock, and corrosion. It has a low friction coefficient and retains strength at elevated temperatures. Thus, it is suitable for high-temperature components in heat engines and also is used as an abrasive. 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.

Nitrides. Another class of ceramics is the nitrides, examples of which are as follows:

- **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 by means of techniques similar to those used in making synthetic diamond (see 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. Consequently, it resists thermal shock. It is suitable for high-temperature structural applications, such as in automotive-engine and gas-turbine components, cam-follower rollers, bearings, sandblast nozzles, and components for the paper industry.

Sialon. Sialon (derived from *si*licon, *al*uminum, *o*xygen, and *n*itrogen) 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. Presently, it 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. They have been developed for high-temperature applications, such as nozzles for jet engines and brakes for aircraft. Cermets can be regarded as composite materials 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 (e.g., 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 good thermal-shock resistance. However, it also has a very low strength and fatigue life; thus, 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 properties of ceramics, the particle size in ceramics has been reduced by means of various techniques, such as gas condensation. Called *nanoceramics* or **nanophase ceramics,** these materials consist of atomic clusters containing a few thousand atoms. Control of particle size, distribution, and contamination are important in nanoceramics, which exhibit ductility at significantly lower temperatures than do conventional ceramics and are stronger and easier to fabricate and to machine with fewer flaws. Present applications are in the automotive industry (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 better tensile strength and creep resistance. (See also *nanomaterials* in Section 6.16.)

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; a high elastic modulus; low toughness, density, and thermal expansion; and low thermal and electrical conductivity. However, because of the wide variety of material compositions and grain sizes, the mechanical and physical properties of ceramics vary considerably.

Properties of ceramics 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. The general mechanical and physical properties are described next.

8.3.1 Mechanical Properties

The mechanical properties of several engineering ceramics are presented in Table 8.2. Note that their strength in tension (transverse rupture strength, Section 2.5) is approximately one order of magnitude lower than their compressive strength. The reason for this is their sensitivity to cracks, impurities, and porosity. Such defects lead to the initiation and propagation of cracks under tensile stresses and significantly reduce the tensile strength of the ceramic. Thus, reproducibility and reliability (acceptable performance over a specified period) is an important aspect 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

$$\text{UTS} = \text{UTS}_o e^{-nP}, \tag{8.1}$$

where P is the volume fraction of pores in the solid (thus, if the porosity is 15%, then $P = 0.15$), UTS_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 its porosity by the expression

$$E \approx E_o(1 - 1.9P + 0.9P^2), \tag{8.2}$$

where E_o is the elastic modulus at zero porosity.

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. In addition to undergoing fatigue failure under cyclic loading, ceramics (particularly glasses) exhibit a phenomenon called **static fatigue**. When subjected to a static tensile load over time, these materials suddenly may fail. This phenomenon 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.

Ceramic components that are to be subjected to tensile stresses may be *prestressed* in much the same way that concrete is prestressed. Prestressing the shaped

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 subjects them to compressive stresses. The methods used include the following:

- Heat treatment and chemical tempering
- Laser treatment of surfaces
- Coating with ceramics having different thermal-expansion coefficients
- Surface-finishing operations (such as grinding) in which compressive residual stresses are induced on the surfaces.

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.

Thermal conductivity in ceramics varies by as much as three orders of magnitude (depending on their composition), whereas in metals it varies by only one order. Like that of 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

$$k = k_o(1 - P), \tag{8.3}$$

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 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.

A familiar example that illustrates the importance of low thermal expansion is heat-resistant ceramics for cookware and electric stove tops. (See also *glass ceramics,* Section 8.5.) They can sustain high thermal gradients, from hot to cold and vice versa. Moreover, the similar thermal expansion of both ceramics and metals is an important reason for the use of ceramic components in heat engines. The fact that the thermal conductivity of partially stabilized zirconia components is close to that of the cast iron in engine blocks is an additional advantage to the use of PSZ in heat engines.

Another characteristic is the **anisotropy of thermal expansion** of oxide ceramics (like that exhibited by hexagonal close-packed metals), wherein the thermal expansion varies in different directions in the ceramic (by as much as 50% for quartz). This behavior causes thermal stresses that can lead to cracking of the ceramic component.

The *optical* properties of ceramics can be controlled by using various formulations and by controlling the 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 a translucent gray.) Porosity influences the optical properties of ceramics in much the same way as air trapped in ice cubes, making the material less transparent and giving it 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 act 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 very attractive for high-temperature applications. The higher operating temperatures made possible by the use of ceramic components mean more efficient combustion of fuel and 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%.

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 also are candidates for ceramics (Section 25.3). Furthermore, the high elastic modulus of ceramics makes them attractive for improving the stiffness of machines, while reducing the weight. Their high resistance to wear 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.

EXAMPLE 8.2 Ceramic Gun Barrels

The wear resistance and low density of ceramics have led to research into their use as liners for gun barrels. Their limited success has led to more recent developments in making composite ceramic gun barrels, which have improved performance over traditional steel barrels. The 50-caliber zirconia ceramic barrel is formed in several separate segments, each 150–200 mm long and with a wall thickness of 3.75 mm, by the shaping and sintering processes described in Chapter 17.

The segments subsequently are machined to the required dimensions and surface finish. Zirconia has been chosen for its high toughness, flexural strength, specific heat, operating temperature, and very low thermal conductivity. The thermal properties are important to the performance of the barrel and the bullet.

The separate ceramic segments are then joined, and the barrel is wrapped with a carbon-fiber/polymer-matrix composite that subjects the ceramic barrel to a compressive stress of 690 MPa, thus greatly improving its capacity to withstand tensile stresses developed during firing. Finally, the inside of the barrel is rifled (cut to produce internal spiral grooves that give rotation to the exiting bullet for gyroscopic stability) and fitted to a breech.

Source: Courtesy of K.H. Kohnken, Surface Conversion Technologies, Inc., Cumming, Georgia.

EXAMPLE 8.3 Ceramic Ball and Roller Bearings

Silicon-nitride ceramic ball and roller bearings are used when high temperature, high speed, or marginally lubricated conditions occur. 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. 8.2). Examples of machines utilizing ceramic and hybrid bearings include high-performance machine tool spindles, metal-can seaming heads, high-speed flow meters, and the Space Shuttle's main booster rocket's liquid oxygen and hydrogen pumps.

The ceramic spheres have a diameter tolerance of 0.13 μm and a surface roughness of 0.02 μm. They have high wear resistance, high fracture toughness, perform well with little or no lubrication, and have low density. 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 can be produced.

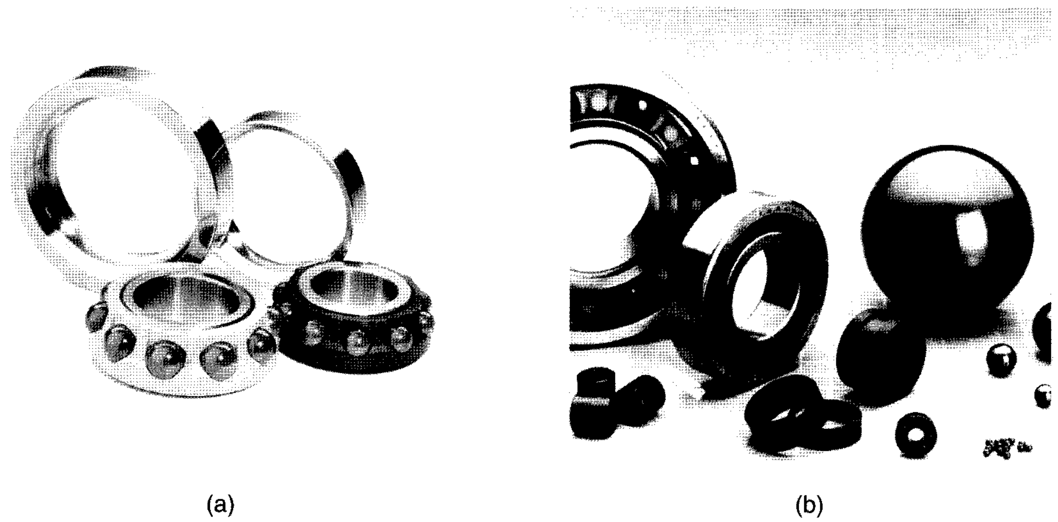

(a) (b)

FIGURE 8.2 A selection of ceramic bearings and races. *Source*: Courtesy of The Timken Company.

Bioceramics. Because of their strength and inertness, ceramics are 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 (likewise with porous titanium implants) and develop a strong bond with structural integrity.

8.4 Glasses

Glass is an amorphous solid with the structure of a liquid. It has been *supercooled* (cooled 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 condition without crystallizing. Glass has no distinct melting or freezing point; thus, its behavior is similar to that of amorphous alloys (see *metallic glasses,* Section 6.14) and amorphous polymers (Section 7.2).

Glass beads first were produced in about 2000 B.C., and the art of glassblowing started in about 200 B.C. Silica was used for all glass products until the late 1600s. Rapid developments in glasses began in the early 1900s. Currently, there are some 750 different types of commercially available glasses, with applications ranging from window glass to glass for containers, cookware, lighting, and view screens for television sets and mobile phones, and to glasses with special mechanical, electrical, high-temperature, antichemical, 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).

All glasses contain at least 50% silica, which is known as a **glass former.** The composition and properties of glasses can be modified greatly 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 type (Table 8.3):

- **Soda-lime glass** (the most common type)
- **Lead-alkali glass**
- **Borosilicate glass**
- **Aluminosilicate glass**
- **96%-silica glass**
- **Fused silica glass.**

Glasses also are classified as colored, opaque (white and translucent), multiform (a variety of shapes), optical, photochromatic (darkens when exposed to light, as in sunglasses), photosensitive (changing from 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, the rest hard.

TABLE 8.3

Properties of Various Glasses

	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

8.4.2 Mechanical Properties

The behavior of glass, like that of most ceramics, generally is regarded as perfectly elastic and brittle. The modulus of elasticity for commercial glasses ranges from 55 to 90 GPa, and their Poisson's ratios 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, 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 inadvertent abrading. 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 freshly drawn into fibers (**fiberglass**), its tensile strength ranges from 0.2 to 7 GPa, with an average value of about 2 GPa. These glass fibers are stronger than steel; they are used to reinforce plastics in such applications as boats, automobile bodies, furniture, and sports equipment (Tables 2.2 and 9.1).

The strength of glass usually is measured by bending it. The surface of the glass is thoroughly abraded (roughened) to ensure that the test gives a strength level that is reliable for actual service under adverse conditions. The phenomenon of **static fatigue**, observed in ceramics, also is exhibited by glasses. As a guide, if a glass item must withstand a load for 1000 hours or longer, the maximum stress that can be applied to it is approximately one-third the maximum stress that the same item can withstand during the first second of loading.

8.4.3 Physical Properties

Glasses are characterized by low thermal conductivity and high electrical resistivity and dielectric strength. Their thermal expansion coefficients are lower than those for metals and plastics and may even approach zero. For example, titanium silicate glass (a clear, synthetic high-silica glass) 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** (such as *Pyroceram,* a trade name) have a high crystalline component to their microstructure. Glass ceramics 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 first are shaped and then heat treated, with **devitrification** (recrystallization) of the glass occurs. Unlike most glasses, which are clear, glass ceramics are generally white or gray in color.

The hardness of glass ceramics ranges approximately from 520 to 650 HK. Because glass ceramics have a near-zero coefficient of thermal expansion, they have high thermal-shock resistance. They are strong, because of the absence of the porosity usually found in conventional ceramics. The properties of glass ceramics can be improved by modifying their composition and by heat-treatment techniques. First developed in 1957, glass ceramics are suitable for cookware, heat exchangers in gas-turbine engines, radomes (housings for radar antennas), and electrical and electronics applications.

8.6 Graphite

Graphite is a crystalline form of carbon having a *layered structure* with basal planes or sheets of close-packed carbon atoms (see Fig. 1.4). Consequently, graphite is weak when sheared along the layers. This characteristic, in turn, gives graphite its low frictional properties as a solid lubricant. However, its frictional properties are low only in an environment of air or moisture; in a vacuum, graphite is abrasive and a poor lubricant. Unlike in other materials, strength and stiffness of graphite increase with temperature. Amorphous graphite is known as **lampblack** (black soot) and is used as a pigment.

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, therefore, an important material for applications such as electrodes, heating elements, brushes for motors, high-temperature fixtures and furnace parts, mold materials (such as crucibles for the melting and casting of metals), and seals (Fig. 8.3). A characteristic of graphite is its resistance to chemicals; thus, it is used in filters for corrosive fluids. Also, its low absorption cross section and high scattering cross section for thermal neutrons make graphite suitable for nuclear applications. Ordinary pencil "lead" is a mixture of graphite and clay.

Graphite is available commercially in square, rectangular, and round shapes of various sizes and generally is graded in decreasing order of grain size: *industrial, fine grain,* and *micrograin.* As in ceramics, the mechanical properties of graphite improve with decreasing grain size. Micrograin graphite can be impregnated with copper. In this form, it is used for electrodes in electrical-discharge machining (Section 27.5) and for furnace fixtures. Graphite usually is processed first by molding or forming, then by oven baking, and finally by machining to the final shape.

Graphite Fibers. An important use of graphite is as fibers in reinforced plastics and composite materials, as described Section 9.2.

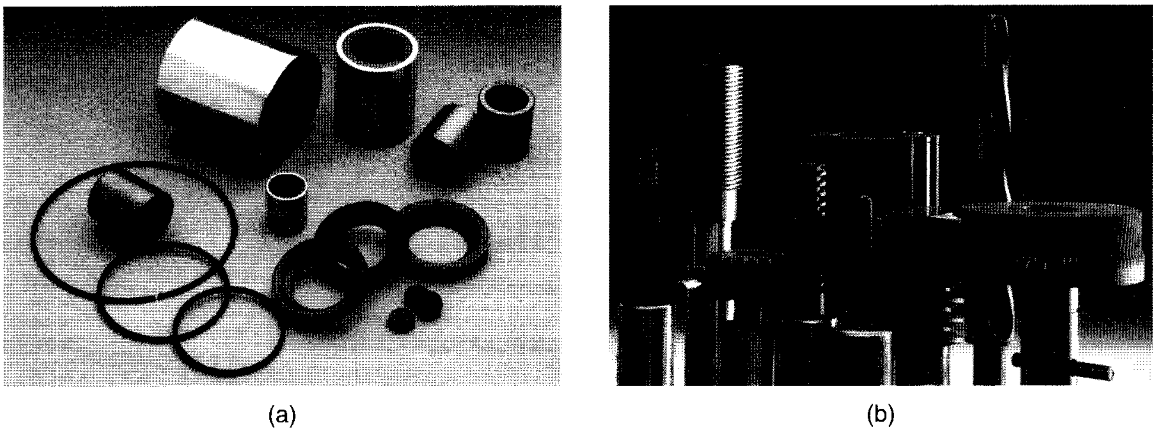

(a) (b)

FIGURE 8.3 (a) Various engineering components made of graphite. *Source*: Courtesy of Poco Graphite, Inc., a Unocal Co. (b) Examples of graphite electrodes for electrical discharge machining. *Source*: Courtesy of Unicor, Inc.

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 pitches) have low density, high thermal and electrical conductivity (but lower mechanical strength), and are much more expensive than nongraphitic foams (produced from coal or organic resins), which are highly amorphous.

These foams have a cellular microstructure with interconnected pores; thus, their mechanical properties depend on density (see also Section 8.3). They easily can be machined into various complex shapes with appropriate tooling. 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

A more recent development is the production of carbon molecules (usually C60) in the shape of a soccer balls, called **fullerenes or buckyballs**, after Buckminster 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. Fullerenes can become low-temperature superconductors when mixed with metals. Despite their promise and significant research investment, 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 by laser ablation of graphite, carbon-arc discharge, and, most often, by chemical vapor deposition. They can be single-walled (SWNTs) or multiwalled (MWNTs) nanotubes and can be doped with various species.

Carbon nanotubes have exceptional strength, which makes them attractive as a reinforcing fiber for composite materials. However, they have very low adhesion with most materials, so that delamination with a matrix limits their reinforcing effectiveness. Also, it is difficult to disperse the nanotubes properly, and their effectiveness as a reinforcement is limited if the nanotubes are clumped. However, a few products have used carbon nanotubes, such as a bicycle frame used in the 2006 Tour de France and specialty baseball bats and tennis racquets. Note that the nanotubes provide only a fraction of the reinforcement by volume or effectiveness in these products, with graphite filling the major role.

The other material 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. 8.4). Armchair nanotubes are theoretically able to carry a current density more than 1,000 times larger than silver or copper, which makes them attractive for electrical connections in nanodevices (see Section 29.5). Carbon nanotubes have been incorporated into polymers to improve their static electricity discharge capability, especially in fuel lines for automotive and aerospace applications. Other proposed uses for carbon nanotubes include storage of hydrogen for use in hydrogen-powered vehicles, flat-panel displays, catalysts, and X-ray and microwave generators. Highly sensitive sensors using aligned carbon nanotubes are now being developed for detecting deadly gases, such as sarin.

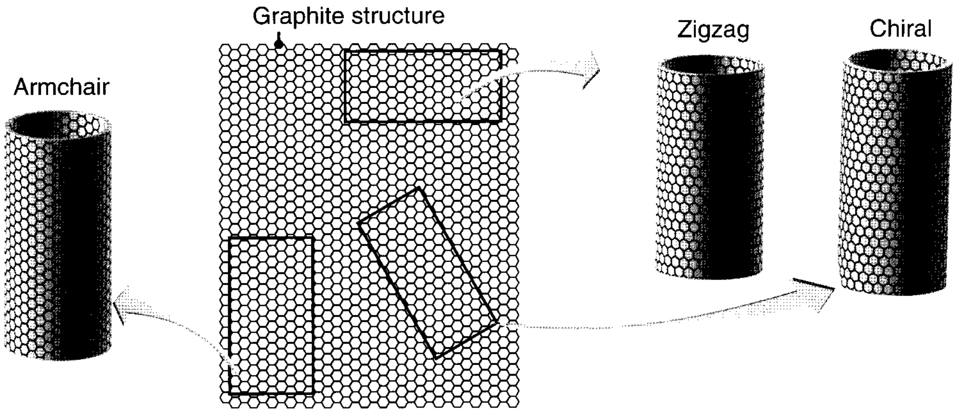


FIGURE 8.4 Forms of carbon nanotubes: armchair, zigzag, and chiral. Armchair nanotubes are noteworthy for their high electrical conductivity, whereas zigzag and chiral nanotubes are semiconductors.

8.7 Diamond

Diamond is a principal form of carbon with a covalently bonded structure. It is the hardest substance known (7000 to 8000 HK). However, it is brittle and begins to decompose in air at about 700°C, but it resists higher temperatures in nonoxidizing environments.

Synthetic (also called **industrial**) **diamond** was first made 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. Synthetic diamond is identical to natural diamond. It has superior properties because of its lack of impurities and is used extensively in industrial applications. Synthetic diamond is available in various sizes and shapes; for abrasive machining, the most common grit size is 0.01 mm in diameter. **Diamond-like carbon** also has been developed and is used as a diamond film coating, as described in Section 34.13. Diamond particles also can be *coated* with nickel, copper, or titanium for improved performance in grinding operations.

In addition to its use in jewelry, gem-quality synthetic diamond has possible 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 grinding hard materials
- Dressing of grinding wheels (i.e., sharpening of the abrasive grains)
- Dies for drawing wire less than 0.06 mm in diameter
- Coatings for cutting tools and dies.

8.8 Nanomaterials

Important developments have taken place in the production of materials as particles, fibers, wire, tube, films, and composites having features that are 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 and commercially available materials. These characteristics include high strength, hardness, ductility, toughness, wear resistance and corrosion resistance, 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 alloys, organic polymers, semiconductors, and various composites. *Nanometal–polymer hybrid* nanomaterials are also being developed for very lightweight components. Recent investigations include the development of *nanopaper*, produced from wood pulp with fibers rearranged into an entangled porous mesh, resulting in very high strength and toughness.

Synthesis methods include inert-gas condensation, sputtering, plasma synthesis, electrode position, sol–gel synthesis, and mechanical alloying or ball milling. 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. Because the synthesis of nanomaterials is done at atomic levels, their purity (on the order of 99.9999%), 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. Control of impurities in nanomaterials produced remains a challenging field. Nanoparticles may also be coated for specific purposes.

Applications of Nanomaterials. The unique properties of nanomaterials enable manufacturing of products that are strong, light, and more efficient. The following are some current and potential applications for nanomaterials:

a. Cutting tools and inserts made of nanocrystalline carbides and other ceramics.
b. Nanophase ceramics that are ductile and machinable.
c. Powders for powder-metallurgy processing.
d. Carbon nanotubes have been used in specialty bicycle frames, baseball bats, and tennis racquets. (See also Section 8.6.2.)
e. Next-generation computer chips using nanocrystalline starting materials with very high purity, better thermal conductivity, and more durable interconnections.
f. Flat-panel displays for laptop computers and televisions, made by synthesizing nanocrystalline phosphorus to improve screen resolution.
g. 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.
h. Switches, valves, motor, and pumps.
i. 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.

Because nanomaterials are very expensive to produce and process into products, their cost-effectiveness is under continued study.

Health Hazards. Because of their small size, nanoparticles can present *health hazards* by virtue of their absoption through the skin, lungs, or the digestive track. They can also penetrate human cells. There is increasing evidence that nanoparticles can pollute 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. On the other hand, they are brittle and have low toughness. Nanophase ceramics have better properties than common ceramics. Porosity in ceramics has important effects on their properties.
- Ceramics generally are categorized as either traditional 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 of mechanical, physical, and optical properties. Glass ceramics are predominantly crystalline in structure and have properties that are more desirable than those of glasses.
- Glass in bulk form has relatively low strength, but glasses can be strengthened by thermal and chemical treatments in order to obtain high strength and toughness. Glass fibers are used widely as a reinforcing material in composite materials, such as in fiber-reinforced plastics.
- Graphite, fullerenes, carbon nanotubes, and diamond are forms of carbon that display unique combinations of properties. These materials have several established and emerging applications in engineering and manufacturing. Graphite has high-temperature and electrical applications; graphite fibers are used to reinforce plastics and other composite materials.
- Diamonds (both natural and synthetic) are used as cutting tools for fine machining operations, as dies for thin-wire drawing, and as abrasives for grinding wheels. Diamond-like carbon has applications as a coating material giving improved wear resistance.
- Nanomaterials in various shapes are a more recent development, with physical, mechanical, optical, chemical, and thermal properties allowing their 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
Graphite
Industrial ceramics
Industrial diamond
Ion implantation
Nanoceramics
Nanophase ceramics
Nanotubes
Nitrides
Oxide ceramics
Partially stabilized zirconia
Porcelain
Porosity
Sialon
Silica
Static fatigue
Transformation-toughened zirconia
White ceramics
Zirconia

BIBLIOGRAPHY

Bansal, N.P. (ed.), **Handbook of Ceramic Composites**, Springer, 2005.

Barsoum, M.W., **Fundamentals of Ceramics**, Institute of Physics Publishing, 2003.

Buchanon, R.C. (ed.), **Ceramic Materials for Electronics: Processing, Properties and Applications**, 3rd ed., Dekker, 2004.

Bhushan, B. (ed.), **Springer Handbook of Nanotechnology**, 2nd ed., Springer, 2006.

Cardelli, F., **Materials Handbook: A Concise Desk Reference**, 2nd ed., Springer, 2008.

Carter, C.B., and Norton, M.G., **Ceramic Materials: Science and Engineering**, Springer, 2008.

Chiang, Y.-M., **Physical Ceramics: Principles for Ceramics Science and Engineering**, Wiley, 1995.

Edinsinghe, M.J., **An Introduction to Structural Engineering Ceramics**, Ashgate Pub. Co., 1997.

Engineered Materials Handbook, Vol. 4: **Ceramics and Glasses**, ASM International, 1991.

Gogotsi, Y., **Nanomaterials Handbook**, CRC press, 2006.

Green, D.J., **An Introduction to the Mechanical Properties of Ceramics**, Cambridge Univ. Press, 1998.

Harper, C.A. (ed.), **Handbook of Ceramics, Glasses, and Diamonds**, McGraw-Hill, 2001.

Holand, W., and Beall, G.H., **Design and Properties of Glass Ceramics**, American Chemical Society, 2001.

Mitura, S., **Nanomaterials**, Elsevier, 2000.

O'Connell, M.J. (ed), **Carbon Nanotubes: Properties and Applications**, CRC Press, 2006.

Park, J., **Bioceramics: Properties, Characterization and Applications**, Springer, 2008.

Parinov, I., **Microstructure and Properties of High-Temperature Superconductors**, Springer, 2008.

Pierson, H.O., **Handbook of Carbon, Graphite, Diamond and Fullerenes: Properties, Processing and Applications**, Noyes Pub., 1993.

Prelas, M.A., Popovichi, G., and Bigelow, L.K. (eds.), **Handbook of Industrial Diamonds and Diamond Films**, Marcel Dekker, 1998.

Richerson, D.W., **Modern Ceramic Engineering: Properties, Processing, and Use in Design**, 3rd ed., Marcel Dekker, 2005.

Shelby, J.E., **Introduction to Glass Science and Technology**, Royal Society of Chemistry, 2005.

Somiya, S., Aldiner, F., Spriggs, R., Uchino, K., Kuomoto, K., and Kaneno, M. (eds.) **Handbook of Advanced Ceramics: Materials, Applications, Processing and Properties**, Academic Press, 2003.

Vollath, D., **Nanomaterials: An Introduction to Synthesis, Properties and Applications**, Wiley, 2008.

Wilks, J., and Wilks, E., **Properties and Applications of Diamond**, Butterworth-Heinemann, 1991.

REVIEW QUESTIONS

8.1. Compare the major differences between the properties of ceramics and those of metals and plastics.

8.2. List the major types of ceramics that are useful in engineering applications.

8.3. What do the following materials typically consist of (a) carbides, (b) cermets, and (c) sialon?

8.4. What is porcelain?

8.5. What is glass? Why is it called a supercooled material?

8.6. What is devitrification?

8.7. List the major types of glasses and their applications.

8.8. What is static fatigue? What is its significance?

8.9. Describe the major uses of graphite.

8.10. How are alumina ceramics produced?

8.11. What features of PSZ differentiate it from other ceramics?

8.12. What are buckyballs?

8.13. List the major uses of diamond.

8.14. What is a carbon nanotube? Explain why they are not as prevalent as other forms of carbon.

QUALITATIVE PROBLEMS

8.15. Explain why ceramics are weaker in tension than in compression.

8.16. What are the advantages of cermets? Suggest applications in addition to those given in this chapter.

8.17. Explain why the electrical and thermal conductivity of ceramics decreases with increasing porosity.

8.18. 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?

8.19. Describe the reasons that have encouraged the development of synthetic diamond.

8.20. Explain why the mechanical properties of ceramics generally differ from those of metals.

8.21. Explain how ceramics can be made tougher.

8.22. List and describe situations in which static fatigue can be important.

8.23. What properties are important in making heat-resistant ceramics for use on oven tops? Why?

8.24. A large variety of glasses is now available. Why is this so?

8.25. What is the difference between the structure of graphite and that of diamond? Is it important? Explain.

8.26. List and explain materials that are suitable for use as a coffee cup.

8.27. Aluminum oxide and PSZ are described as white in appearance. Can they be colored? If so, how would you accomplish this?

8.28. Why does the strength of a ceramic part depend on its size?

8.29. In old castles and churches in Europe, the glass windows display pronounced ripples and are thicker at the bottom than at the top. Explain.

8.30. Ceramics are hard and strong in both compression and shear. Why, then, are they not used as nails or other fasteners? Explain.

8.31. 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.

8.32. Investigate and list the ceramics used for high-temperature superconductor applications.

QUANTITATIVE PROBLEMS

8.33. In a fully dense ceramic, $\text{UTS}_o = 180$ MPa and $E_o = 300$ GPa. What are these properties at 15% porosity for values of $n = 4$, 5, 6, and 7, respectively?

8.34. Plot the UTS, E, and k values for ceramics as a function of porosity P, describe and explain the trends that you observe in their behavior.

8.35. What would be the tensile strength and the modulus of elasticity of the ceramic in Problem 8.33 for porosities of 20% and 40%, for the four n values given?

8.36. Calculate the thermal conductivities for ceramics at porosities of 5%, 10%, and 20% for $k_o = 0.7$ W/m·K.

8.37. A ceramic has $k_o = 0.80$ 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.

8.38. 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

8.39. 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.

8.40. 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?

8.41. Describe applications in which a ceramic material with a near-zero coefficient of thermal expansion would be desirable.

8.42. The modulus of elasticity of ceramics is typically maintained at elevated temperatures. What engineering applications could benefit from this characteristic?

8.43. List and discuss the factors that you would take into account when replacing a metal component with a ceramic component in a specific product.

8.44. 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.

8.45. Assume that the cantilever beam in Quantitative Problem 3.16 in Chapter 3 is made of ceramic. How different would your answer to that problem be if the beam were made of metal? Explain clearly, giving numerical examples.

8.46. 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.

8.47. Ceramic pistons are being considered for high-speed combustion engines. List the benefits and concerns that you would have regarding this application.

8.48. 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.

8.49. 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.

CHAPTER 9

Composite Materials: Structure, General Properties, and Applications

EXAMPLES:

- With their attractive properties—especially high strength-to-weight and stiffness-to-weight ratios—reinforced plastics and composites are among the most important engineered materials.
- Composites are widely used as structural components, especially in the aerospace industry, 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 a discussion of the selection and applications of a variety of reinforced plastics and composites.

9.1 Introduction

A **composite material** is 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 **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, which was 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 composite.

Composites, in the most general sense, can be thought of as a wide variety of materials, such as cermets (Section 8.2), two-phase alloys (Section 4.2), natural materials such as wood and bone, and general reinforced or combined materials

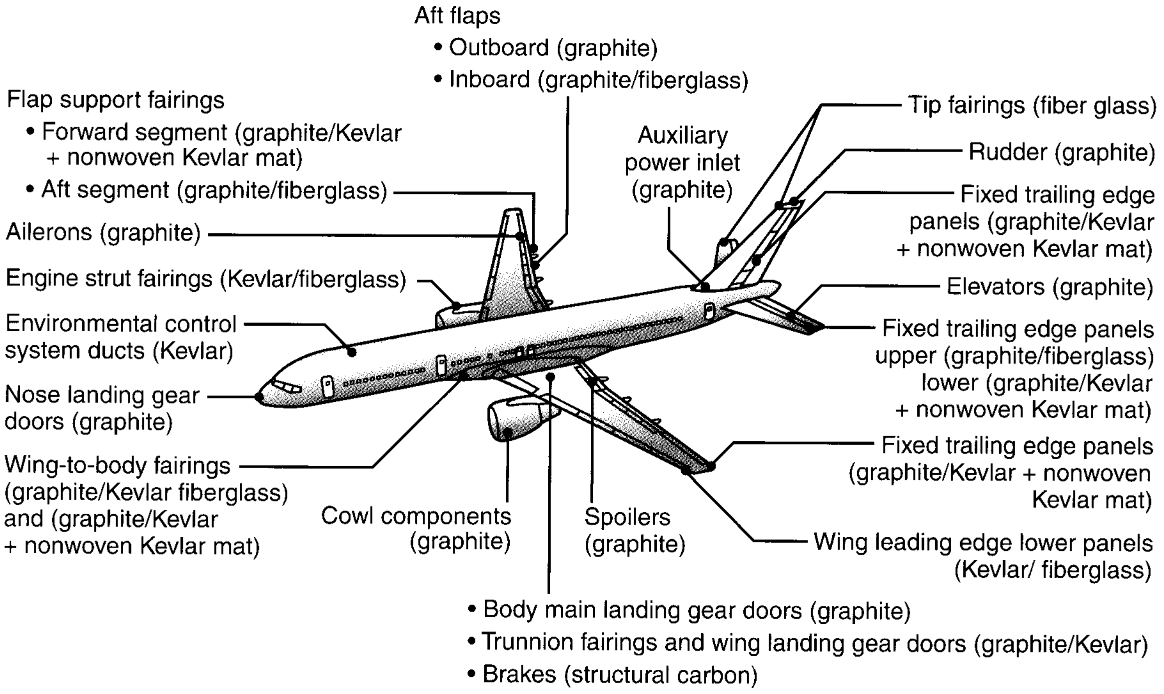


FIGURE 9.1 Application of advanced composite materials in Boeing 757-200 commercial aircraft. *Source*: Courtesy of Boeing Commercial Airplane Company.

such as steel-wire reinforced automobile tires. These materials should be recognized as composite materials, even if they are not emphasized in this chapter.

This chapter describes the structure of composite materials that have a continuous matrix and fibers as the reinforcement. The chapter also discusses the types and characteristics of reinforcing fibers used and typical major applications of these materials. The processing and the 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 shown in Fig. 9.2. These fibers are strong and stiff (Table 9.1), and they 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, greater toughness, and higher creep resistance than those made of unreinforced plastics. Reinforced-plastic structures are relatively easy to design, fabricate, and repair.

By themselves, the fibers in reinforced plastics have little structural value; they have stiffness in their longitudinal direction, but no transverse stiffness or strength. The plastic matrix is less strong and less stiff, but it is tougher and often more chemically inert, than the fibers. Reinforced plastics combine the advantages of each of the two constituents. The percentage of fibers (by volume) in reinforced plastics usually ranges between 10 and 60%. Practically, the percentage of fiber in a matrix is limited by the average distance between adjacent fibers or particles. The highest

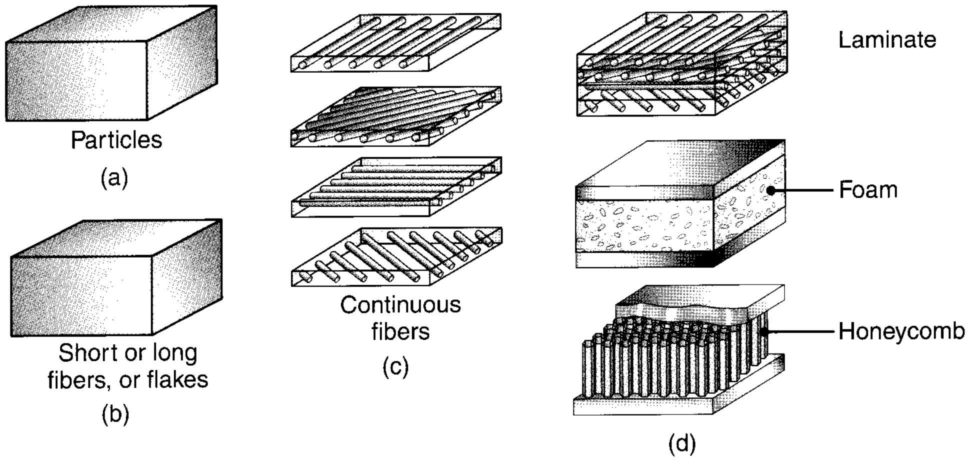


FIGURE 9.2 Schematic illustration of methods of reinforcing plastics (matrix) with (a) particles, (b) short or long fibers or flakes, and (c) continuous fibers. The laminate structures shown in (d) can be produced from layers of continuous fibers or sandwich structures using a foam or honeycomb core (see also Fig. 16.50).

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

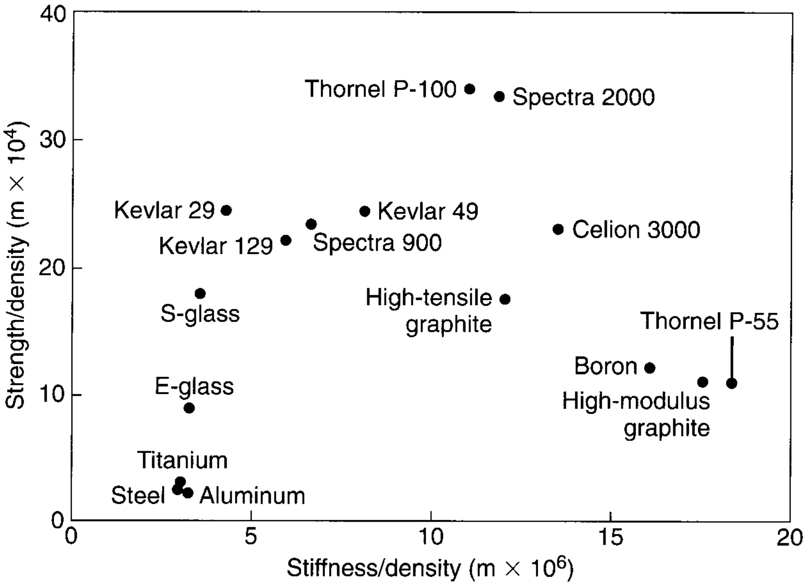


FIGURE 9.3 Specific tensile strength (tensile-strength-to-density ratio) and specific tensile modulus (modulus-of-elasticity-to-density ratio) for various fibers used in reinforced plastics. Note the wide range of specific strengths and stiffnesses available.

practical fiber content is 65%; higher percentages generally result in poorer structural properties.

When more than one type of fiber is used in a reinforced plastic, the composite is called a **hybrid**. Hybrids generally have even better properties than single-fiber composites, but they are more costly. Glass or carbon fiber-reinforced hybrid plastics have been developed for high-temperature applications up to about 300°C. Generally, however, these fibers are brittle, abrasive, lack toughness, and can degrade chemically when exposed to the environment. Also, the properties of fibers can vary significantly with the quality of the material and the method of processing.

9.2.1 Reinforcing Fibers

Glass, carbon, ceramics, aramids, and boron are the most common reinforcing fibers for polymer-matrix composites (Table 9.2).

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 by volume. Fibers are made by drawing molten glass through small openings in a platinum die (see Section 18.3.4). The molten glass is then elongated mechanically, cooled, and wound on a roll. A protective coating or sizing may be applied to facilitate their passage through the machinery. The glass fibers are treated with **silane** (a silicon hydride) for improved wetting and bonding between the fiber and the matrix.

The following are the principal types of glass fibers:

- **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, offering higher resistance to elevated temperatures and acid corrosion than does the E glass.

TABLE 9.2

Typical Properties of Reinforcing Fibers

Type	Tensile strength (MPa)	Elastic modulus (GPa)	Density (kg/m^3)	Relative cost
Boron	3500	380	2600	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	2800	62	1440	High
49	2800	117	1440	High
129	3200	85	1440	High
Nextel				
312	1630	135	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.

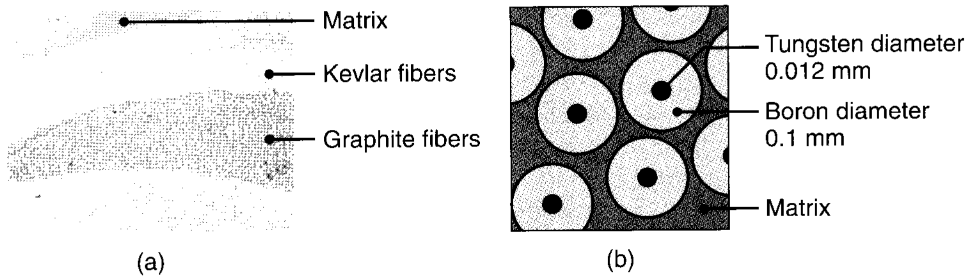


FIGURE 9.4 (a) Cross section of a tennis racket, showing graphite and aramid (Kevlar) reinforcing fibers. *Source*: Courtesy of J. Dvorak, Mercury Marine Corporation; and F. Garrett, Wilson Sporting Goods Co. (b) Cross section of boron fiber-reinforced composite material.

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 product is called **carbon-fiber reinforced plastic (CFRP)**. Although the words often are used interchangeably, the difference between *carbon* and *graphite* depends on the purity of the material and the temperature at which it was processed. Carbon fibers are at least 90% carbon; graphite fibers are usually more than 99% carbon. 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 *low*, *intermediate*, *high*, and *very high modulus*. Some trade names for carbon fibers are Celion and Thornel; see Fig. 9.3.

All carbon fibers are made by **pyrolysis** of organic **precursors**, commonly *polyacrylonitrile* (PAN) because of its low cost. *Rayon* and *pitch* (the residue from catalytic crackers in petroleum refining) also can be used as precursors. Pyrolysis is the process of inducing chemical changes by heat—for instance, by burning a length of yarn and causing the material to carbonize and become black in color. 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 point, the fibers are *carburized*: They are exposed to elevated temperatures to expel the hydrogen (dehydrogenation) and the nitrogen (denitrogenation) from the PAN. The temperatures for carbonizing range up to about 1500°C; for graphitizing up to 3000°C. PAN fibers, which have an oriented long-chain molecules, are white and have a density of 1170 kg/m^3.

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) using a continuous electroplating process; the coating is typically 0.5-μm thick on a 7-μm-diameter graphite fiber core. Available in a chopped or a continuous form, the conductive fibers are incorporated directly into injection-molded plastic parts. Applications include electromagnetic and radio-frequency shielding and lightning-strike protection.

Ceramic Fibers. One family of ceramic fibers is *Nextel*, a trade name. Nextels are oval in cross section and consist of alumina, silica, and boric oxide. Typical mechanical properties are given in Table 9.2. These fibers have low elongation, low thermal conductiity, and good chemical resistance, in addition to being suitable for high-temperature applications.

Polymer Fibers. Polymer fibers may 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 and have very high specific strength (Fig. 9.3). Aramids can undergo some plastic deformation prior to fracture and, hence, have higher toughness than brittle fibers. However, aramids absorb moisture (*hygroscopic*), thus degrading their properties and complicating their application.

Another high-performance polyethylene fiber is *Spectra* (a trade name); it has an ultra-high molecular weight and high molecular-chain orientation. Spectra, a bright white polyethylene, has better abrasion resistance and flexural-fatigue resistance than aramid fibers 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 fibers (see Table 9.2). However, a low melting point and poor adhesion characteristics as compared to other polymers are its major limitations to applications.

Polymer fibers are made by two processes: melt spinning and dry spinning (see Section 19.2.2). **Melt spinning** involves extruding a liquid polymer through small holes in a die (*spinnerets*). The fibers are then cooled before being gathered and wound onto bobbins. The fibers may be stretched to further orient and strengthen the polymer. In **dry spinning**, the polymer is dissolved in a liquid solution to form a partially oriented liquid-crystal form. As the polymer passes through the spinnerette, it is oriented further, and at this point, the fibers are washed, dried, and wound. Aramids are oriented in solution and are oriented fully when they pass through the spinnerette and therefore, do not need to be drawn any further.

Boron Fibers. These fibers consist of boron deposited (by chemical vapor-deposition techniques) onto tungsten fibers (Fig. 9.4b). Boron also can be deposited onto carbon fibers. Boron fibers have desirable properties, such as 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 also are expensive.

Other Fibers. 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). Whiskers are tiny needle-like single crystals that grow to 1 to 10 μm in diameter, with high aspect ratios (the ratio of fiber length to its diameter) ranging from 100 to 15,000. Because of their small size, either whiskers are free of imperfections, or the imperfections they contain do not significantly affect their strength, which approaches the theoretical strength of the material (size effect). The elastic moduli of whiskers range between 400 and 700 GPa, and their tensile strength is on the order of 15 to 20 GPa, although this varies with the material.

9.2.2 Fiber Size and Length

Fibers are very strong and stiff in tension. The reason is that the molecules in the fibers are oriented in the longitudinal direction, and their cross sections are 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, for example, can have tensile strengths as high as 4600 MPa, whereas the strength of glass in bulk form (Section 8.4) is much lower.

Fibers generally 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 the average fiber length, then it is called a *short fiber*. If no such improvement in properties occurs, it is called a *long fiber*. Short fibers typically have aspect ratios between 20 and 60, long fibers between 200 and 500.

Reinforcing elements also may be in the form of *chopped fibers*, *particles*, or *flakes*, or in the form of continuous *roving* (slightly twisted strands) fibers, *woven fabric* (similar to cloth), *yarn* (twisted strands), and *mats* of various combinations. Various hybrid yarns also are available.

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 the environment.
3. Reduce the propagation of cracks in the composite by virtue of the greater ductility and toughness of the plastic matrix.

Matrix materials are usually *thermoplastics* or *thermosets* and commonly consist of epoxy, polyester, phenolic, fluorocarbon, polyethersulfone, or silicon. The most commonly used are the epoxies (80% of all reinforced plastics) and the polyesters (less expensive than the epoxies). Polyimides, which resist exposure to temperatures in excess of 300°C, continue to be developed 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 material, the length of the fibers, and the volume fraction (percentage) of the reinforcing material. Short fibers are less

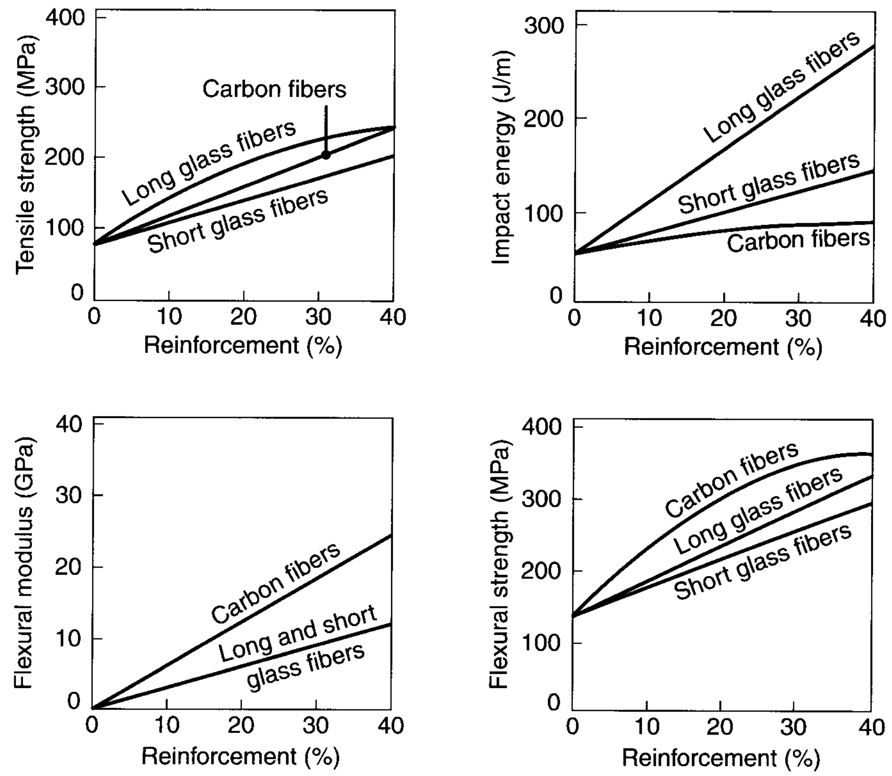
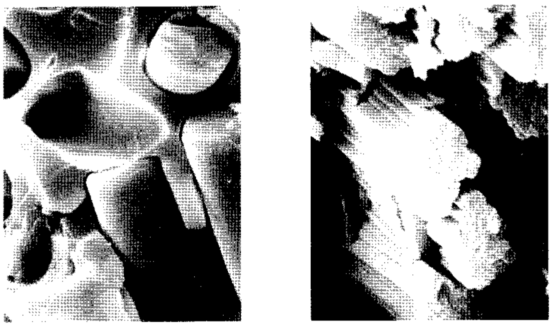


FIGURE 9.5 The effect of the type of fiber on various properties of fiber-reinforced nylon (6,6). *Source*: Courtesy of NASA.

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; thus, they are commonly 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), to make processing easier, and to reduce production costs.

A critical factor in reinforced plastics is the strength of the bond between the fiber and the polymer matrix, because the load is transmitted through the fiber–matrix interface. Weak interfacial bonding causes **fiber pullout** and **delamination** of the structure, 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** (a silicon hydride) 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, shown in Figs. 9.6a and b. Note, for example, the separation between the fibers and the matrix; obviously, better adhesion between them improves the overall strength of the composite.

Generally, the highest stiffness and strength in reinforced plastics are obtained when the fibers are aligned in the direction of the tension force. The composite is then highly anisotropic (Fig. 9.7); that is, it has properties, such as strength and stiffness, that depend on direction. As a result, other properties, such as stiffness, creep resistance, thermal and electrical conductivity, and thermal expansion, also are anisotropic. The transverse properties of such a unidirectionally reinforced structure are much lower than the longitudinal properties. For example, note how strong



(a) (b)

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 μm to 11 μm in diameter, are in bundles and are all aligned in the same direction. *Source*: Courtesy of L.J. Broutman, Illinois Institute of Technology, Chicago.

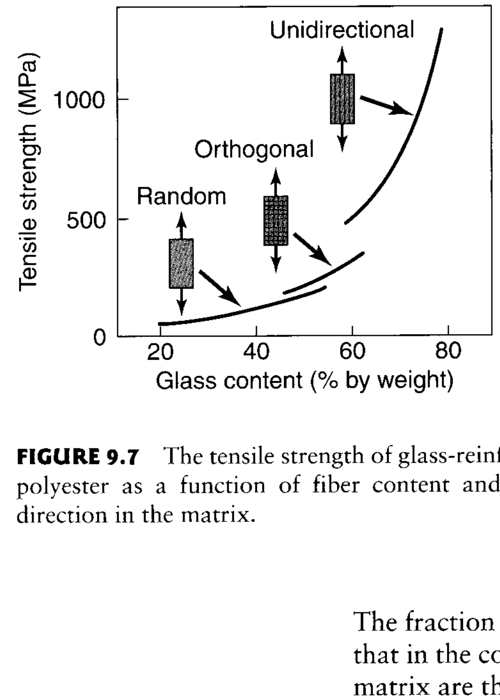


FIGURE 9.7 The tensile strength of glass-reinforced polyester as a function of fiber content and fiber direction in the matrix.

fiber-reinforced packaging tape is when pulled in tension, yet how easily it can split when pulling in the width direction.

Because it is an engineered material, a reinforced plastic part can be given an optimal configuration for a specific service condition. For example, if 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) 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*, Section 19.13.3.) Also, a composite flywheel rotor has been produced using a special weaving technique in which the reinforcing fibers (E-glass) are aligned in the radial direction as well as in the hoop direction. Designed for mechanical-energy storage systems in low-emission electric and hybrid vehicles, the flywheel can operate at speeds up to 50,000 rpm.

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 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 Equation (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 now can 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 Section 2.2 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,

$$E_c = xE_f + (1 - x)E_m. \qquad (9.5)$$

EXAMPLE 9.1 Calculation of Stiffness of a Composite and Load Supported by Fibers

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. 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

$$P_c = P_f + P_m \quad \text{and} \quad P_m = \frac{P_f}{0.75},$$

we obtain,

$$P_c = P_f + \frac{P_f}{0.75} = 2.33P_f, \quad \text{or} \quad P_f = 0.43P_c.$$

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 with asbestos fibers. In the 1920s, *formica*, a trade name, was developed and used commonly for countertops. Epoxies first were 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, resulting in materials that are now called **advanced composites**. Glass or carbon fiber-reinforced hybrid plastics are available for high-temperature applications, with continuous use ranging up to about 300°C.

Reinforced plastics typically are 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. The Boeing 777 is made of about 9% composites by total weight; that proportion is triple the composite content of previous Boeing transport aircraft. The fuselage of the new Boeing 787 Dreamliner is made of composites, for reduced weight and higher fatigue resistance.

By virtue of the resulting weight savings, reinforced plastics have reduced fuel consumption by about 2%. The newly designed Airbus jumbo jet A380, with a capacity of 550 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 upper fuselage will be made of alternating layers of aluminum and glass fiber-reinforced epoxy prepregs.

The hood of the 2008 Mustang automobile is made of a carbon–fiber composite. The contoured frame of the Stealth bomber is fashioned out of composites

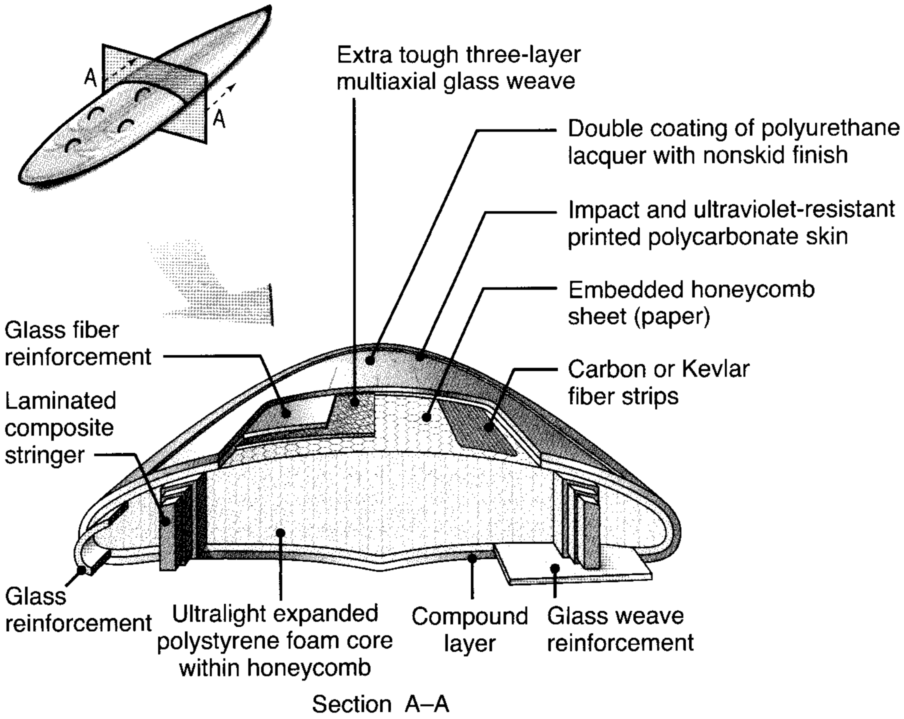


FIGURE 9.8 Cross section of a composite sailboard, an example of advanced materials construction. *Source*: K. Easterling, *Tomorrow's Materials*, 2nd ed., Institute of Metals, 1990.

consisting of carbon and glass fibers, epoxy-resin matrices, high-temperature polyimides, and other advanced materials. The structure of the Lear Fan 2100 passenger aircraft is almost totally made of graphite–epoxy reinforced plastic. 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 50 knots (93 km/h). More recent developments include (a) reinforcing bars (*rebar*) for concrete, replacing steel bars and thus lowering the costs involved due to their corrosion, and (b) rollers for papermaking and similar industries, with lower deflections as compared to traditional steel rollers.

EXAMPLE 9.2 Composite Military Helmets and Body Armor

Personal protective equipment in the form of body armor and composite helmets have become widespread for military and police applications. Body armor depends on high-strength woven fibers to prevent the penetration of projectiles. To stop a bullet, a composite material first must deform or flatten it; this process 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, of course, by the user of the armor, but successful designs will contain bullets and shrapnel and prevent serious and fatal injuries.

There are two main types of body armor: *soft armor*, which relies upon many layers of woven, high-strength fibers and is designed mainly to defeat handguns; and *hard armor*, which uses a metal, ceramic, or polymer plate in addition to the woven fiber and is intended to provide protection against rifle rounds and shrapnel. A schematic of body armor is shown in Fig. 9.9.

A number of fiber meshes have been used in body armor applications. Different suppliers employ different combinations of fiber meshes and may include additional layers to provide protection against blunt trauma. Historically, the first fiber used for

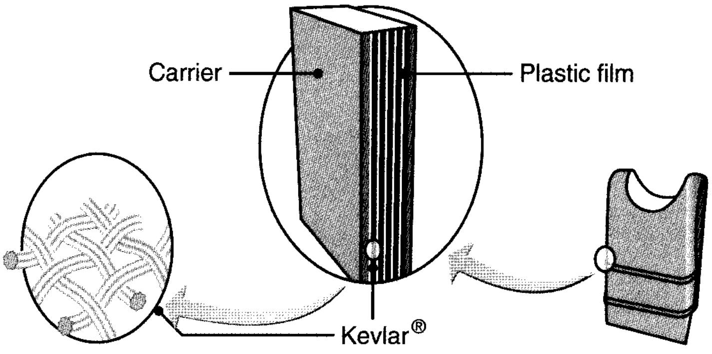


FIGURE 9.9 Schematic illustration of body armor, showing the layers of woven fibers.

flexible body armor was Kevlar 29 (an aramid), which has been improved through a number of versions. Other forms include Kevlar 49, Kevlar 129, and Kevlar Protera—a form whose 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. Honeywell also produces an aramid fiber-based body armor, but other designs, such as Akzo Noble's TWARON aramid fiber, use over a thousand finely spun filaments that interact with each other to dissipate the impact energy.

Spectra fiber is used to make the Spectra Shield composite for use in 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 resin layers are sealed between two thin sheets of polyethylene film, which is similar in appearance to plastic food wrap.

Hard armor uses a number of designs, but typically consists of steel, ceramic (usually aluminum oxide and silica), or polyethylene plates strategically located to prevent penetration of ballistic particles to critical areas. Designs currently being evaluated utilize fluids with suspended nanoparticles of silica. At low strain rates, these fluids are inviscid and flow readily. At high strain rates typical of ballistic particles, these fluids are very resistant to deformation and can provide additional protection. The fluid is contained by the woven fiber mesh (it acts like a sponge holding the fluid in place) and is contained by the outer fabric.

In addition, a composite military helmet has been developed that, although weighing about the same as a conventional manganese–steel helmet, covers more of the head and offers twice the ballistic and fragmentation protection. A composite helmet has a nonwoven fiber construction made with Spectra fibers in a thermosetting polymer matrix, which effectively stops the bullet by flattening it as it strikes the first layer of material.

Source: Courtesy of Pinnacle Armor, AlliedSignal Corp., and CGS Gallet SA.

9.5 Metal-matrix Composites

The advantages 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 parts. Matrix materials in *metal-matrix composites* (MMC) are usually aluminum, aluminum–lithium alloy (lighter than aluminum), magnesium, copper, titanium, and superalloys (Fig. 9.10). Fiber materials are graphite, aluminum oxide, silicon carbide, boron, molybdenum, and tungsten. The elastic modulus of nonmetallic fibers ranges between 200 and 400 GPa, with tensile strengths in the range from 2000 to 3000 MPa.

Typical compositions and applications for metal-matrix composites are given in Table 9.3.

FIGURE 9.10 Examples of metal-matrix composite parts. *Source*: Courtesy of Metal Matrix Cast Composites, LLC.

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

EXAMPLE 9.3 Aluminum-matrix Composite Brake Calipers

One of the trends in automobile design and manufacture is the increased push towards 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 from 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 that would require a larger volume, and the space available between the wheel and rotor is highly 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 material 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 that matched deflections of cast-iron calipers in a

TABLE 9.4

Summary of Fiber and Material Properties for an Automotive Brake Caliper

Property	Alumina fiber	Aluminum-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

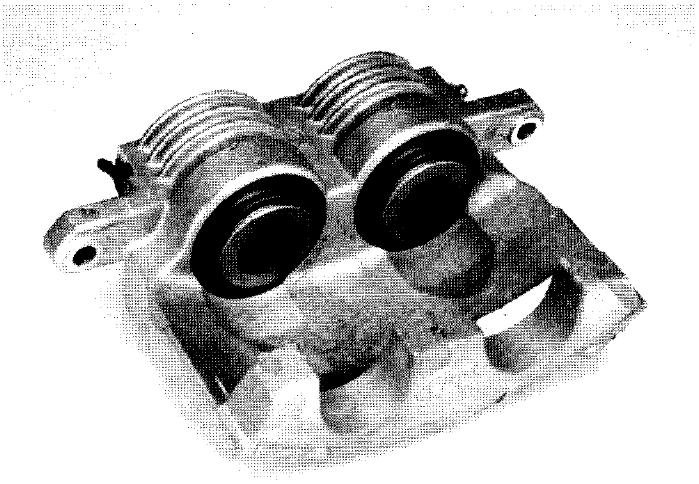


FIGURE 9.11 Aluminum-matrix composite brake caliper using nanocrystalline alumina fiber reinforcement. *Source*: Courtesy of 3M Speciality Materials Division.

packaging-constrained environment. The new brake caliper is shown in Fig. 9.11. It produces a weight savings of 50% and brings added benefits of easy recyclability and corrosion resistance.

9.6 Ceramic-matrix Composites

Ceramic-matrix composites (CMC) are important because of their resistance to high temperatures and corrosive environments. As described in Section 8.3, ceramics are strong and stiff, they resist high temperatures, but they 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 on base metals or substrates (Chapter 34). Examples are

- Plating of aluminum or other metals over plastics, generally for decorative purposes,
- Enamels, and
- Vitreous (glasslike) coatings on metal surfaces for various functional or ornamental purposes.

Composites also are made into cutting tools and dies, such as cement 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 of granite particles in an epoxy matrix has high strength, good vibration-damping capacity (better than gray cast iron), and good frictional characteristics. It is used in machine-tool beds for some precision grinders.

EXAMPLE 9.4 Composites in the Aircraft Industry

High fuel prices have significantly affected the operations of the aerospace industry, and any design advantages that lead to increased efficiency or reduced fuel consumption are being 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 a number of 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 sandwiched between thin sheets of

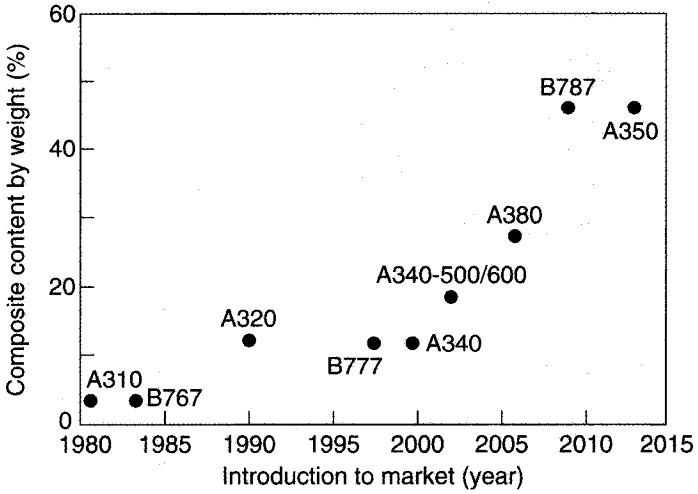


FIGURE 9.12 Composite content in selected commercial aircraft as a function of time (measured by date first introduced into the market). A = Airbus, B = Boeing.

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. In addition to weight savings, GLARE 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 barrel 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 usually are 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 typically are graphite, boron, aluminum oxide, silicon carbide, molybdenum, or tungsten. Matrix materials generally consist of aluminum, aluminum–lithium alloy, magnesium, copper, titanium, and superalloys.
- For 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 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.

ASM Handbook, Vol. 21: **Composites**, ASM International, 2001.

Bansal, N.P. (ed.), **Handbook of Ceramic Composites**, Springer, 2004.

Barbero, E.J., **Introduction to Composite Materials Design**, CRC Press, 1998.

Cantor, B., Dunne, F.P.E., and Stone, I.C. (eds.), **Metal and Ceramic Matrix Composites**, Taylor & Francis, 2003.

Cardelli, F., **Materials Handbook: A Concise Desk Reference**, 2nd ed., Springer, 2008.

Chawla, K.K., **Composite Materials: Science and Engineering**, 3rd ed., Springer, 2008.

Cheremisinoff, N.P., and Cheremisinoff, P.N., **Fiberglass Reinforced Plastics**, Noyes Publications, 1995.

Daniel, I.M., and Ishai, O., **Engineering Mechanics of Composite Materials**, 2nd ed., Oxford, 2005.

Fitzer, E., and Manocha, L.M., **Carbon Reinforcements and Carbon/Carbon Composites**, Springer, 1998.

Gay, D., and Hoa, S.V., **Composite Materials: Design and Applications**, 2nd ed., CRC Press, 2007.

Harper, C.A, **Handbook of Plastics, Elastomers, and Composites**, 3rd ed., McGraw-Hill, 1996.

Hull, D., and Clyne, T.W., **An Introduction to Composite Materials**, Cambridge Univ. Press, 1996.

Kelley, A., Cahn, R.W., and Bever, M.B. (eds.), **Concise Encyclopedia of Composite Materials**, rev. ed., Pergamon, 1994.

Krenker, W. (ed.), **Ceramic Matrix Composites: Fiber Reinforced Materials and Their Applications**, Wiley, 2008.

Mallick, P.K., **Composites Engineering Handbook**, Marcel Dekker, 1997.

Miller, E., **Introduction to Plastics and Composites: Mechanical Properties and Engineering Applications**, Marcel Dekker, 1995.

Peters, S. (ed.), **Handbook of Composites**, Chapman & Hall, 1997.

Rosato, D.V., DiMattia, D.P., and Rosato, D.V., **Designing with Plastics and Composites: A Handbook**, Van Nostrand Reinhold, 1991.

Schwartz, M., **Composite Materials**, Vol. 1: *Properties, Nondestructive Testing, and Repair*; Vol. 2: *Processing, Fabrication, and Applications*. Prentice Hall, 1997.

Strong, A.B., **Fundamentals of Composites Manufacturing: Materials, Methods and Applications**, 2nd ed., Society of Manufacturing Engineers, 2007.

REVIEW QUESTIONS

9.1. Distinguish between composites and metal alloys.

9.2. Describe the functions of the matrix and the reinforcing fibers. What fundamental differences are there in the characteristics of the two materials?

9.3. Name the reinforcing fibers generally used to make composites. Which type of fiber is the strongest? Which type is the weakest?

9.4. What is the range in length and diameter of typical reinforcing fibers?

9.5. List the important factors that determine the properties of reinforced plastics.

9.6. Comment on the advantages and limitations of metal-matrix composites, reinforced plastics, and ceramic-matrix composites.

9.7. What are the most commonly used matrix materials? Why?

9.8. Describe the advantages of hybrid composites over other composites.

9.9. What material properties are improved by the addition of reinforcing fibers?

9.10. Describe the purpose of the matrix material.

9.11. What are the most common types of glass fibers?

9.12. Explain the difference between a carbon fiber and a graphite fiber.

9.13. How can a graphite fiber be made electrically and thermally conductive?

9.14. What is a whisker? What is the difference between a whisker and a fiber?

9.15. Explain the composition of boron fibers. Why are they heavy?

QUALITATIVE PROBLEMS

9.16. How do you think the use of straw mixed with clay originally came about in making brick for dwellings?

9.17. What products have you personally seen that are made of reinforced plastics? How can you tell?

9.18. Describe applications that are not well suited for composite materials.

9.19. Is there a difference between a composite material and a coated material? Explain.

9.20. Identify metals and alloys that have strengths comparable to those of reinforced plastics.

9.21. What limitations or disadvantages do composite materials have? What suggestions would you make to overcome the limitations?

9.22. Give examples of composite materials other than those stated in this chapter.

9.23. Explain why the behavior of the materials depicted in Fig. 9.5 is as shown.

9.24. Explain why fibers are so capable of supporting a major portion of the load in composite materials.

9.25. Do metal-matrix composites have any advantages over reinforced plastics? Explain.

9.26. Give reasons for the development of ceramic-matrix composites. Name some applications and explain why they should be effective.

9.27. 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.

9.28. How would you go about trying to determine the strength of a fiber?

9.29. Glass fibers are said to be much stronger than bulk glass. Why is this so?

9.30. Describe situations in which a glass could be used as a matrix material.

9.31. 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.

9.32. 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.

QUANTITATIVE PROBLEMS

▶**9.33.** 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.

9.34. 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 200 MPa? What would be the answer if the fiber stiffness were doubled and the matrix stiffness were halved?

▶**9.35.** Calculate the percent increase in the mechanical properties of reinforced nylon from the data shown in Fig. 9.5.

▶**9.36.** Plot E/ρ and $E/\rho^{0.5}$ for the composite materials listed in Table 9.1, and compare your results with the properties of the materials described in Chapters 4 through 8. (See also Table 9.2.)

▶**9.37.** Calculate the stress in the fibers and in the matrix in Example 9.1. Assume that the cross-sectional area is 150 mm^2 and $P_c = 220$ kg.

▶**9.38.** Repeat the calculations in Example 9.1 (a) if Nextel 610 fiber is used and (b) if Spectra 2000 is used.

9.39. 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?

▶**9.40.** Plot the elastic modulus and strength of an aluminum metal-matrix composite with high-modulus carbon fibers as a function of fiber content.

9.41. 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 200 MPa and a matrix strength of 30 MPa.

SYNTHESIS, DESIGN, AND PROJECTS

9.42. 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?

9.43. Using the information given in this chapter, develop special designs and shapes for possible new applications of composite materials.

9.44. Would a composite material with a strong and stiff matrix and a soft and flexible reinforcement have any practical uses? Explain.

9.45. Make a list of products for which the use of composite materials could be advantageous because of their anisotropic properties.

9.46. Inspect Fig. 9.1 and explain what other components of an aircraft, including the cabin, could be made of composites.

9.47. Name applications in which both specific strength and specific stiffness are important.

9.48. What applications for composite materials can you think of in which high thermal conductivity would be desirable? Explain.

9.49. 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.

9.50. 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.

9.51. 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 diection.)

9.52. 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.

9.53. Comment on your observations on the design of the sailboard illustrated in Fig. 9.8.

9.54. 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.

9.55. 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.

9.56. 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.

9.57. Derive a general expression for the coefficient of thermal expansion for a continuous fiber-reinforced composite in the fiber direction.

9.58. It is possible to make fibers or whiskers with a varying cross section, or a "wavy" fiber. What advantages would such fibers have?

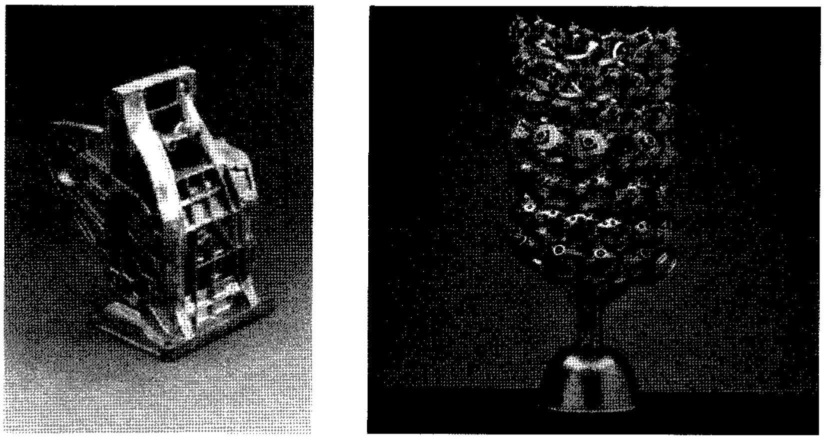
9.59. Describe how you can produce some simple composite materials using raw materials available around your home. Explain.

9.60. *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.

PART II

Metal-Casting Processes and Equipment

As described throughout the rest of this book, several different methods are available to shape metals into useful 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. Examples of cast parts are shown in Figure II.1. Casting first was used around 4000 B.C. to make ornaments, arrowheads, and various other objects. A wide variety of products can be cast, and the process is capable of producing intricate shapes in one piece, 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.



(a) (b)

FIGURE II.1 Examples of cast parts. (a) A die-cast aluminum transmission housing. (b) A tree of rings produced through investment casting. *Source*: (a) Courtesy of North American Die Casting Association, (b) Courtesy of Romanoff, Inc.

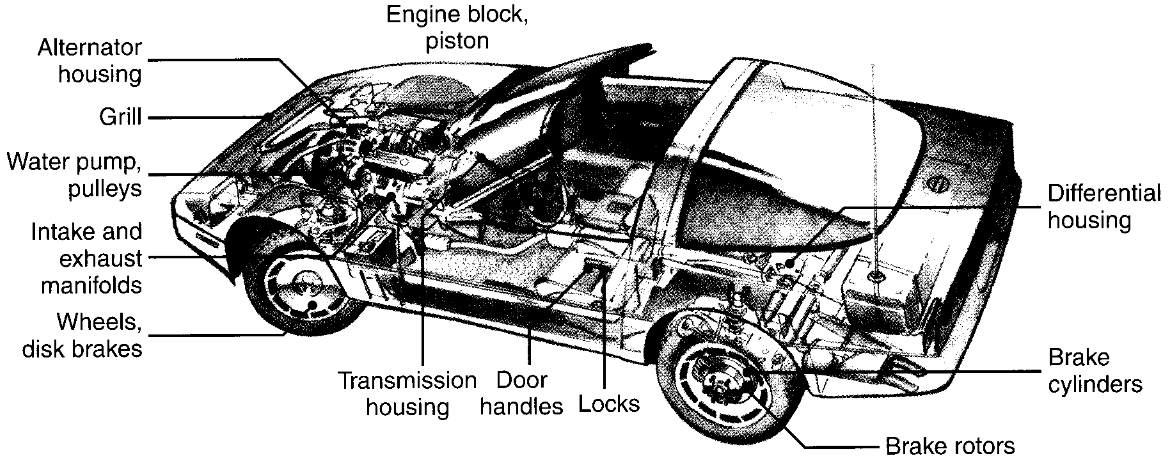


FIGURE II.2 Cast parts in a typical automobile.

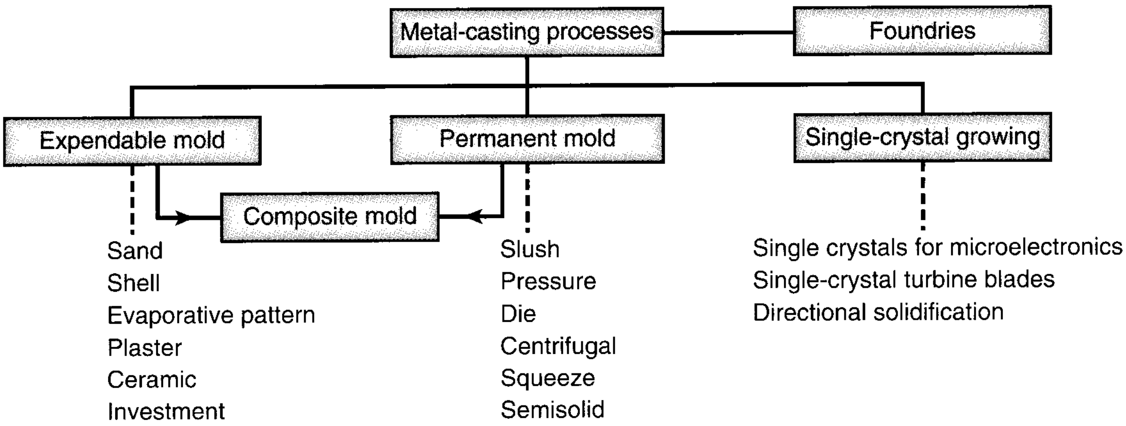


FIGURE II.3 Outline of metal-casting processes described in Part II.

As in all manufacturing operations, each casting process has its own characteristics, applications, advantages, limitations, and costs. Casting processes are most often selected over other manufacturing methods for the following reasons:

- Casting can produce complex shapes and can incorporate internal cavities or hollow sections.
- Very large parts can be produced in one piece.
- Casting can utilize materials that are difficult or uneconomical to process by other means.
- The casting process can be economically competitive with other manufacturing processes.

Almost all metals can be cast in (or nearly in) the final shape desired, 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 the control of chemical composition, mechanical properties of castings can equal those made by other manufacturing processes.

Fundamentals of Metal Casting

CHAPTER 10

- First used about six thousand years ago, casting continues to be an important manufacturing process for producing very small, as well as very large and complex, parts.
- The first topic discussed is solidification of molten metals, including the differences between solidification of pure metals and alloys, as well as concepts such as grain nucleation and freezing range.
- 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 through a mold.
- The concept and importance of turbulence versus laminar flow is introduced.
- Heat transfer and shrinkage of castings are also discussed, 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 cast-metal properties.

EXAMPLES:

10.1 Introduction

The **casting** process basically involves (a) pouring molten metal into a mold patterned after the part to be manufactured, (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 the production of good-quality, economical castings and for establishing proper techniques for mold design and casting practice.

Important considerations in casting operations are as follows:

- Flow of the molten metal into the mold cavity
- Solidification and cooling of the metal in the mold
- Influence of the type of mold material.

This chapter describes relationships among the many factors involved in casting. The flow of molten metal into the mold cavity first is discussed 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. The type of mold also has an important influence, because it affects the rate of cooling. The chapter finishes with a description of the factors influencing defect formation.

Industrial metal-casting processes, design considerations, and casting materials are described in Chapters 11 and 12. The casting of ceramics and of plastics, which involves 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 series of events takes place during the solidification of the metal and its cooling 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 metal. The significant factors affecting these events are the type of metal, 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 a 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 and Fig. 4.4.) After the temperature of the molten metal drops to its freezing point, its temperature remains constant while the *latent heat of fusion* is given off. The *solidification front* (solid–liquid interface) moves through the molten metal from the mold walls in toward the center. The solidified metal, called the casting, is taken out of 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 shrink while cooling and generally also shrink when they solidify (see Table 10.1).

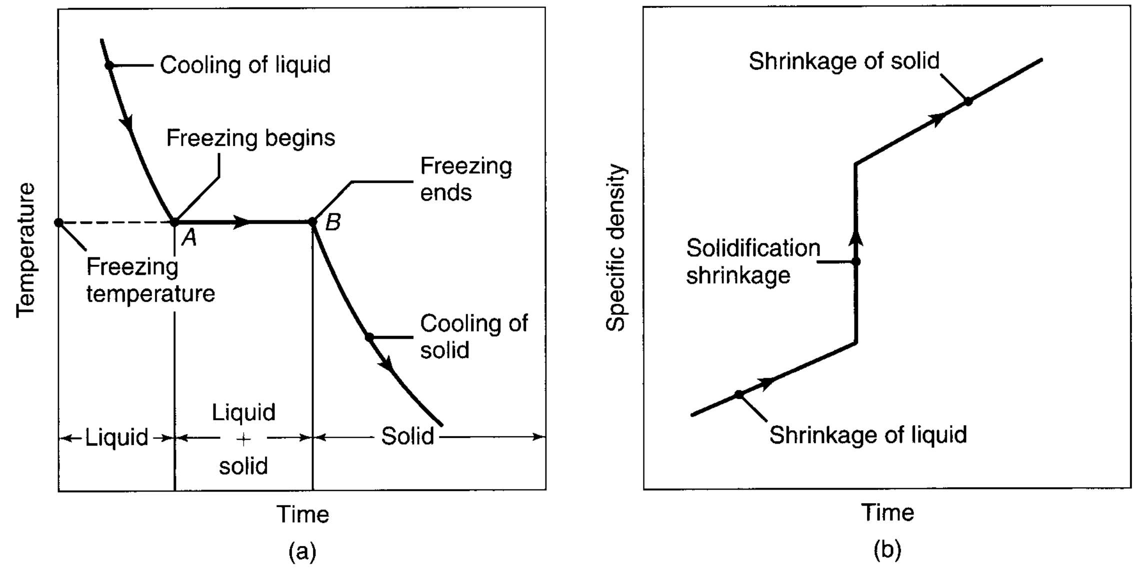


FIGURE 10.1 (a) Temperature as a function of time for the solidification of pure metals. Note that freezing takes place at a constant temperature. (b) Density as a function of time.

This is an important consideration, because shrinkage can lead to microcracking and associated porosity, which can in turn compromise 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 or typically are much cooler than the molten metal, the metal cools rapidly and produces 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 growth. As the driving force of the heat transfer is reduced away from the mold walls, the grains become equiaxed and coarse. This process of grain development is known as **homogenous nucleation**, meaning that the grains (crystals) grow upon themselves, starting at the mold wall.

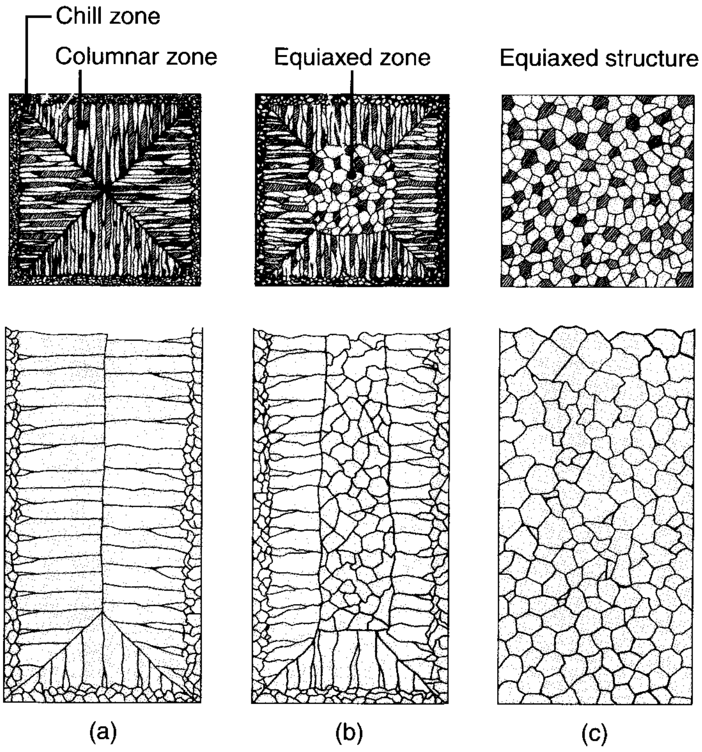


FIGURE 10.2 Schematic illustration of three cast structures of metals solidified in a square mold: (a) pure metals; (b) solid–solution alloys; and (c) structure obtained by using nucleating agents.
Source: After G.W. Form, J.F. Wallace, J.L. Walker, and A. Cibula.

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 the presence of liquid metal between the dendrite arms. 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 contribute to detrimental factors, such as 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**:

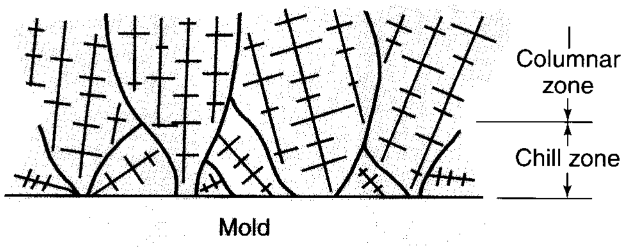


FIGURE 10.3 Development of a preferred texture at a cool mold wall. Note that only favorably oriented grains grow away from the surface of the mold.

$$\text{Freezing range} = T_L - T_S. \qquad (10.1)$$

It can be seen in Fig. 10.4 that pure metals have a freezing range that approaches zero and that the solidification front moves as a plane without forming a mushy zone. *Eutectics* (Section 4.3) solidify in a similar manner, with an essentially plane front. The type of structure developed after solidification depends on the composition of the eutectic. In alloys with a nearly symmetrical phase diagram, the structure is generally lamellar, with two or more solid phases present, depending on

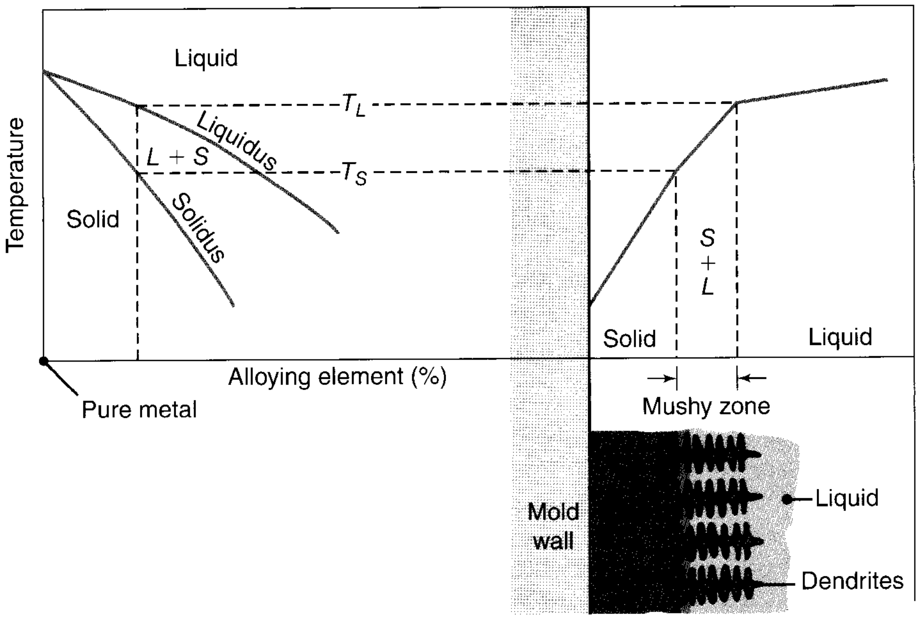


FIGURE 10.4 Schematic illustration of alloy solidification and temperature distribution in the solidifying metal. Note the formation of dendrites in the mushy zone.

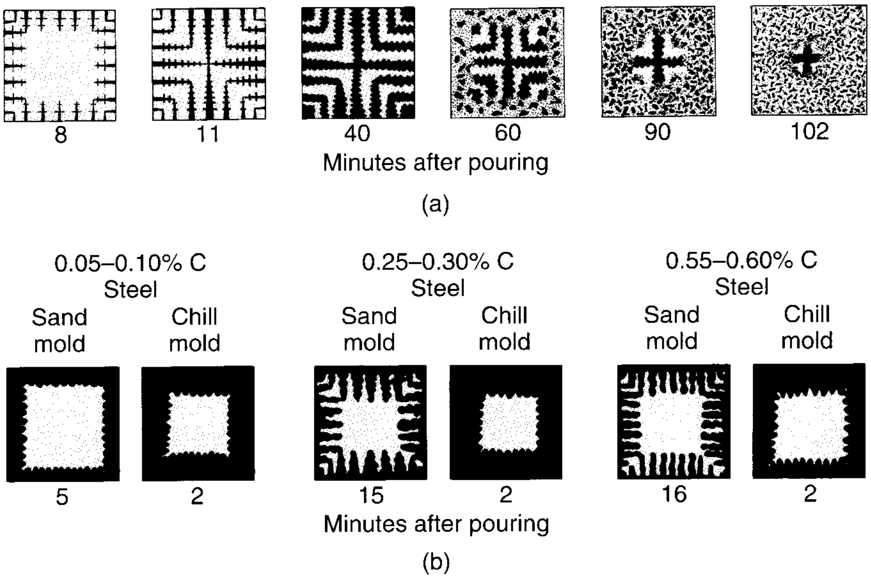


FIGURE 10.5 (a) Solidification patterns for gray cast iron in a 180-mm square casting. Note that after 11 minutes of cooling, dendrites reach each other, but the casting is still mushy throughout. It takes about 2 hours for this casting to solidify completely. (b) Solidification of carbon steels in sand and chill (metal) molds. Note the difference in solidification patterns as the carbon content increases. *Source*: After H.F. Bishop and W.S. Pellini.

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 important for cast irons.

For alloys, a *short freezing range* generally involves a temperature difference of less than 50°C, and a *long freezing range* greater than 110°C. Ferrous castings generally have narrow mushy zones, whereas aluminum and magnesium alloys have wide mushy zones. Consequently, these alloys are in a mushy state throughout most of the solidification process.

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 K/s) the structures developed are *amorphous*, as described in Section 6.14.

The structures developed and the resulting grain size influence the properties of the casting. As grain size decreases, the strength and the 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. 10.12) during solidification decreases. Lack of uniformity in grain size and grain distribution results in castings with *anisotropic properties*.

A criterion describing the kinetics of the liquid–solid interface is the ratio G/R, where G is the *thermal gradient* and R is the *rate* at which the interface moves. Typical values for G range from 10^2 to 10^3 K/m and for R range from 10^{-3} to 10^{-4} m/s. Dendritic-type structures (Fig. 10.6a and b) typically have a G/R ratio in the range from 10^5 to 10^7, whereas ratios of 10^{10} to 10^{12} produce a plane-front, nondendritic liquid–solid interface (Fig. 10.7).

10.2.3 Structure–property Relationships

Because all castings are expected to possess certain properties to meet design and service requirements, the relationships between properties and the structures developed during solidification are important aspects of casting. This section describes these relationships in terms of dendrite morphology and the concentration of alloying elements in various regions within the metal.

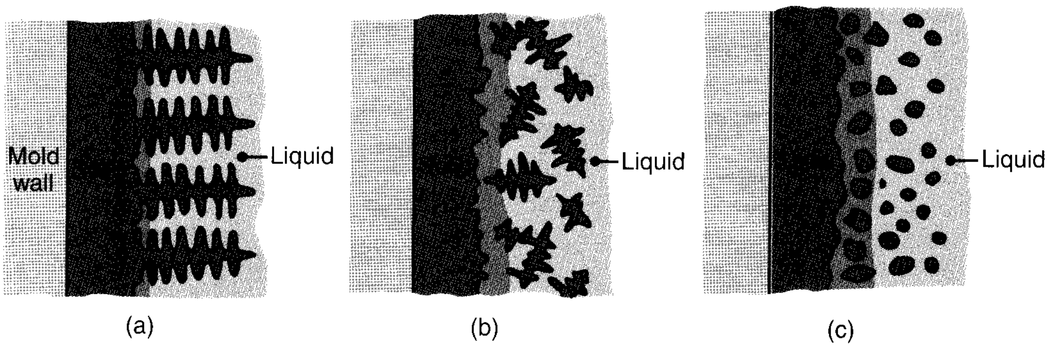


FIGURE 10.6 Schematic illustration of three basic types of cast structures: (a) columnar dendritic; (b) equiaxed dendritic; and (c) equiaxed nondendritic. *Source*: Courtesy of D. Apelian.

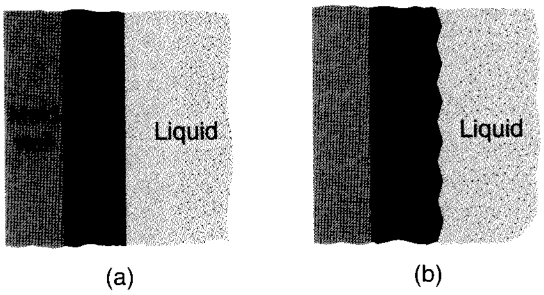


FIGURE 10.7 Schematic illustration of cast structures in (a) plane front, single phase, and (b) plane front, two phase. *Source*: Courtesy of D. Apelian.

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. Cored dendrites have a surface composition different from that at their centers, a difference referred to as a *concentration gradient*. The surface of the dendrite has a higher concentration of alloying elements than does 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 (Fig. 10.2b), the opposite occurs; that is, the center of the casting has a lower concentration of alloying elements (**inverse segregation**) than does 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. **Gravity segregation** describes the process whereby higher density inclusions or compounds sink and lighter elements (such as antimony in an antimony–lead alloy) float to the surface.

A typical cast structure of a solid–solution alloy with an inner zone of equiaxed grains is shown in Fig. 10.2b. This 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 (**heterogeneous nucleation**).

Because of the presence of *thermal gradients* in a solidifying mass of liquid metal, and due to gravity and the resultant density differences, *convection* has a strong influence on the structures developed. 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 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. Another form of semisolid metal forming is *thixotropic casting*, where a solid billet is heated to the semisolid state and then injected into a die-casting mold (Section 11.4.5). The heating is usually by convection in a furnace, but can be enhanced by the use of mechanical or electromagnetic methods.

10.3 Fluid Flow

To emphasize the importance of fluid flow in casting, let's briefly describe a basic gravity-casting system, as shown in Fig. 10.8. 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 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 to supply any 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 problems such 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.

Two basic principles of fluid flow are relevant to gating design: Bernoulli's theorem and the law of mass continuity.

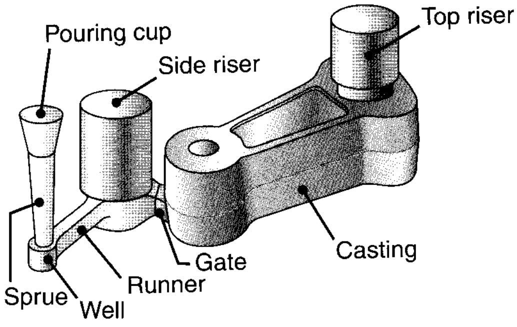


FIGURE 10.8 Schematic illustration of a typical riser-gated casting. Risers serve as reservoirs, supplying molten metal to the casting as it shrinks during solidification.

Bernoulli's Theorem. This theorem is based on the principle of the conservation of energy and relates pressure, velocity, the elevation of the fluid at any location in the system, and the frictional losses in a system that is full of liquid. 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, at a particular location in the system, 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 often are used to inhibit such behavior in sand molds.

Sprue Design. An application of the two principles just stated is the traditional tapered design of sprues (shown in Fig. 10.8). 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 downward. Thus, if we design a sprue with a constant cross sectional area and pour the molten metal into it, regions can develop where the liquid loses contact with the sprue walls. As a result, **aspiration** (a process whereby air is sucked in or 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. Moving downward from the top, the cross sectional area of the sprue must therefore decrease. Depending on the assumptions made, expressions other than Eq. (10.5) can also be obtained. For example, we may 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*. For example, consider the situation shown in Fig. 10.7 where molten metal is poured into a pouring 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 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 can include energy losses due to turbulence, as well as 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 this 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 stay 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 material (referred to as *mold wash*) or to result in too high of a Reynolds number (see the following). Otherwise, turbulence and associated air entrainment results. Many computational tools are now available to evaluate gating designs and assist in the sizing of components such as Magmasoft, ProCast, Quikcast, and Powercast.

Flow Characteristics. An important consideration of the fluid flow in gating systems is the presence of **turbulence**, as opposed to the *laminar flow* of fluids. Turbulence is flow that is highly chaotic; in casting systems such flow can lead to aspiration. 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

$$\mathrm{Re} = \frac{vD\rho}{\eta}, \tag{10.6}$$

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 to occur.

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 generally is 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 geometry of channel cross sections in gating system design.

Dross or slag can be eliminated only by *vacuum casting* (Section 11.4.2). Conventional atmospheric casting mitigates dross or slag by (a) skimming, (b) using properly designed pouring basins and runner systems, or (c) using filters, which also can eliminate turbulent flow in the runner system. Filters usually are 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*, which consists of two basic factors: (1) characteristics of the molten metal and (2) casting parameters. The following characteristics of molten metal influence fluidity:

Viscosity. As viscosity and its sensitivity to temperature (*viscosity index*) increase, fluidity decreases.

Surface Tension. A high surface tension of the liquid metal reduces fluidity. Because of this, oxide films on the surface of the molten metal have a significant 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 adverse effect on fluidity. This effect can be verified by observing the viscosity of a liquid (such as oil) with and without sand particles in it; the liquid with sand in it has a higher viscosity and, hence, lower fluidity.

Solidification Pattern of the Alloy. The manner in which solidification takes place (Section 10.2) can influence fluidity. Moreover, fluidity is inversely proportional to the freezing range: 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 and also influence the 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 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.

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 specified more easily.

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).

Although complex, the term **castability** generally is used to describe the ease with which a metal can be cast to produce a part with good quality. Castability includes not only fluidity, but the nature of casting practices as well.

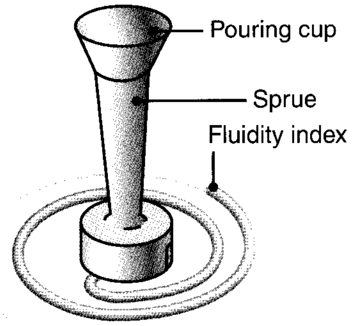


FIGURE 10.9 A test method for fluidity using a spiral mold. The *fluidity index* is the length of the solidified metal in the spiral passage. The greater the length of the solidified metal, the greater is the metal's fluidity.

10.4.1 Tests for Fluidity

Several tests have been developed to quantify fluidity, although none is accepted universally. In one such common test, the molten metal is made to flow along a channel that is at room temperature (Fig. 10.9); the distance the metal flows before it solidifies and stops flowing is a measure of its fluidity. Obviously, this length is a function of the thermal properties of the metal and the mold, as well as of the design of the channel. Still, such fluidity tests are useful and simulate casting situations to a reasonable degree.

10.5 Heat Transfer

The heat transfer during the complete cycle (from pouring, to solidification, and to cooling to room temperature) is another important consideration in metal casting. Heat flow at different locations in the system is a complex phenomenon and depends on several factors relating to the material cast and the mold and process parameters. For instance, in casting thin sections, the metal flow rates must be high enough to avoid premature chilling and solidification. On the other hand, the flow rate must not be so high as to cause excessive turbulence—with its detrimental effects on the casting process.

A typical temperature distribution at the mold liquid–metal interface is shown in Fig. 10.10. 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.

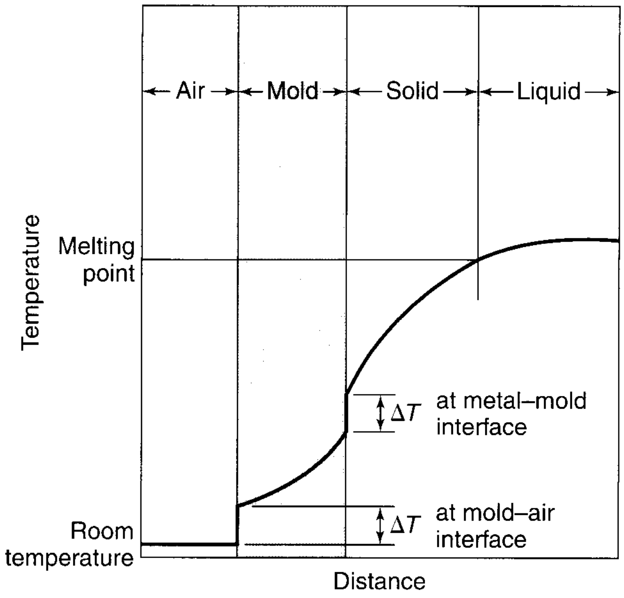


FIGURE 10.10 Temperature distribution at the interface of the mold wall and the liquid metal during the solidification of metals in casting.

10.5.1 Solidification Time

During the early stages of solidification, a thin skin begins to form at the relatively cool mold walls, and as time passes, the thickness of the skin increases (Fig. 10.11). With flat mold walls, this 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*):

$$\text{Solidification time} = C\left(\frac{\text{Volume}}{\text{Surface area}}\right)^n, \qquad (10.7)$$

where C is a constant that reflects (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 is taken as 2. Thus, a large solid sphere will solidify and cool to ambient temperature at a much slower rate than will a smaller solid sphere. The reason for this is that the volume of a sphere is proportional to the cube of its diameter, and the surface area 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. 10.11. As illustrated, the unsolidified molten metal has been poured from the mold at different time intervals ranging from 5 seconds to 6 minutes. Note that (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.

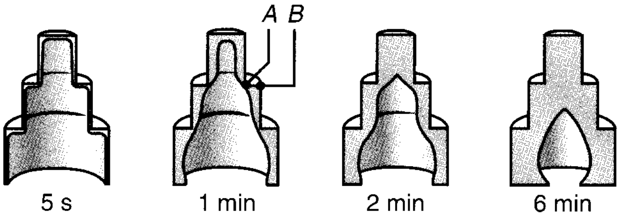


FIGURE 10.11 Solidified skin on a steel casting. The remaining molten metal is poured out at the times indicated in the figure. Hollow ornamental and decorative objects are made by a process called slush casting, which is based on this principle. *Source*: After H.F. Taylor, J. Wulff, and M.C. Flemings.

EXAMPLE 10.1 Solidification Times for Various Shapes

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. Which piece will solidify the fastest, and which one the slowest? Assume that $n = 2$.

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:

$$V = \left(\frac{4}{3}\right)\pi r^3,\ r = \left(\frac{3}{4\pi}\right)^{1/3}.$$

$$A = 4\pi r^2 = 4\pi\left(\frac{3}{4\pi}\right)^{2/3} = 4.84.$$

Cube:

$$V = a^3,\ a = 1, \text{ and } A = 6a^2 = 6.$$

Cylinder:

$$V = \pi r^2 h = 2\pi r^3,\ r = \left(\frac{1}{2\pi}\right)^{1/3},$$

$$A = 2\pi r^2 + 2\pi r h = 6\pi r^2 = 6\pi\left(\frac{1}{2\pi}\right)^{2/3} = 5.54.$$

The respective solidification times are therefore

$$t_{\text{sphere}} = 0.043C,\ t_{\text{cube}} = 0.028C,\ t_{\text{cylinder}} = 0.033C.$$

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 (latent heat of fusion).
3. Contraction of the solidified metal (the casting) as its temperature drops to ambient temperature.

The largest potential amount of shrinkage occurs during the cooling of the casting to ambient temperature. The amount of contraction during the solidification of various metals is shown in Table 10.1. Note that some metals (such as gray cast

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		

iron) expand. (The reason 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 is further discussed in Section 12.2.1 in connection with design considerations in casting.

10.6 Defects

As will be seen in this section (as well as in other sections throughout Parts II through VI), various defects can develop during manufacturing that depend on factors such as materials, part design, and processing techniques. While some defects affect only the appearance of the parts made, others can have major adverse effects on the structural integrity of the parts.

Several defects can develop in castings (Figs. 10.12 and 10.13). Because different names have been used in the past to describe the same defect, 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 (*intergranular*) 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

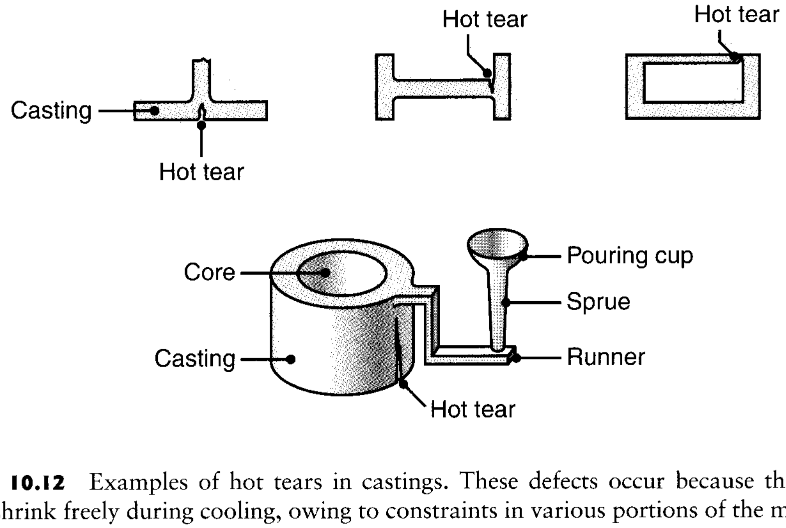


FIGURE 10.12 Examples of hot tears in castings. These defects occur because the casting cannot shrink freely during cooling, owing to constraints in various portions of the molds and cores. Exothermic (heat-producing) compounds may be used (as exothermic padding) to control cooling at critical sections to avoid hot tearing.

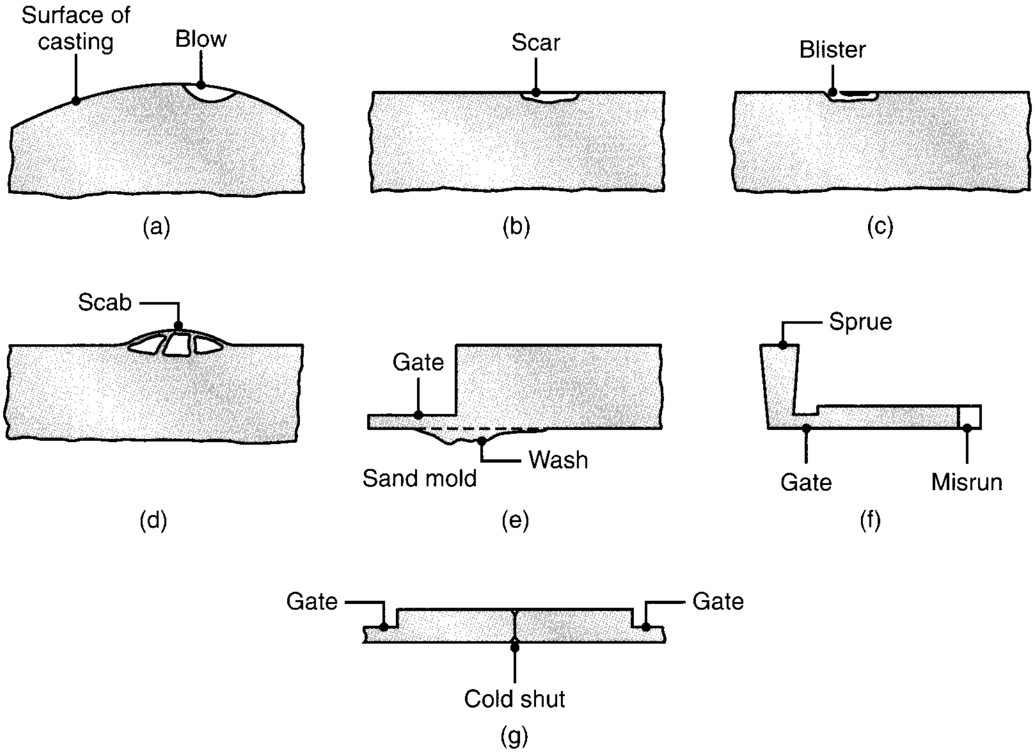


FIGURE 10.13 Examples of common defects in castings. These defects can be minimized or eliminated by proper design and preparation of molds and control of pouring procedures. *Source*: After J. Datsko.

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 factors such as improper shrinkage allowance, pattern-mounting error, irregular contraction, deformed pattern, or warped casting.

G—**Inclusions**, which form during melting, solidification, and molding; these are generally nonmetallic. They are regarded as harmful because they act as stress raisers and thus reduce the strength of the casting. Inclusions may form during melting when the molten metal reacts with the environment (usually oxygen) or with the crucible or mold material. Chemical reactions among components in the molten metal itself may produce inclusions; slags and other foreign material entrapped in the molten metal also become inclusions, although filtering can remove particles as small as 30 μm. Finally, spalling of the mold and core surfaces can produce inclusions, thus indicating the importance of the quality of molds and of their 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; as a result, 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 region begin to solidify first. *Microporosity* also can develop when the liquid metal solidifies and shrinks between dendrites and between dendrite branches.

Porosity is detrimental to the ductility of a casting and its surface finish, making the casting permeable and thus affecting the pressure tightness of a cast pressure vessel. Porosity caused by shrinkage can be reduced or eliminated by various means, including the following:

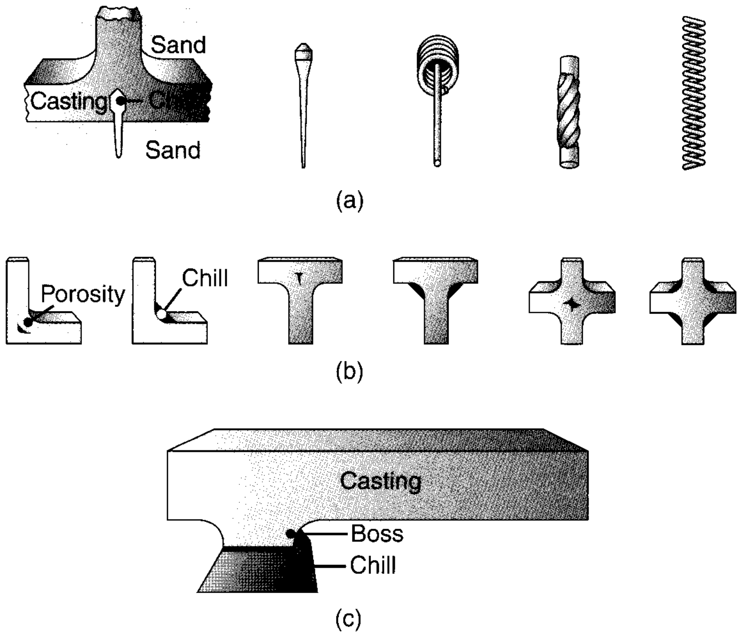


FIGURE 10.14 Various types of (a) internal and (b) external chills (dark areas at corners) used in castings to eliminate porosity caused by shrinkage. Chills are placed in regions where there is a larger volume of metal, as shown in (c).

- Adequate liquid metal should be provided to avoid cavities caused by shrinkage.
- Internal or external **chills**, as those used in sand casting (Fig. 10.14), also are an effective means of reducing shrinkage porosity. The function of chills is to increase the rate of solidification in critical regions. Internal chills usually are made of the same material as the casting and are left in the casting. However, problems may arise that involve proper fusion of the internal chills with the casting; thus, foundries generally avoid the use of internal chills for this reason. External chills may be made of the same material as the casting or may be iron, copper, or graphite.
- With alloys, porosity can be reduced or eliminated by making the temperature gradient steep. For example, mold materials that have higher thermal conductivity may be used.
- Subjecting the casting to *hot isostatic pressing* is another method of reducing porosity (see Section 17.3.2).

Because **gases** are more soluble in liquid metals than solid metals (Fig. 10.15), when a metal begins to solidify, the dissolved gases are expelled from the solution. Gases also may result from reactions of the molten metal with the mold materials. Gases either accumulate in regions of existing porosity (such as in interdendritic regions) 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*. Steel usually is deoxidized with aluminum, silicon, copper-based alloys with phosphorus, copper, titanium, or zirconium-bearing materials.

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. 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**.

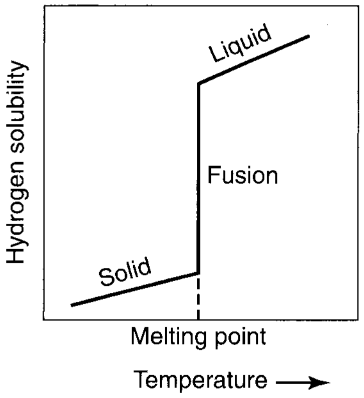


FIGURE 10.15 Solubility of hydrogen in aluminum. Note the sharp decrease in solubility as the molten metal begins to solidify.

EXAMPLE 10.2 Casting of Aluminum Automotive Pistons

Figure 10.16 shows an aluminum piston used in automotive internal combustion engines. These products must be manufactured at very high rates with very tight dimensional tolerances and strict material requirements in order to achieve proper operation. Economic concerns are obviously paramount, and it is essential that pistons be produced with a minimum of expensive finishing operations and with few rejected parts.

Aluminum pistons are manufactured through casting because of the capability to produce near-net shaped parts at the required production rates. However, with poorly designed molds, underfills or excess porosity can cause parts to be rejected, adding to the cost. These defects were traditionally controlled through the use of large machining allowances coupled with the intuitive design of molds based on experience.

The pistons are produced from high-silicon alloys, such as 413.0 aluminum alloy. This alloy has high fluidity and can create high-definition surfaces through permanent mold casting (see Section 11.4); it also has high resistance to corrosion, good weldability, and low specific gravity. The universal acceptance of aluminum pistons for internal combustion engine applications is due mainly to their light weight and high thermal conductivity. Their low inertia allows for higher engine speeds and reduced counterweighting in the crankshaft, and the higher thermal conductivity allows for more efficient heat transfer from the engine.

The H13 tool-steel mold is preheated 200° to 450°C, depending on the cast alloy and part size. Initially, preheat is achieved with a hand-held torch, but after a few castings, the mold reaches a steady-state temperature profile. The molten aluminum is heated to between 100° and 200°C above its liquidus temperature, and then a shot is placed into the infeed section of the mold. Once the molten metal shot is in place, a piston drives the metal into the mold. Because of the high thermal conductivity of the mold material, heat extraction from the molten metal is rapid, and the metal can solidify in small channels before filling the mold completely. Solidification usually starts at one end of the casting, before the mold is fully filled.

As with most alloys, it is desired to begin solidification at one extreme end of the casting and have the solidification front proceed across the volume of the casting. This results in a directionally solidified microstructure and the elimination of gross porosity that arises when two solidification fronts meet inside a casting. Regardless, casting defects such as undercuts, hot spots, porosity, cracking, and entrapped air zone defects (such as blowholes and scabs) can occur.

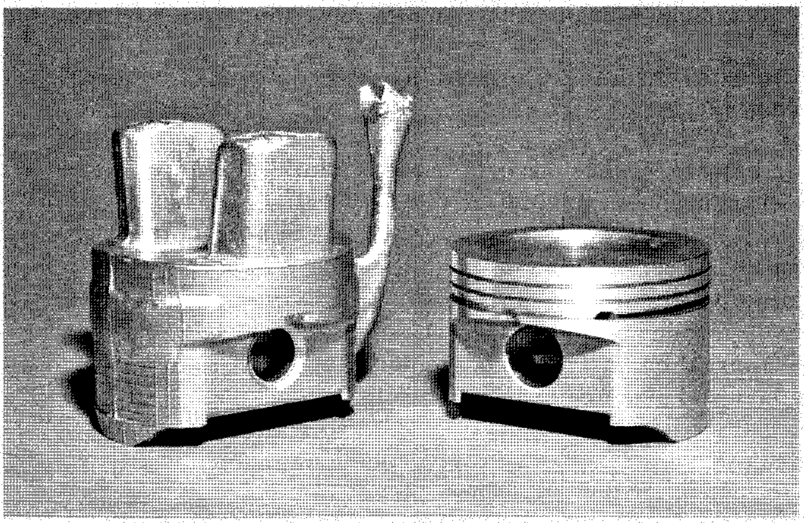


FIGURE 10.16 Aluminum piston for an internal combustion engine: (a) as cast and (b) after machining. The part on the left is as cast, including risers, sprue, and well, as well as a machining allowance; the part on the right is the piston after machining. *Source*: After S. Paolucci.

In order to improve the reliability of, and reduce the costs associated with, permanent mold casting, computer-based modeling of mold filling can suggest potential causes of defects. The computer models use the Bernoulli and continuity equations, coupled with heat transfer and solidification, to model the casting process and identify potential shortcomings. For example, Fig. 10.17 shows a result from a simulation of mold filling in which a volume of entrapped air remains in the mold. This defect is corrected by placing a vent in the area of concern, to allow air to escape during casting. Computer simulation allows designers to evaluate mold features and geometries before purchasing expensive tooling and has become an indispensible process in reducing costs and eliminating defects in casting.

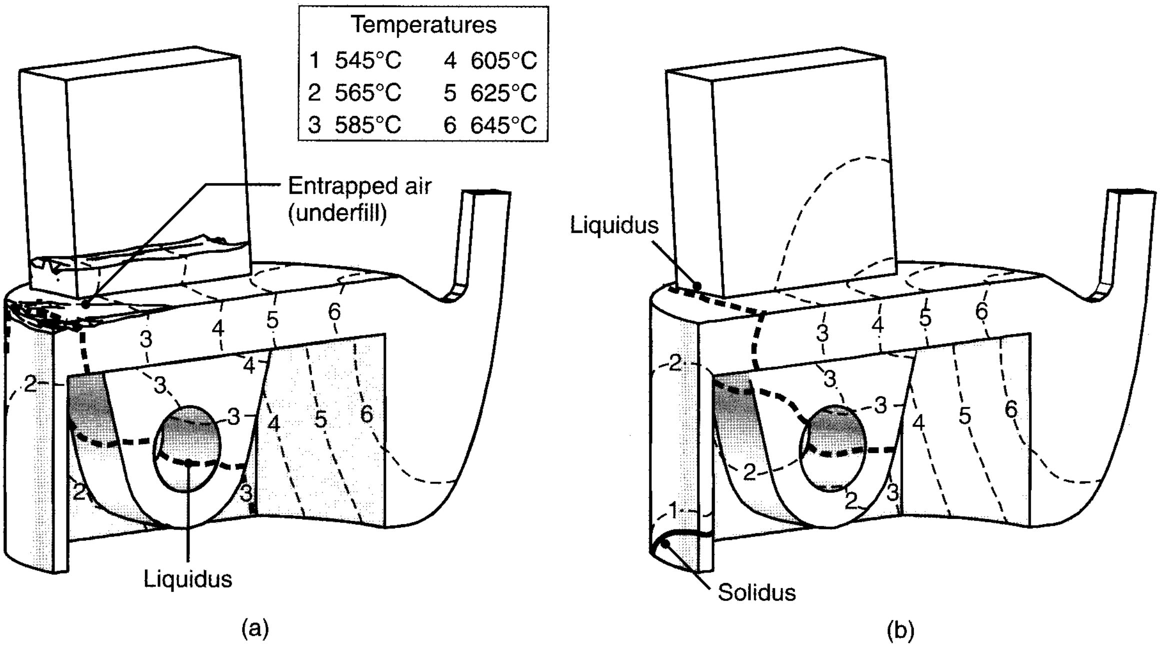


FIGURE 10.17 Simulation of mold filling and solidification. (a) 3.7 seconds after start of pour. Note that the mushy zone has been established before the mold is completely filled. (b) Using a vent in the mold for removal of entrapped air, 5 seconds after pouring. *Source*: After S. Paolucci.

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 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 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.
- Although most metals shrink during solidification, gray cast iron and some aluminum alloys actually expand. Dimensional changes and cracking (hot tearing) are difficulties that can arise during solidification and cooling. Seven 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 of the melt, removal of 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
Freezing range
Gate
Gating system
Heterogeneous nucleation
Homogeneous nucleation
Inoculant
Macrosegregation
Microsegregation
Mold
Mushy zone
Normal segregation
Porosity
Pouring basin
Reynolds number
Rheocasting
Riser
Runner
Segregation
Shrinkage
Skin
Solidification
Sprue
Turbulence

BIBLIOGRAPHY

Alexiades, V., **Mathematical Modeling of Melting and Freezing Processes**, Hemisphere, 1993.

Ammen, C.W., **Metalcasting**, McGraw-Hill, 1999.

Analysis of Casting Defects, American Foundrymen's Society, 1974.

ASM Handbook, Vol. 15: **Casting**, ASM International, 2008.

Campbell, J., **Castings**, Butterworth-Heinemann, 2nd ed., 2003.

Campbell, J., **Castings Practice: The Ten Rules of Casting**, Butterworth-Heinemann, 2004.

Cantor, B., O'Reilly, K., **Solidification and Casting**, Taylor & Francis, 2002.

"Casting," in *Tool and Manufacturing Engineers Handbook*, Volume II: *Forming*, Society of Manufacturing Engineers, 1984.

Kurz, W., and Fisher, D.J., **Fundamentals of Solidification**, Trans Tech Pub., 1994.

Liebermann, H.H. (ed.), **Rapidly Solidified Alloys**, Marcel Dekker, 1993.

Reikher, A., and Barkhudarov, M., **Casting: An Analytical Approach**, Springer, 2008.

Steel Castings Handbook, 6th ed., Steel Founders' Society of America, 1995.

Stefanescu, D.M., **Science and Engineering of Casting Solidification**, Springer, 2002.

Yu, K.-O. (ed.), **Modeling for Casting & Solidification Processing**, CRC Press, 2001.

REVIEW QUESTIONS

10.1. Explain why casting is an important manufacturing process.

10.2. What are the differences between the solidification of pure metals and metal alloys?

10.3. What are dendrites? Why are they called so?

10.4. Describe the difference between short and long freezing ranges.

10.5. What is superheat? Is it important?

10.6. Define shrinkage and porosity. How can you tell whether cavities in a casting are due to porosity or to shrinkage?

10.7. What is the function of chills? What are they made of?

10.8. Why is the Reynolds number important in casting?

10.9. How is fluidity defined? Why is it important?

10.10. Explain the reasons for hot tearing in castings.

10.11. 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?

10.12. Why is Bernoulli's equation important in casting?

10.13. Describe thixocasting and rheocasting.

QUALITATIVE PROBLEMS

10.14. Describe the stages involved in the contraction of metals during casting.

10.15. Explain the effects of mold materials on fluid flow and heat transfer in casting operations.

10.16. 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?

10.17. Describe the events depicted in Fig. 10.5.

10.18. Would you be concerned about the fact that portions of internal chills are left within the casting? Explain.

10.19. What practical demonstrations can you suggest to indicate the relationship of the solidification time to the volume and surface area of a casting?

10.20. Explain why a casting may have to be subjected to various heat treatments.

10.21. 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.

10.22. 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.

10.23. 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.

10.24. Explain why the constant C in Eq. (10.7) depends on mold material, metal properties, and temperature.

10.25. Are external chills as effective as internal chills? Explain.

10.26. Explain why, as shown in Table 10.1, gray cast iron undergoes expansion rather than contraction during solidification.

10.27. Referring to Fig. 10.11, explain why internal corners, such as A, develop a thinner skin than external corners, such as B, during solidification.

10.28. Note the shape of the two risers shown in Fig. 10.8, and discuss your observations with respect to Eq. (10.7).

10.29. Is there any difference in the tendency for shrinkage void formation in metals with short and long freezing ranges, respectively? Explain.

10.30. What is the influence of the cross sectional area of the spiral channel shown in Fig. 10.9 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.

10.31. 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.

10.32. 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?

10.33. 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.

10.34. How can you tell whether cavities in a casting are due to shrinkage or entrained air bubbles?

10.35. Describe the drawbacks to having a riser that is (a) too large and (b) too small.

10.36. Reproduce Fig. 10.2 for a casting that is spherical in shape.

10.37. List the process variables that affect the fluidity index as shown in Fig. 10.9.

10.38. 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

10.39. 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.

▶ **10.40.** A round casting is 0.2 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?

10.41. A 100-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.

▶ **10.42.** 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 6 mm at the bottom.

▶ **10.43.** Pure aluminum is poured into a sand mold. The metal level in the pouring basin is 200 mm above the metal level in the mold, and the runner is circular with a 7.5-mm diameter. What are the velocity and rate of the flow of the metal into the mold? Is the flow turbulent or laminar?

10.44. A cylinder with a diameter of 40 mm and height of 80 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?

▶ **10.45.** The volume flow rate of metal into a mold is 0.02 m^3/s. The top of the sprue has a diameter of 20 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?

10.46. A rectangular mold with dimensions 100 mm × 200 mm × 400 mm is filled with copper with no superheat. Determine the final dimensions of the part as it cools to room temperature. Repeat the analysis for gray cast iron.

▶ **10.47.** The constant C in Chvorinov's rule is given as 3 s/mm^2 and is used to produce a cylindrical casting with a diameter of 50 mm and height of 125 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?

▶ **10.48.** A sprue is 300 mm long and has a diameter of 125 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 650 mm^3/s is to be achieved, what should be the diameter at the bottom of the sprue? Will the sprue aspirate? Explain.

▶ **10.49.** 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 10-mm diameter. What are the velocity and rate of the flow of the metal into the mold? Is the flow turbulent or laminar?

▶ **10.50.** For the sprue described in Problem 10.49, what runner diameter is needed to ensure a Reynolds number of 2000? How long will a 330 cm^3 casting take to fill with such a runner?

▶ **10.51.** How long would it take for the sprue in Problem 10.49 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.

10.52. 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.

SYNTHESIS, DESIGN, AND PROJECTS

10.53. Can you devise fluidity tests other than that shown in Fig. 10.9? Explain the features of your test methods.

10.54. Figure P10.54 indicates various defects and discontinuities in cast products. Review each defect and offer solutions to avoid it.

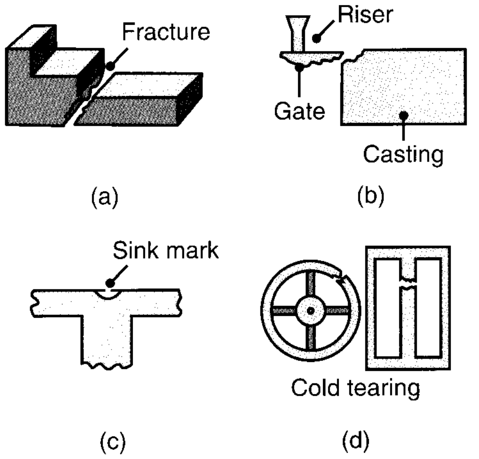


FIGURE P10.54

10.55. The fluidity test shown in Fig. 10.9 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.

10.56. Utilizing the equipment and materials available in a typical kitchen, design an experiment to reproduce results similar to those shown in Fig. 10.11. Comment on your observations.

10.57. 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.

▶ **10.58.** 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.

10.59. 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.

10.60. 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.

10.61. 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

- Building upon the fundamentals of solidification, fluid flow, and heat transfer discussed in Chapter 10, this chapter presents the principles of 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 discussion of inspection techniques for castings.

Typical products made by casting: engine blocks, crankshafts, hubcaps, power tool housings, turbine blades, plumbing parts, zipper teeth, dies and molds, gears, railroad wheels, propellers, office equipment, and statues.

Alternative processes: forging, powder metallurgy, machining, and fabrication.

11.1 Introduction

The first metal castings were made during the period from 4000 to 3000 B.C., using stone and metal molds for casting copper. Various casting processes have been developed over time, each with its own characteristics and applications (see also Fig. I.6a), to meet specific design requirements (Table 11.1). A large variety of parts and components are made by casting, such as engine blocks, crankshafts, automotive components and powertrains (Fig. 11.1), agricultural and railroad equipment, pipes and plumbing fixtures, power-tool housings, gun barrels, frying pans, jewelry, orthopedic implants, and very large components for hydraulic turbines.

Two trends have had a major impact on the casting industry. The first is the 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 traditional methods of casting. The second major trend has been the increasing demand for high-quality castings with close dimensional tolerances.

This chapter is organized around the major classifications of casting practices (see Fig. II.3 in the Introduction to Part II). These classifications are related to mold

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	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

materials, pattern production, molding processes, and methods of feeding the mold with molten metal. The major categories are as follows:

1. **Expendable molds**, which typically are made of sand, plaster, ceramics, and similar materials and generally are mixed with various binders (*bonding agents*) for improved properties. A typical sand mold consists of 90% sand, 7% clay, and 3% water. As described in Chapter 8, 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.

   The mold is produced from a pattern; in some processes, such as sand and shell casting, 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 consumes a pattern for each mold produced; it is an example of an *expendable-mold, expendable-pattern process.*

2. **Permanent molds**, which are made of metals that maintain their strength at high temperatures. As the name implies, they are used repeatedly and are designed in such a way 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); hence, the solidifying casting is subjected to a higher rate of cooling, which in turn affects the microstructure and grain size within the casting.

3. **Composite molds**, which are 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 various casting processes to improve mold strength, control the cooling rates, and optimize the overall economics of the casting process.

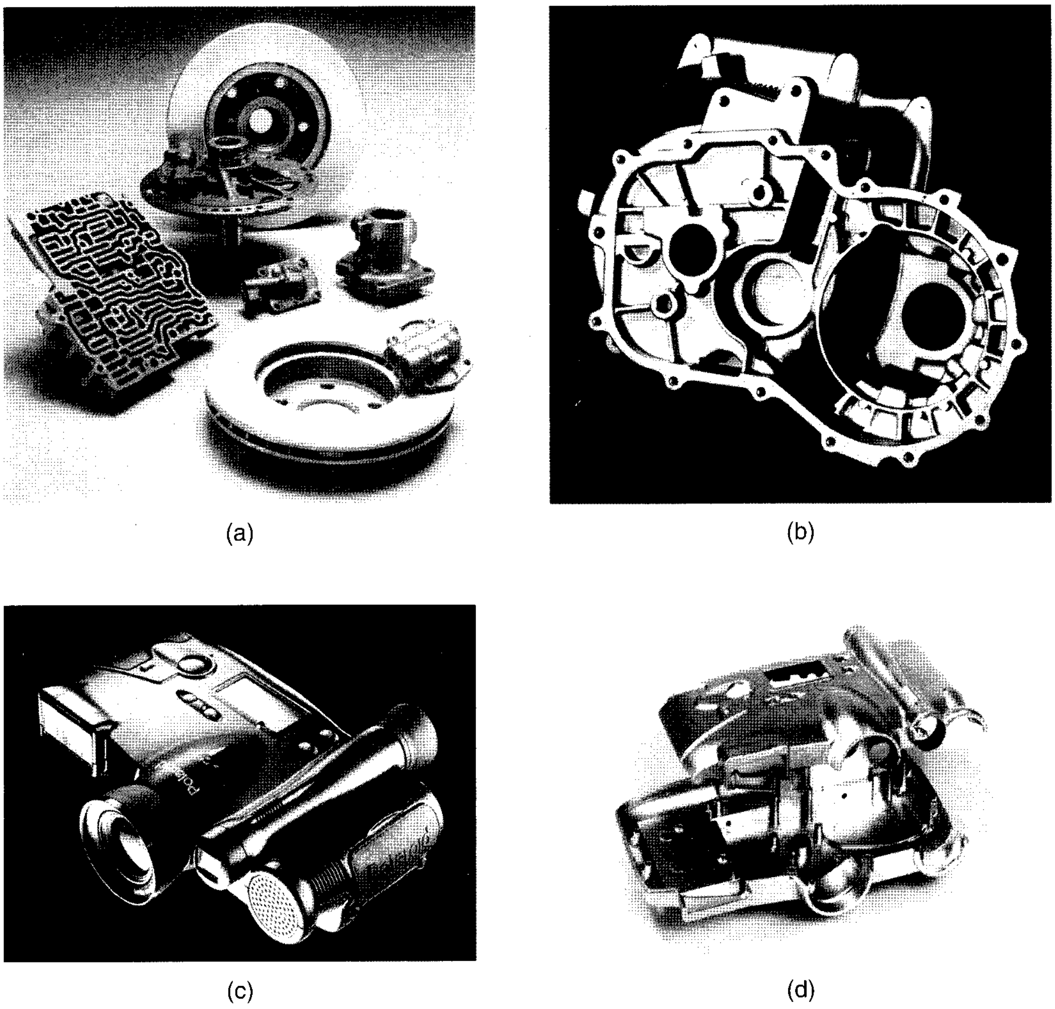

(a) (b) (c) (d)

FIGURE 11.1 (a) Typical gray-iron castings used in automobiles, including the transmission valve body (left) and the hub rotor with disk-brake cylinder (front). *Source*: Courtesy of Central Foundry Division of General Motors Corporation. (b) A cast transmission housing. (c) The Polaroid PDC-2000 digital camera with an AZ191D die-cast, high-purity magnesium case. (d) A two-piece Polaroid camera case made by the hot-chamber die-casting process. *Source*: (c) and (d) Courtesy of Polaroid Corporation and Chicago White Metal Casting, Inc.

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 cast-iron engine blocks and very large propellers for ocean liners, can be made by casting.

Because of their unique characteristics and applications, particularly in manufacturing microelectronic devices (Part V), basic crystal-growing techniques also are described in this chapter, which concludes with a brief overview of modern foundries.

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). For example, 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 only general; significant variations can occur, depending on the manufacturing methods used.

2. Approximate values without the use of rapid prototyping technologies. Minimum quantity is 1 when applying rapid prototyping.

Source: Data taken from J.A. Schey, *Introduction to Manufacturing Processes*, 3d ed., McGraw-Hill, 2000.

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.

11.2.1 Sand Casting

The traditional method of casting metals is in sand molds and has been used for millennia. Sand casting is still the most prevalent form of casting; in the United States alone, about 15 million tons of metal are cast by this method each year. Typical applications of sand casting include machine bases, large turbine impellers, propellers, plumbing fixtures, and a wide variety of other products and components. The capabilities of sand casting are given in Table 11.2.

Basically, *sand casting* consists of (a) placing a pattern (having the shape of the desired casting) in sand to make an imprint, (b) incorporating a gating system, (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).

Sands. Most sand-casting operations use silica sand (SiO_2) as the mold material. 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. For proper functioning, mold sand must be clean and preferably new.

Several factors are important in the selection of sand for molds, and certain tradeoffs with respect to properties are involved. Sand having fine, round grains can be packed closely and, thus, forms a smooth mold surface. Although fine-grained sand enhances mold strength, the fine grains also lower mold *permeability* (where fluids and gases penetrate through pores). Good permeability of molds and

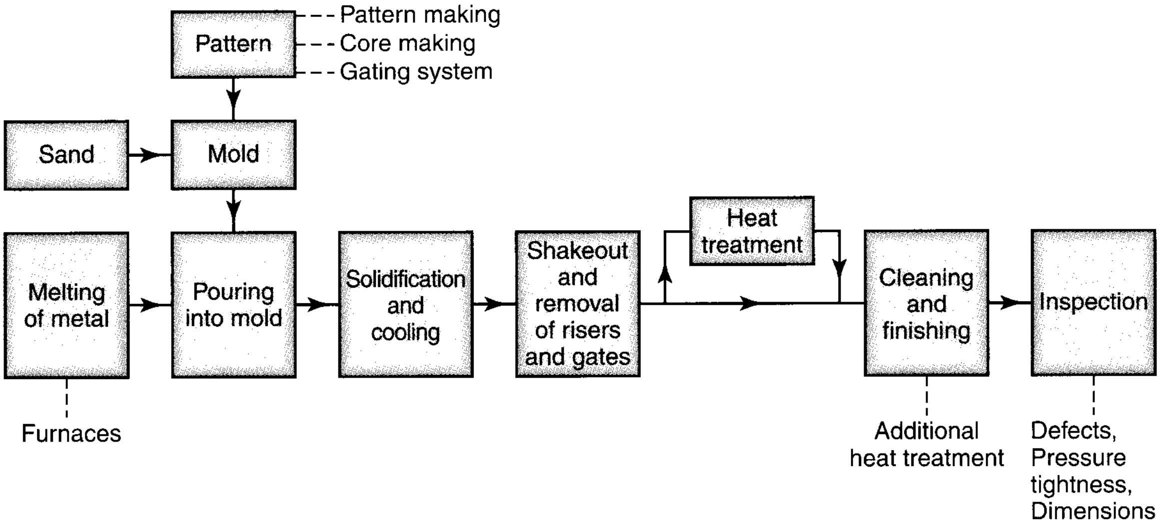


FIGURE 11.2 Outline of production steps in a typical sand-casting operation.

cores allows gases and steam evolved during the casting to escape easily. The mold also should have good *collapsibility* to allow the casting to shrink while cooling and, thus, to avoid defects in the casting, such as hot tearing and cracking (see Fig. 10.12).

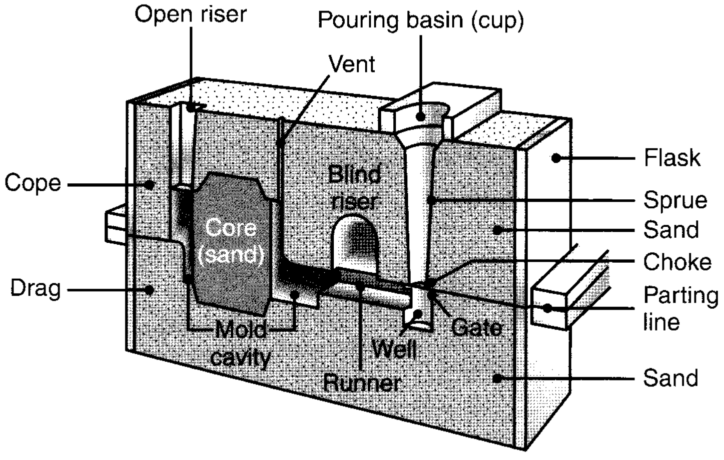


FIGURE 11.3 Schematic illustration of a sand mold, showing various features.

Types of Sand Molds. Sand molds (Fig. 11.3) are characterized by the types of sand that comprise them and by the methods used to produce them. There are three basic types of sand molds: green-sand, cold-box, and no-bake molds. The most common mold material is **green molding sand**, which is 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 generally are 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 heat, they are called **cold-setting processes.**

Sand molds can be oven dried (*baked*) prior to pouring the molten metal; they are then 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) the production rate is lower because of the considerable drying time required.

The major features of molds in sand casting are as follows:

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.
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.
6. **Cores**, which are inserts made from sand. They are 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 form features such as lettering on the surface or deep external pockets.

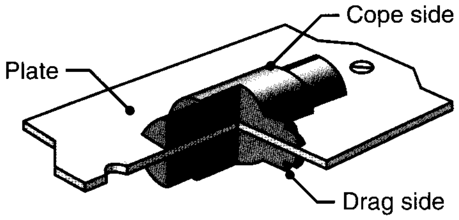


FIGURE 11.4 A typical metal match-plate pattern used in sand casting.

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. Vents 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, or metal. The selection of a pattern material 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 will produce. Patterns may be made of a combination of materials to reduce wear in critical regions, and they usually are coated with a **parting agent** to facilitate the removal of the casting from the molds.

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*, generally are used for simpler shapes and low-quantity production; they generally are made of wood and are inexpensive. **Split patterns** are two-piece patterns, made such that 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 of pattern is used most often in conjunction with molding machines and large production runs to produce smaller castings.

An important development in molding and pattern making is the application of **rapid prototyping** (Chapter 20). In sand casting, for example, a pattern can be fabricated in a rapid-prototyping machine and fastened to a backing plate at a fraction of the time and cost of machining a pattern. There are several rapid prototyping techniques with which these tools can be produced quickly.

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. (These topics are described in greater detail in Chapter 12.)

Cores. For castings with internal cavities or passages, such as those found in an automotive engine block or a valve body, *cores* are utilized. Cores are placed in the mold cavity to form the interior surfaces of the casting and are removed from the finished part during shakeout and further processing. Like molds, cores must possess strength, permeability, the ability to withstand heat, and collapsibility; hence, cores are made of sand aggregates. The core is anchored by **core prints**, which are recesses added to the pattern to locate and support the core and to provide vents for the escape of gases (Fig. 11.6a). A common problem 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 generally are made in a manner similar to that used in sand moldmaking; the majority are made

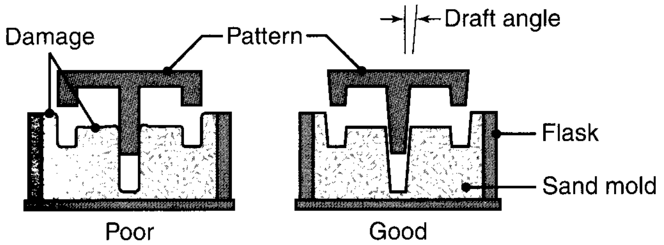


FIGURE 11.5 Taper on patterns for ease of removal from the sand mold.

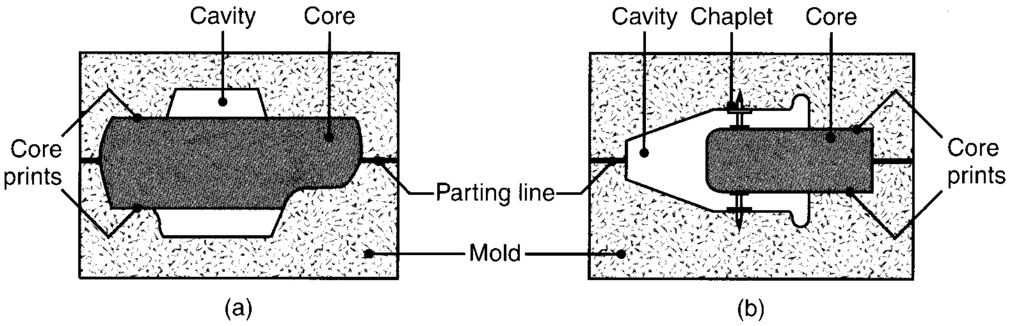


FIGURE 11.6 Examples of sand cores, showing core prints and chaplets to support the cores.

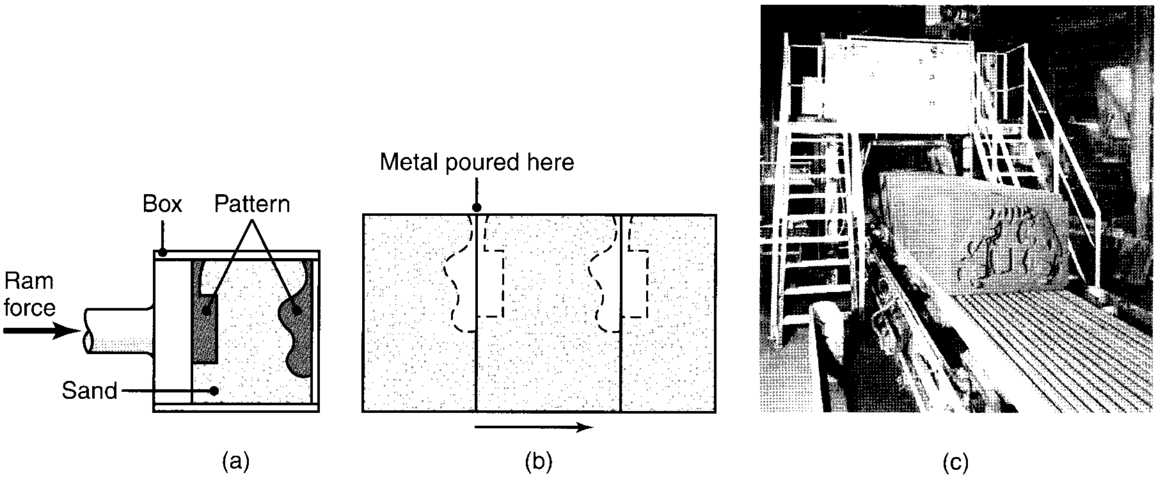


FIGURE 11.7 Vertical flaskless molding. (a) Sand is squeezed between two halves of the pattern. (b) Assembled molds pass along an assembly line for pouring. (c) A photograph of a vertical flaskless molding line. *Source*: Courtesy of American Foundry Society.

with shell (see Section 11.2.2), no-bake, or cold-box processes. Cores are shaped in *core boxes*, which are 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, is to compact the sand by hand hammering (*tamping*) or ramming it around the pattern. For most operations, however, the sand mixture is compacted around the pattern by *molding machines*. These machines eliminate arduous labor, offer high-quality casting by improving the application and distribution of forces, manipulate the mold in a carefully controlled manner, and increase production rate.

In **vertical flaskless molding**, the halves of the pattern form a vertical chamber wall against which sand is blown and compacted (Fig. 11.7). Then the mold halves are packed horizontally, with the parting line oriented vertically, and moved along a pouring conveyor. This 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 automated.

Sandslingers fill the flask uniformly with sand under a high-pressure stream; they are used to fill large flasks and are operated typically by machine. 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 appropriately.

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*), the pattern is covered tightly with a thin sheet of plastic. A flask is placed over the coated pattern and is filled with dry, binderless sand. A second sheet of plastic then is placed on top of the sand, a vacuum action compacts the sand, and the pattern can then be withdrawn. Both halves of the mold are made in this manner and subsequently assembled. During pouring, the mold remains under vacuum, but the casting cavity does not. When the metal has solidified, the vacuum is turned off and the sand falls away, releasing the casting. Vacuum molding produces castings with high-quality surface detail and dimensional accuracy; it is suited especially well for large, relatively flat (plane) castings.

The Sand-casting Operation. After the mold has been shaped and the cores have been placed in position, the two 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. 11.8.

After solidification, the casting is shaken out of its mold, and the sand and oxide layers adhering to the casting are removed by vibration (using a shaker) or by sand blasting. Castings also are cleaned by blasting with steel shot or grit (*shot blasting*; Section 26.8). The risers and gates are cut off by oxyfuel-gas cutting, sawing, shearing, or abrasive wheels; or they are trimmed in dies. Gates and risers on steel castings also may be removed with air carbon-arc cutting (Section 30.8) or torches. Castings may be cleaned further by electrochemical means or by pickling with chemicals to remove surface oxides.

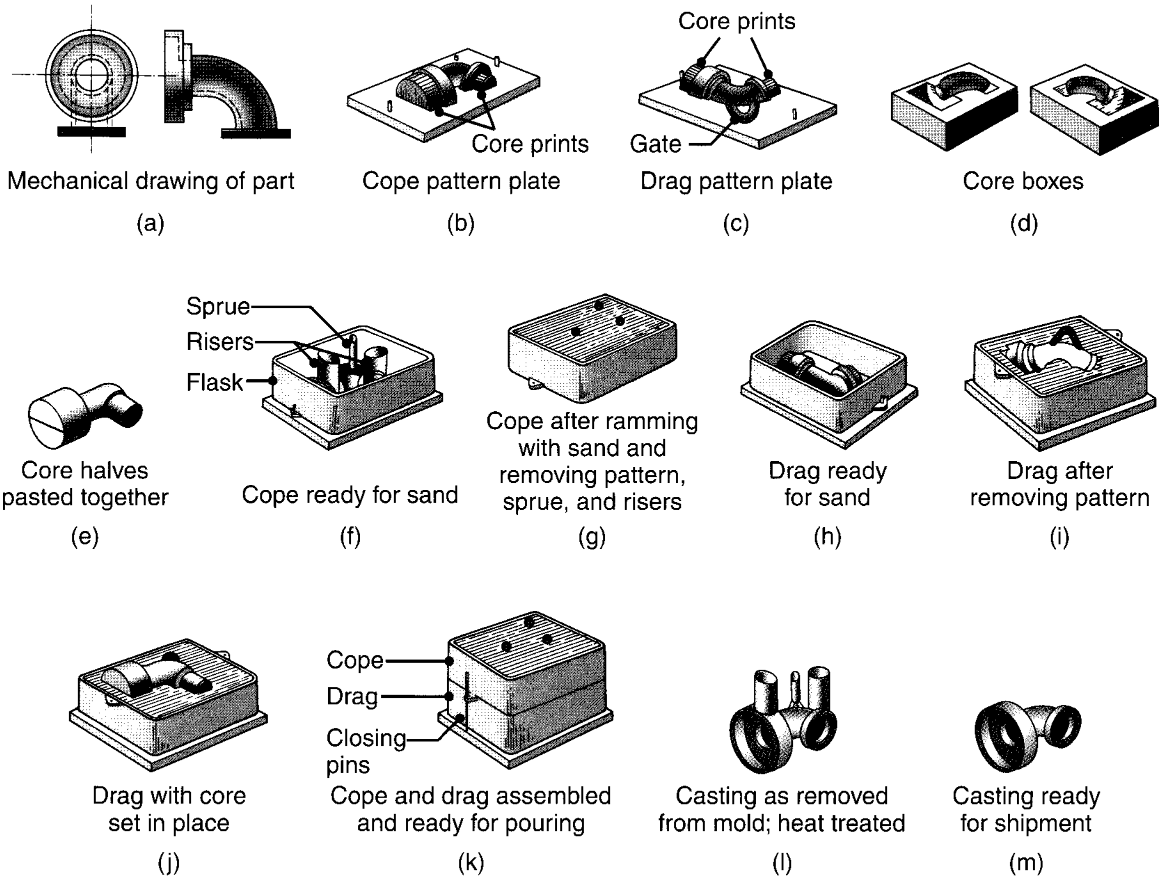
The casting subsequently may be *heat treated* to improve certain properties required for its intended use; heat-treatment is particularly important for steel castings. *Finishing operations* may involve machining, straightening, or forging with dies (sizing) to obtain final dimensions. *Inspection* 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.

11.2.2 Shell Molding

Shell molding was first developed in the 1940s and has grown significantly because it can produce many types of castings with close dimensional tolerances and a good surface finish at low cost. Shell-molding applications include small mechanical parts requiring high precision, such as gear housings, cylinder heads, and connecting rods. The process also is 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° to 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



Mechanical drawing of part (a)

Cope pattern plate (b)

Drag pattern plate (c)

Core boxes (d)

Core halves pasted together (e)

Cope ready for sand (f)

Cope after ramming with sand and removing pattern, sprue, and risers (g)

Drag ready for sand (h)

Drag after removing pattern (i)

Drag with core set in place (j)

Cope and drag assembled and ready for pouring (k)

Casting as removed from mold; heat treated (l)

Casting ready for shipment (m)

FIGURE 11.8 Schematic illustration of the sequence of operations for sand casting. (a) A mechanical drawing of the part is used to generate a design for the pattern. Considerations such as part shrinkage and draft must be built into the drawing. (b–c) Patterns have been mounted on plates equipped with pins for alignment. Note 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 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 liquid, 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 (when necessary). *Source*: Courtesy of Steel Founders' Society of America.

with 2.5 to 4% of a thermosetting resin binder (such as phenol-formaldehyde) that coats the sand particles. Either the box is rotated upside down (Fig. 11.9), 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 the 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 to cure it. The shell hard-

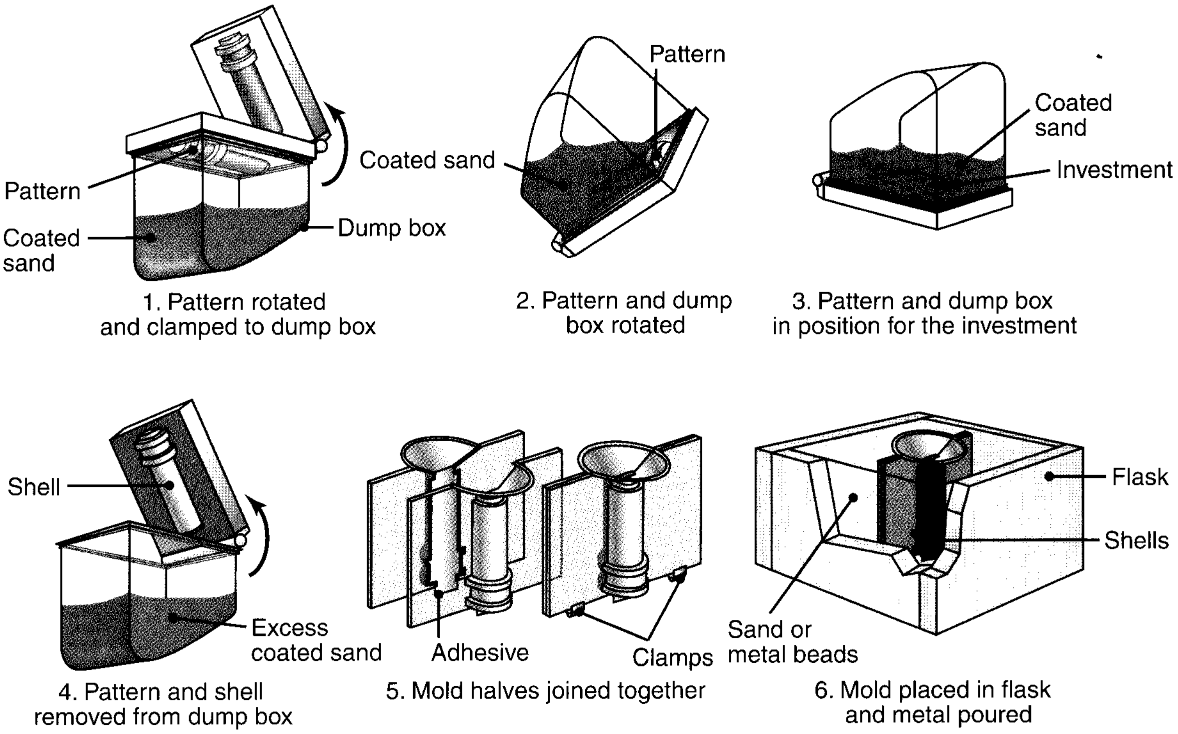


FIGURE 11.9 The shell-molding process, also called the *dump-box* technique.

ens around the pattern and is removed from the pattern using 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.

Shell sand has a much lower permeability than the sand used for green-sand molding, because a sand of much smaller grain size is used for shell molding. The decomposition of the shell-sand binder also produces a high volume of gas. Consequently, unless the molds are vented properly, trapped air and gas can cause serious problems in the 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 fairly easily.

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. The castings usually weigh less than 10 kg and are typically in the range of 125 to 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 flour to improve strength and to control the time required for the plaster to set. These components are mixed with water, and the resulting slurry is poured over the pattern. After the plaster sets (usually within 15 minutes), it is removed, and the mold is dried at a temperature range of 120° to 260°C. Higher drying temperatures may be used, depending on the type of plaster. The mold halves are 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 substantially 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 generally are made of materials such as aluminum alloys, thermosetting plastics, brass, or zinc alloys. Wood patterns are not suitable for making a large number of molds, because they are repeatedly in contact with the water-based plaster slurry and warp or degrade quickly. 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. Because plaster molds have lower thermal conductivity than other mold materials, the castings cool slowly, and thus a more uniform grain structure is obtained with less warpage. The wall thickness of the cast parts can be 1 to 2.5 mm.

11.2.4 Ceramic-mold Casting

The *ceramic-mold casting* 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. Typical parts made are impellers, cutters for machining operations, dies for metalworking, and molds for making plastic and rubber components. Parts weighing as much as 700 kg have been cast by this process.

The slurry is a mixture of fine-grained zircon ($ZrSiO_4$), aluminum oxide, and fused silica, which are mixed with bonding agents and poured over the pattern (Fig. 11.10), which has been placed in a flask.

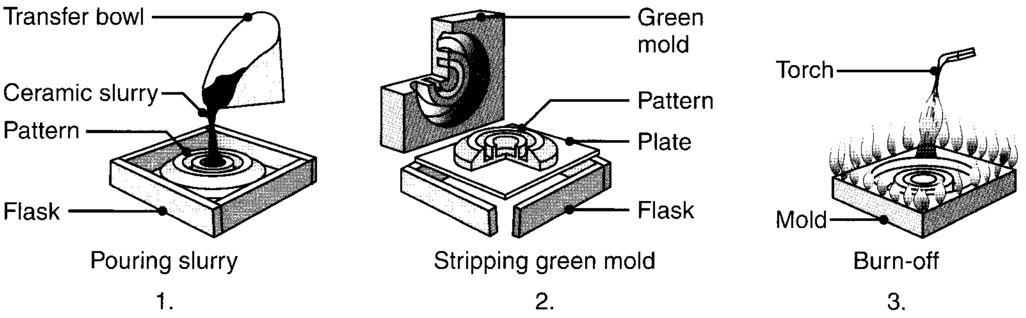


FIGURE 11.10 Sequence of operations in making a ceramic mold. *Source*: *Metals Handbook*, Vol. 5, 8th ed.

The pattern may be made of wood or metal. After setting, the molds (ceramic facings) are removed, dried, ignited to burn off volatile matter, and baked. The molds are 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 then are assembled into a complete mold, ready to be poured.

The high-temperature resistance of the refractory molding materials allows these 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 sometimes referred to as *expendable-pattern* casting processes or *expendable mold–expendable pattern* processes. They are unique in that a mold and a pattern must be produced for each casting, whereas the patterns in the processes described in the preceding section are reusable. Typical applications are cylinder heads, engine blocks, crankshafts, brake components, manifolds, and machine bases.

11.3.1 Evaporative-pattern Casting (Lost-foam Process)

The *evaporative-pattern casting* 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* and falls under the trade name *full-mold* 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 to 8% pentane (a volatile hydrocarbon) are placed in a preheated die that is usually made of aluminum. 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 then cooled and opened, and the polystyrene pattern is removed. Complex patterns also may be made by bonding various individual pattern sections using hot-melt adhesive (Section 32.4.1).

The pattern is coated with a water-based refractory slurry, dried, and placed in a flask. The flask is then filled with loose, fine sand, which surrounds and supports the pattern (Fig 11.11) 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 products from the polystyrene are vented into the surrounding sand.

The flow velocity of the molten metal in the mold depends on the rate of degradation of the polymer. Studies have shown that the flow of the metal is basically laminar, with Reynolds numbers in the range of 400 to 3000. The velocity of the molten metal at the metal–polymer pattern front (interface) is in the range of 0.1 to 1.0 m/s and can be controlled by producing patterns with cavities or hollow sections. Thus, the velocity will increase as the molten metal crosses these hollow regions, similar to pouring the metal into an empty cavity.

Because the polymer requires considerable energy to degrade, large thermal gradients are present at the metal–polymer interface. 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

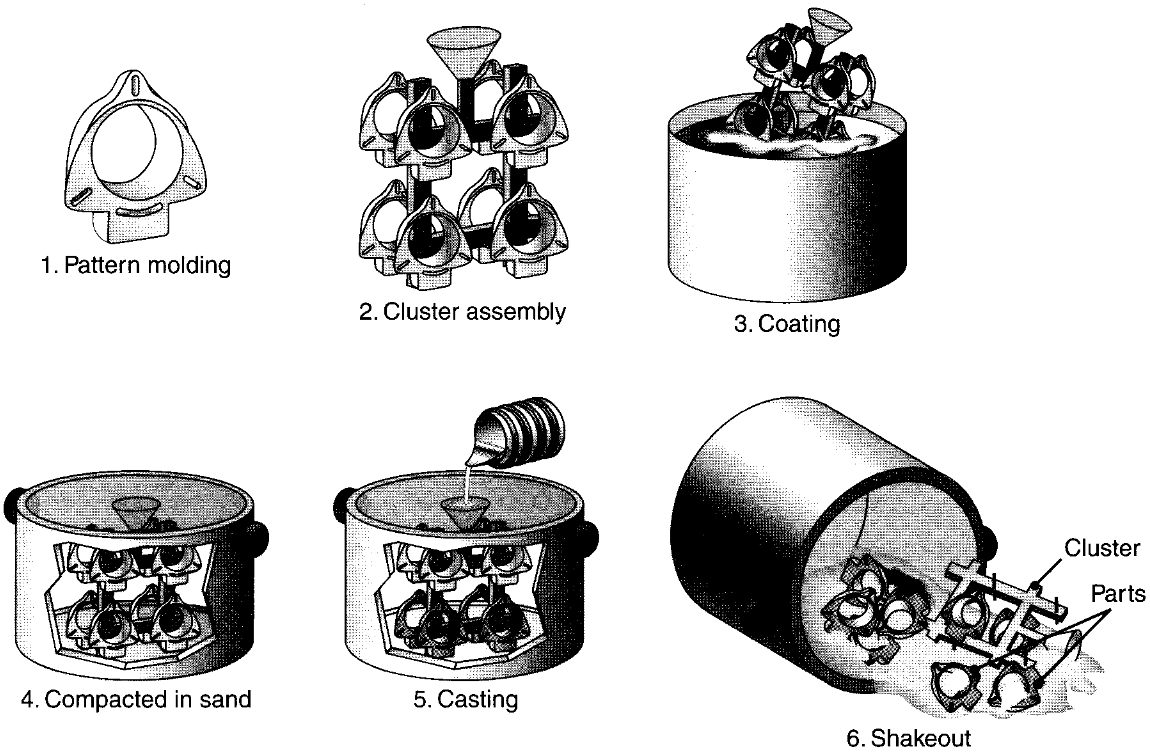


FIGURE 11.11 Schematic illustration of the expendable-pattern casting process, also known as lost-foam or evaporative-pattern casting.

microstructure throughout the casting and also leads to directional solidification of the metal. Polymethylmethacrylate (PMMA) and polyalkylene carbonate also may be used as pattern materials for ferrous castings.

The evaporative-pattern process has a number of advantages over other casting methods:

- The process is relatively simple because there are no parting lines, cores, or riser systems. Hence, it has design flexibility.
- Inexpensive flasks are satisfactory for the process.
- 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, major factors are the cost to produce the die used for expanding the polystyrene beads to make the pattern and the need for two sets of tooling.

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 (using wax patterns, Section 11.3.2) is that carbon pickup into the metal is avoided entirely. Further developments in evaporative-pattern casting include the production of metal-matrix composites (Sections 9.5 and 19.14). During the molding of the polymer pattern, fibers or particles are embedded throughout, which then become an integral part of the casting. Additional techniques include the modification and grain refinement of the casting by using grain refiners and modifier master alloys within the pattern while it is being molded.

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; it provides the basic structure that encloses the pistons and cylinders and thus encounters significant pressure during operation. Recent industry trends have focused upon high-quality, low-cost lightweight designs; additional economic benefits can be attained through casting more complex geometries and by incorporating multiple components into one part. Recognizing that evaporative-pattern (lost-foam) casting 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 and illustrated in Fig. 11.12a. 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 on each block, respectively. Lost-foam casting also allowed consolidation of the engine's cylinder head and exhaust and cooling systems into the block, thus eliminating the associated machining operations and fasteners required in sand-cast or die-cast designs. In addition, since the pattern contained holes, and these could be cast without the use of 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 a cylinder block and head integration, featuring hollow sections for water jacket cooling that could not be cored out in die casting or semipermanent mold processes (used for its other V6 blocks). Based on the success the foundry had with the three-cylinder lost-foam block, engineers applied this process for casting the V6 die block (Fig. 11.12b). The new engine block involves only one casting that is lighter and cheaper than the previous designs. Produced with an integrated cylinder head and exhaust and cooling system, this component is cast hollow to create more efficient water jacket cooling of the engine during its operation.

The company also has 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 as compared to conventional lost-foam casting.

Source: Courtesy of Mercury Marine.

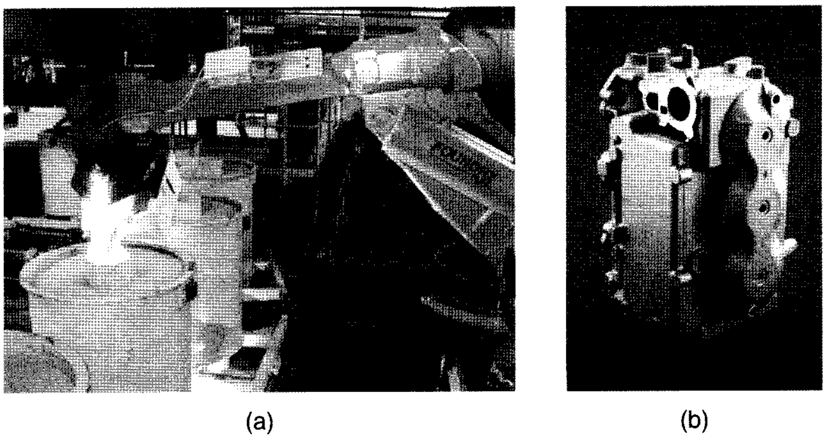

(a) (b)

FIGURE 11.12 (a) Metal is poured into a mold for lost-foam casting of a 60-hp, three-cylinder marine engine; (b) finished engine block. *Source*: 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, as well as 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 by this process. The capabilities of investment casting are given in Table 11.2.

The sequence involved in investment casting is shown in Fig. 11.13. The pattern is made of wax, or of a plastic such as polystyrene, by molding or rapid-prototyping techniques. The pattern 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 better strength. Note that the initial coating can use smaller particles to develop a better surface finish in the casting; subsequent layers use larger particles and are intended to build 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 strong enough to withstand the forces encountered during mold making; however, unlike plastic patterns, wax can be recovered and reused.

The one-piece mold is 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

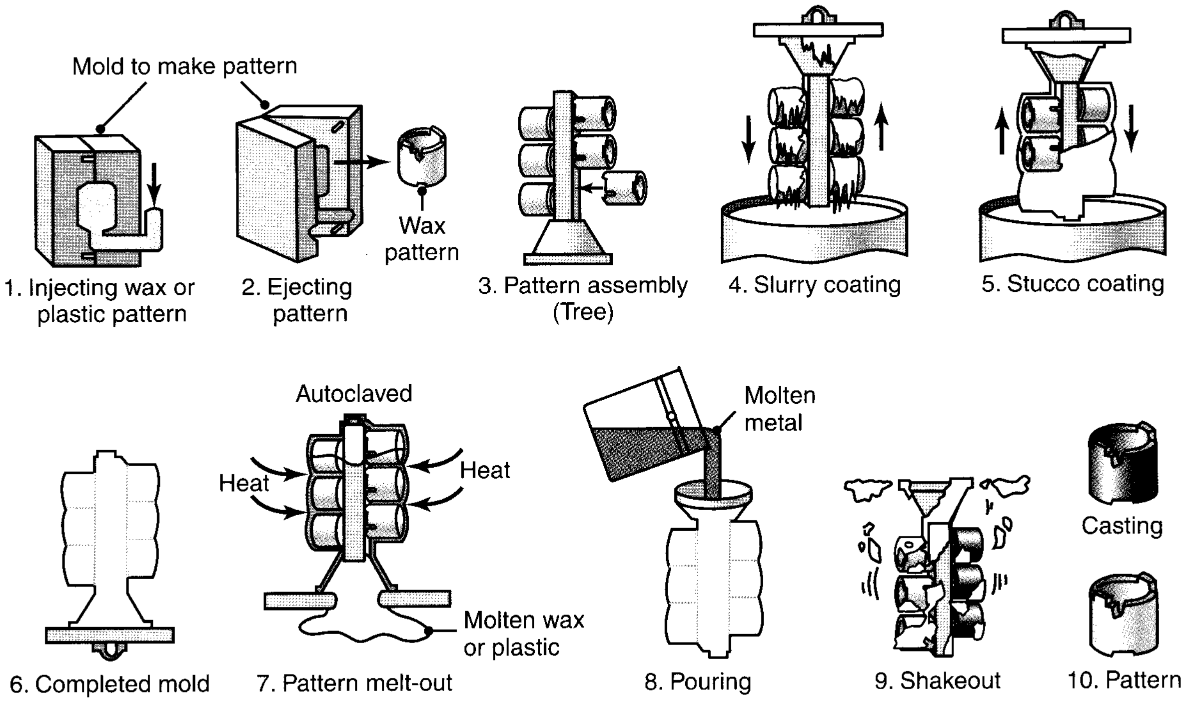


FIGURE 11.13 Schematic illustration of the investment-casting (lost-wax) process. Castings produced by this method can be made with very fine detail and from a variety of metals. *Source*: Courtesy of Steel Founders' Society of America.

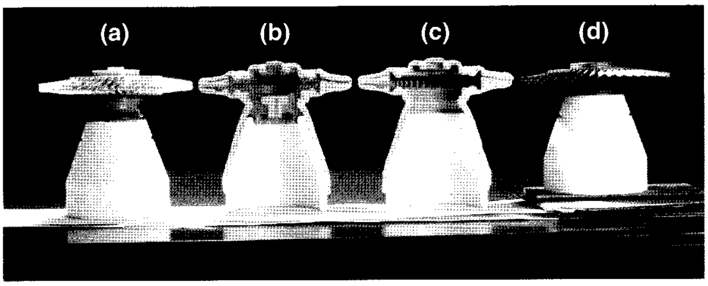


FIGURE 11.14 Investment casting of an integrally cast rotor for a gas turbine. (a) Wax pattern assembly. (b) Ceramic shell around wax pattern. (c) Wax is melted out and the mold is filled, under a vacuum, with molten superalloy. (d) The cast rotor, produced to net or near-net shape. *Source*: Courtesy of Howmet Corporation.

up and the casting is removed. A number of patterns can be joined to make one mold, called a **tree** (Fig. 11.13), significantly increasing the production rate. For small parts, the tree can be inserted into a permeable flask and filled with a liquid slurry investment. The investment then is placed into a chamber and evacuated (to remove the air bubbles in it) until the mold solidifies. The flask usually is placed in a vacuum-casting machine, so that molten metal is drawn into the permeable mold and onto the part, producing fine detail.

Although the mold materials and 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, which otherwise would add significantly to the total cost of the casting, are required. The process is capable of producing intricate shapes, with parts weighing from 1 g to 35 kg, from a wide variety of ferrous and nonferrous metals and alloys. Recent 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.

Ceramic-shell Investment Casting. A variation of the investment-casting process is *ceramic-shell casting*. It uses the same type of wax or plastic pattern, which is dipped first in ethyl silicate gel and subsequently into a fluidized bed (see Section 4.12) of fine-grained fused silica or zircon flour. 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. 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. 11.14. If ceramic cores are used in the casting, they are removed by leaching with caustic solutions under high pressure and temperature. The molten metal may 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. Aluminum castings, for example, are subjected to a gas pressure up to 100 MPa at 500°C.

EXAMPLE 11.1 Investment-cast Superalloy Components for Gas Turbines

Since the 1960s, investment-cast superalloys have been replacing wrought counterparts in high-performance gas turbines. The microstructure of an integrally investment-cast gas-turbine rotor is shown in the upper half of Fig. 11.15. Note the fine, uniform equiaxed grains throughout the rotor cross section. Casting procedures include the use of a nucleant addition to the molten metal, as well as close control of its superheat, pouring techniques, and the cooling rate of the casting (see Section 10.2). In contrast, note the coarse-grained structure in the lower half of Fig. 11.15 showing the same type of rotor cast conventionally. This rotor has inferior properties compared with the fine-grained rotor. Due to developments in these processes, the proportion of cast parts to other parts in aircraft engines has increased from 20% to about 45% by weight.

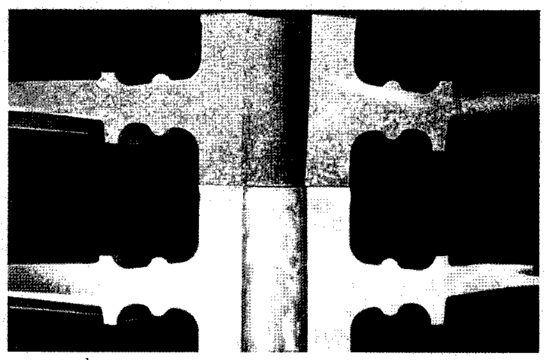


FIGURE 11.15 Cross section and microstructure of two rotors: (top) investment cast; (bottom) conventionally cast. *Source*: Courtesy of ASM International.

CASE STUDY 11.2 Investment Casting of Total Knee Replacements

With great advances in medical care in the past few decades, life expectancies have increased considerably, but so have expectations that the quality of life in later years will remain high. One of the reasons for improvements in the quality of life in the past 40 years has been the great success of orthopedic implants. Hip, knee, shoulder, spine, and other implants have resulted in increased activity and reduced pain for millions of people worldwide.

An example of an orthopedic implant that has greatly improved patient quality of life is the total knee replacement (TKR), as shown in Fig. 11.16a. 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; Section 7.6). Each material is chosen for specific properties that are important in the application of the device.

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 the injection molding of patterns, which are then hand assembled onto trees, as shown in Fig. 11.16b. The patterns are spaced properly on a central wax sprue and then are 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 is shown in Fig. 11.17a, which contains 12 knee implants arranged into four rows.

The completed trees are 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 a dilute slurry and 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 that 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 then placed into another furnace to preheat them in preparation for the casting process.

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. The machine then rotates so that the metal flows into the mold, as shown in Fig. 11.16d. The tree is then allowed to cool and the mold is removed. The cast parts are machined from the tree and then are further machined and

(*continued*)

polished to the required dimensional tolerance and surface finish. Figure 11.17 shows the progression of investment casting, from tree, to investment, to casting. The parts are then removed from the tree and subjected to finishing operations to obtain the final product.

Source: Courtesy of M. Hawkins, Zimmer, Inc.

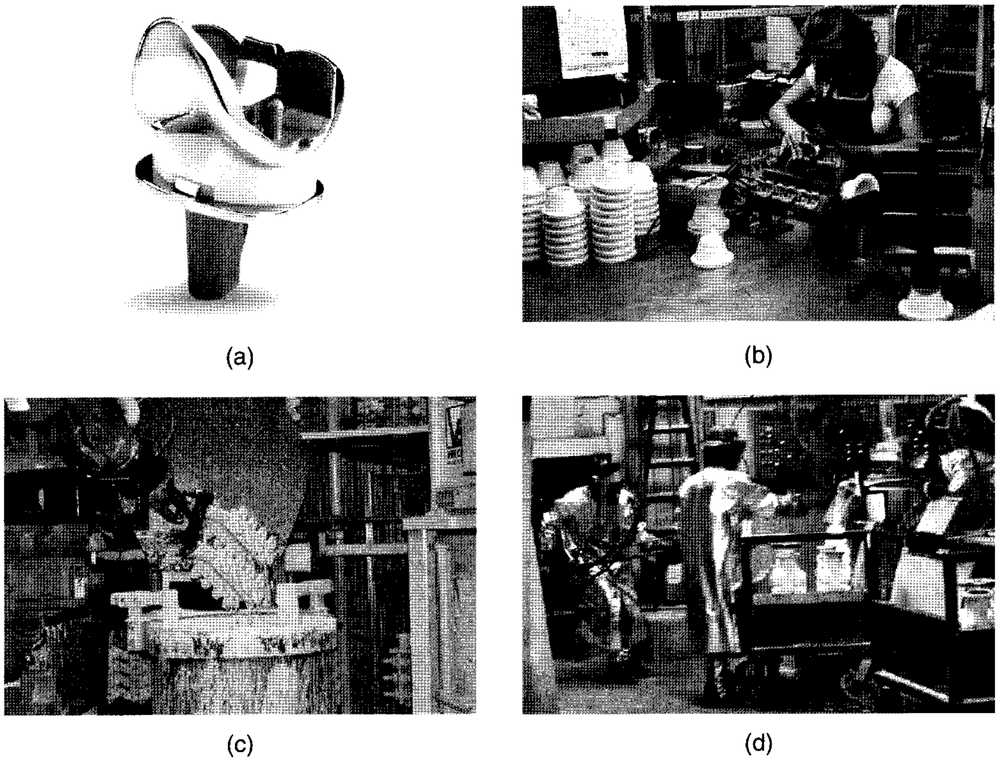

(a) (b) (c) (d)

FIGURE 11.16 Manufacture of total knee replacements. (a) The Zimmer NexGen mobile-bearing knee (MBK); the femoral portion of the total knee replacement is the subject of the case study. (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.

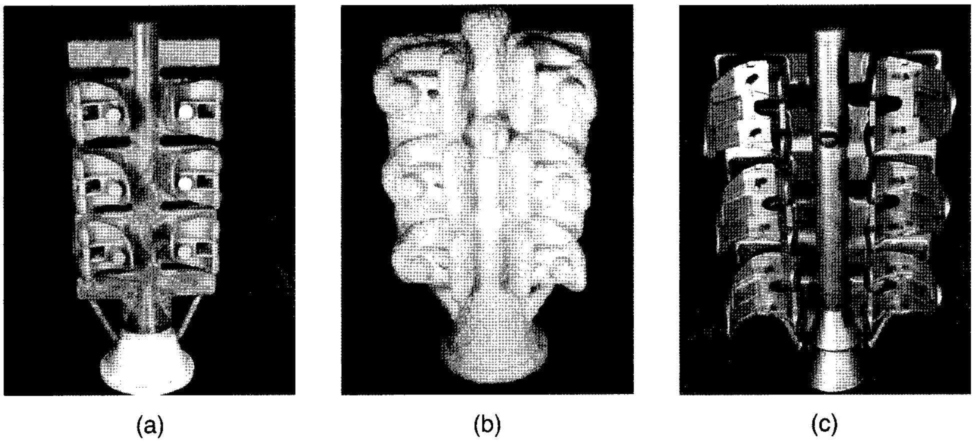

(a) (b) (c)

FIGURE 11.17 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, Inc.

11.4 Permanent-mold Casting Processes

Permanent-mold casting processes have certain advantages over other casting processes.

11.4.1 Permanent-mold Casting

In *permanent-mold casting* (also called *hard-mold casting*), two halves of a mold are made from materials with high resistance to erosion and thermal fatigue, such as cast iron, steel, bronze, graphite, or refractory metal alloys. Typical parts made are automobile pistons, cylinder heads, connecting rods, gear blanks for appliances, and kitchenware. Parts that can be made economically generally 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 11.2.

The mold cavity and gating system are machined into the mold and thus become an integral part of it. 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 for various parts of the mold.

In order to increase the life of permanent molds, the surfaces of the mold cavity usually are coated with a refractory slurry (such as sodium silicate and clay) or 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.

The 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 due to high-temperature gradients. 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 of 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. The process is used mostly for aluminum, magnesium, and copper alloys, as well as for gray iron, because of their generally lower melting points, although 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, easily collapsable sand cores can be used, which are then removed from castings, leaving intricate internal cavities. This process then is called **semipermanent-mold casting.**

11.4.2 Vacuum Casting

A schematic illustration of the *vacuum-casting* process, or *countergravity low-pressure* (CL) *process* (not to be confused with the vacuum-molding process described in Section 11.2.1) is shown in Fig. 11.18. Vacuum casting is an alternative to investment, shell-mold, and green-sand casting and is suitable particularly for thin-walled (0.75 mm) complex shapes with uniform properties. Typical parts made are superalloy gas-turbine components with walls as thin as 0.5 mm.

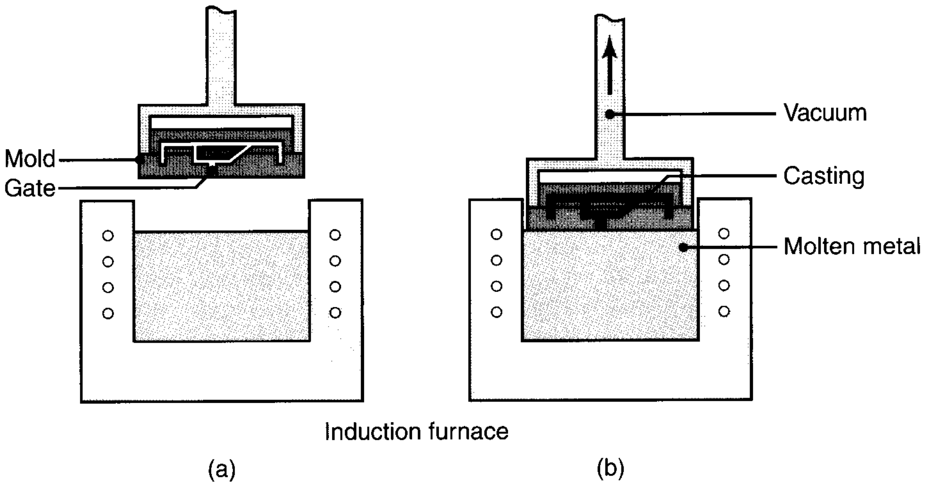


FIGURE 11.18 Schematic illustration of the vacuum-casting process. Note that the mold has a bottom gate. (a) Before and (b) after immersion of the mold into the molten metal. *Source*: After R. Blackburn.

In the vacuum-casting 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 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 usually at a temperature of 55°C above the liquidus temperature of the alloy. Consequently, it begins to solidify within a very short time. After the mold is filled, it is withdrawn from the molten metal.

The process can be automated, and production costs 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 parts 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

It was noted in Fig. 10.11 that a solidified skin develops in a casting and becomes thicker with time. Hollow castings with thin walls can be made by permanent-mold casting using this principle: a process called *slush casting*. This process is suitable for small production runs and generally is 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.

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 then are opened and the casting is removed. Note that this operation is similar to making hollow chocolate shapes, eggs, and other confectionaries.

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 upward by gas pressure into a graphite or metal mold. The pressure is maintained until the metal has solidified completely in the mold. The molten metal also may be forced upward by a vacuum, which also removes dissolved gases and produces a casting with lower porosity. Pressure casting generally is used for high-quality castings, such as steel railroad-car wheels, although 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, business-machine and appliance components, hand-tool components, and toys. The weight of most castings 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, because the process is semi- or fully automated. Die casting is economical for large production runs. The capabilities of die casting are given in Table 11.2.

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-chamber and cold-chamber machines.

The **hot-chamber process** (Fig. 11.19) involves the use of a piston, which forces a certain 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 usually are cooled by circulating water or oil through various passageways in the die block. Low-melting-point alloys (such as zinc, magnesium, tin, and lead) commonly are 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. 11.20), 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.

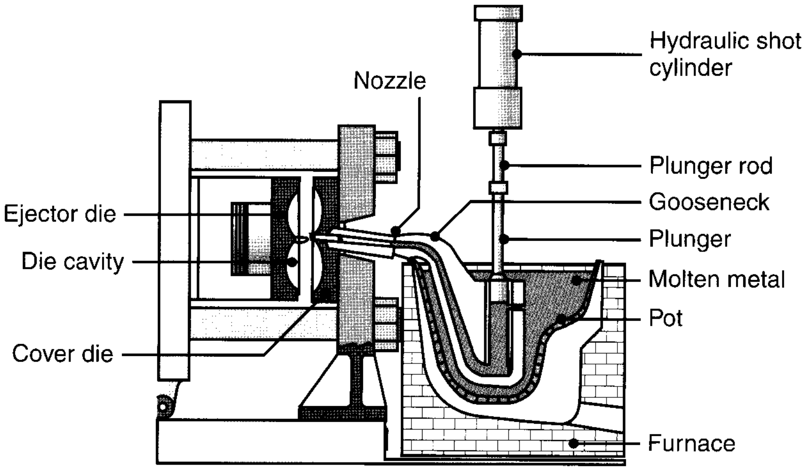


FIGURE 11.19 Schematic illustration of the hot-chamber die-casting process.

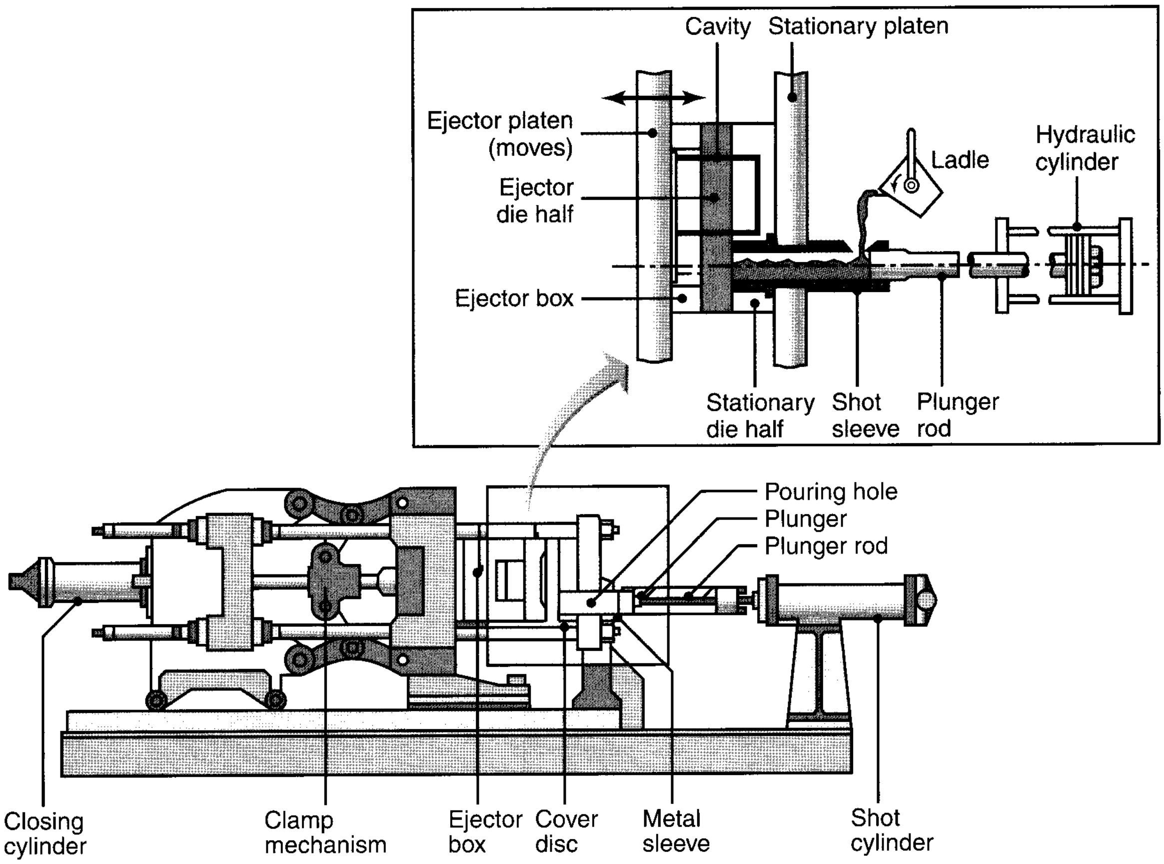


FIGURE 11.20 Schematic illustration of the cold-chamber die-casting process. These machines are large compared to the size of the casting, because high forces are required to keep the two halves of the dies closed under pressure.

The machines may be horizontal (as 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 other metals (including ferrous 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.

Process Capabilities and Machine Selection. Die casting has the capability for rapid production of strong, 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 machining or finishing operations (net-shape forming). 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, as may small amounts of flash (thin material squeezed out between the dies) at the die parting line.

A typical part made by die casting is shown in Fig. 11.1d; note the intricate shape and fine surface detail. In the fabrication of 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

TABLE 11.3

Properties and Typical Applications of Some Common Die-casting Alloys

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

Source: American Die Casting Institute.

wall thickness. With a good surface finish and dimensional accuracy, die casting can produce smooth surfaces for bearings that otherwise normally would be machined.

Components such as pins, shafts, and threaded fasteners can be die cast integrally. Called **insert molding**, this process is similar to placing wooden sticks in popsicles prior to freezing (see also Section 19.3). For good interfacial strength, insert surfaces may be knurled (see Fig. 23.11 on page 616), grooved, or splined. Steel, brass, and bronze inserts are used commonly in 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 part unless clamped together tightly. Die-casting machines are thus rated according to the clamping force that can be exerted to keep the dies closed. The capacities of commercially available machines range from about 23 to 2700 metric tons. Other factors involved in the selection of die-casting machines are die size, piston stroke, shot pressure, and cost.

Die-casting dies (Fig. 11.21) may be *single cavity*, *multiple cavity* (with several identical cavities), *combination cavity* (with several different cavities), or *unit dies*

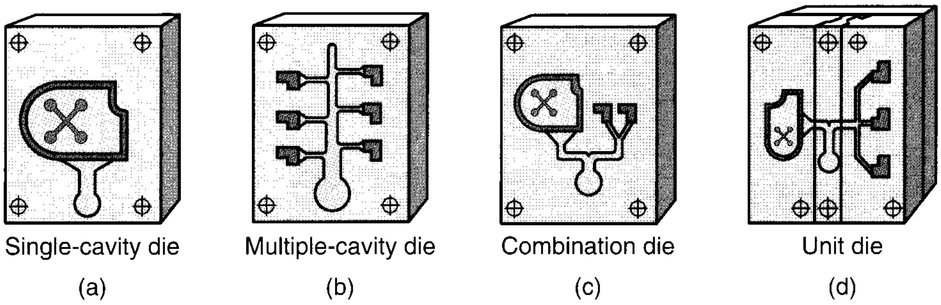

Single-cavity die (a) Multiple-cavity die (b) Combination die (c) Unit die (d)

FIGURE 11.21 Various types of cavities in a die-casting die. *Source*: Courtesy of American Die Casting Institute.

(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, the die for a casting weighing 2 kg would weigh about 2000 kg. The dies usually are made of hot-work die steels or mold steels (see Section 5.7). Die wear increases with the temperature of the molten metal. **Heat checking** of dies (surface cracking from repeated heating and cooling of the die, discussed in Section 3.6) can be a problem. When die materials are selected and maintained properly, dies may last more than a half million shots before any significant die wear takes place.

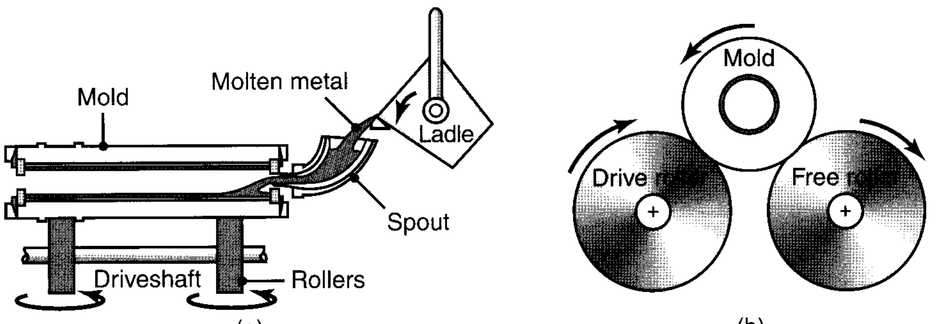
11.4.6 Centrifugal Casting

As its name implies, the *centrifugal-casting* process utilizes inertial forces (caused by rotation) to distribute the molten metal into the mold cavities—a method that was first suggested in the early 1800s. There are three types of centrifugal casting: true centrifugal casting, semicentrifugal casting, and centrifuging.

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) 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 (as much as 150 g); 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. The capabilities of centrifugal casting are given in Table 11.2.

Semicentrifugal Casting. An example of *semicentrifugal casting* is shown in Fig. 11.23a. This method is used to cast parts with rotational symmetry, such as a wheel with spokes.



(a) (b)

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.

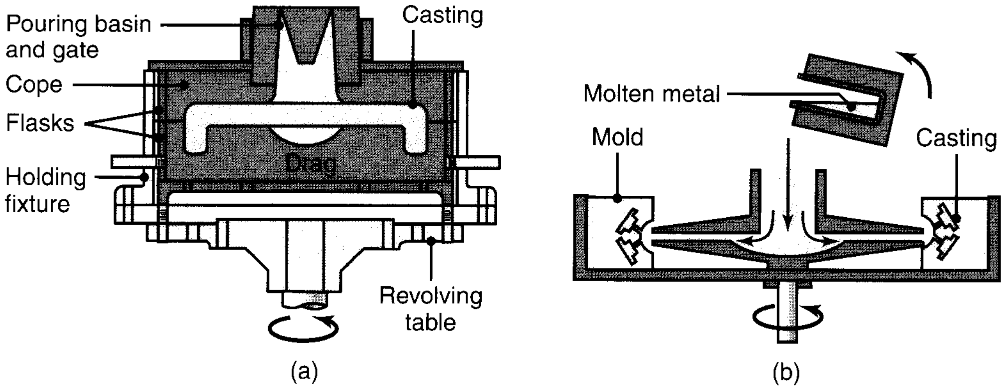


FIGURE 11.23 (a) Schematic illustration of the semicentrifugal casting process. Wheels with spokes can be cast by this process. (b) Schematic illustration of casting by centrifuging. The molds are placed at the periphery of the machine, and the molten metal is forced into the molds by centrifugal force.

Centrifuging. In *centrifuging* (also called *centrifuge casting*), mold cavities of any shape 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. 11.23b). 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 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 developed 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 short-barrelled cannon). The machinery includes a die, punch,

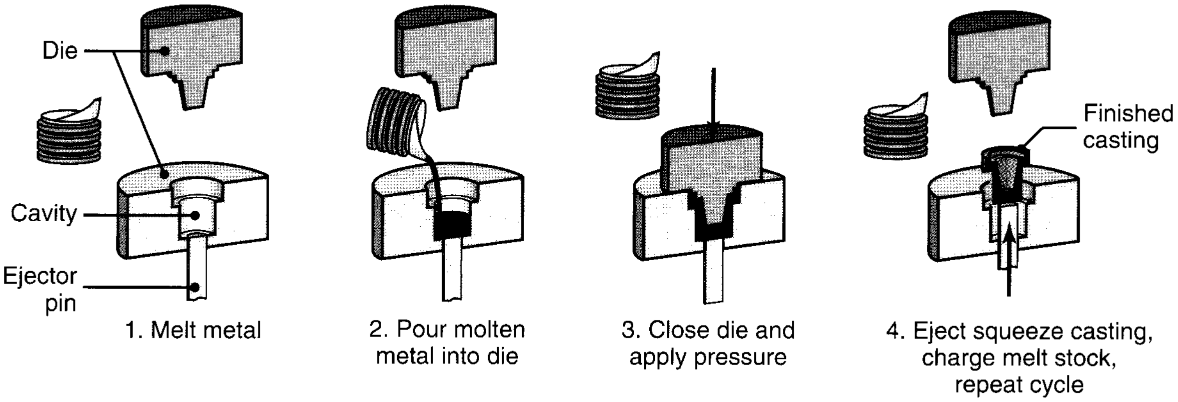


FIGURE 11.24 Sequence of operations in the squeeze-casting process. This process combines the advantages of casting and forging.

and ejector pin. The pressure applied by the punch keeps the entrapped gases in solution, and the contact under high pressure at the die–metal interface promotes rapid heat transfer, thus resulting in a fine microstructure with good mechanical properties.

The application of pressure also overcomes feeding difficulties that may arise when casting metals with a long freezing range (Section 10.2.2). The pressures required in squeeze casting are lower than those for hot or cold forging. 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 and put into commercial production by 1981. When it enters the die, the metal (consisting of liquid and solid components) is stirred so that all of the dendrites are crushed into fine solids, and when cooled in the die it develops into a fine-grained structure. The alloy exhibits *thixotropic* behavior, described in Section 10.2.3; hence, the process also is called **thixoforming** or **thixomolding**, meaning its viscosity decreases when agitated. Thus, at rest and above its solidus temperature, the alloy has the consistency of butter, but when agitated vigorously, its consistency becomes more like motor oil. Processing metals in their mushy state also has led to developments in *mushy-state extrusion*, similar to injection molding (see Section 19.3), *forging*, and *rolling* (hence the term *semisolid metalworking*). These processes also are used in making parts with specially designed casting or wrought alloys and metal-matrix composites. They also have the capability for blending granukes of different alloys, called *thixoblending*, for specific applications.

Thixotropic behavior has been utilized in developing technologies that combine casting and forging of parts using cast billets that are forged when 30 to 40% liquid. Parts made include 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 as well as wrought alloys can be used; (d) parts subsequently can 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 the 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 other casting processes. They generally are 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 castings, and can help reduce overall costs and processing time. Molding materials commonly used are shells (made as described previously), plaster, sand with binder, metal, and graphite. These molds also may include cores and chills to control the rate of solidification in critical areas of castings.

11.5 Casting Techniques for Single-crystal Components

The characteristics of single-crystal and polycrystalline structures in metals were described in Section 1.3. 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. The procedures involved also can be used for other alloys and components.

Conventional Casting of Turbine Blades. The *conventional-casting process* uses a ceramic mold. The molten metal is poured into the mold and begins to solidify at the ceramic 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. 11.25a) was first developed in 1960. The ceramic mold is preheated by radiant heating, and the mold is supported by a water-cooled chill plate. 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 thus is solidified directionally, with longitudinal, but no transverse, grain boundaries. Consequently, the blade is stronger in the direction of centrifugal forces developed in the gas turbine.

Single-crystal Blades. In *crystal growing*, developed in 1967, the mold has a constriction in the shape of a corkscrew or helix (Figs. 11.25b and c). The cross section is so small that it allows only one crystal to fit through. The mechanism of crystal growth is such that only the most favorably oriented crystals are able to grow (a situation similar to that shown in Fig. 10.3) through the helix, because all others are intercepted by the walls of the helical passage.

As the assembly is lowered slowly, a single crystal grows upward through the constriction and begins to grow in the mold. Strict control of the rate of movement is important. The resultant casting is a single-crystal blade. Although these blades

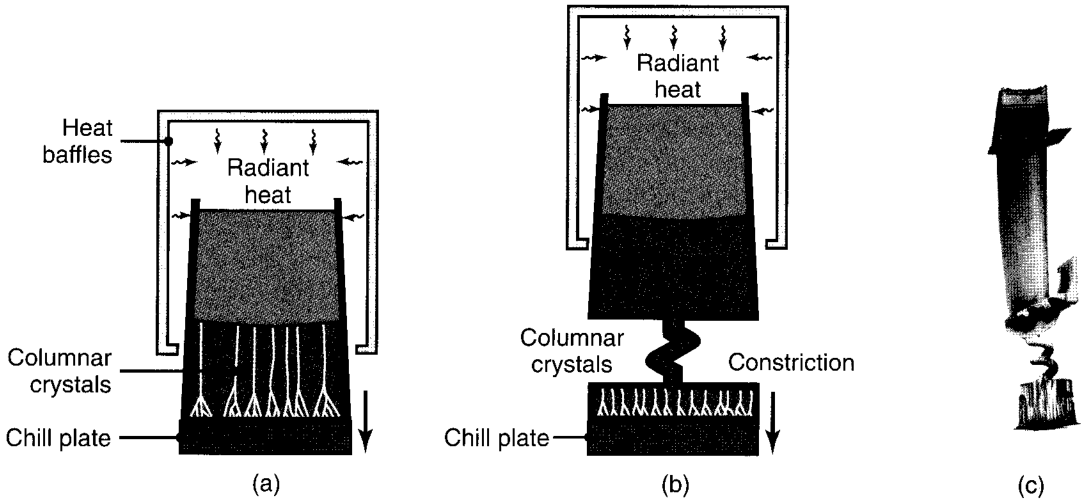


FIGURE 11.25 Methods of casting turbine blades: (a) directional solidification; (b) method to produce a single-crystal blade; and (c) a single-crystal blade with the constriction portion still attached. *Source*: (a) and (b) After B.H. Kear, (c) Courtesy of ASM International.

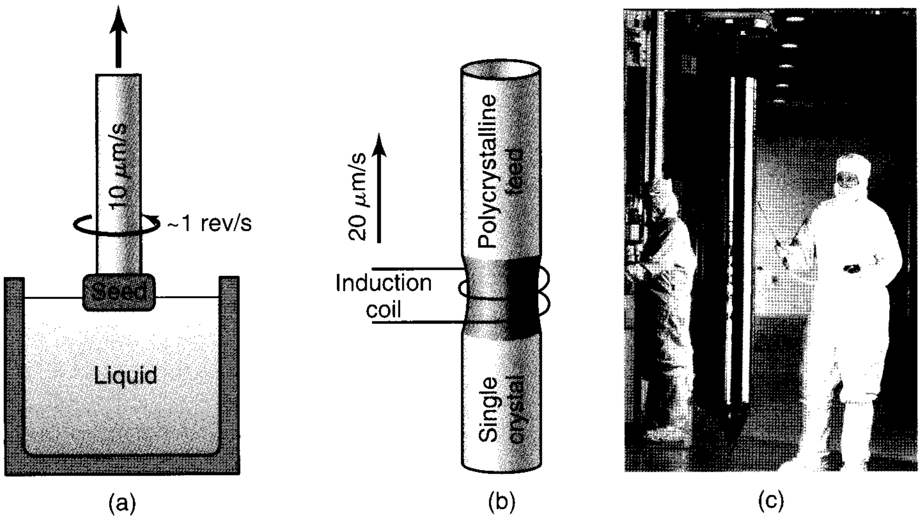


FIGURE 11.26 Two methods of crystal growing: (a) crystal pulling (Czochralski process) and (b) the floating-zone method. Crystal growing is especially important in the semiconductor industry. (c) A single-crystal ingot produced by the Czochralski process. *Source*: Courtesy of Intel Corp.

are more expensive than other types, the lack of grain boundaries makes them resistant to creep and thermal shock, so 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 in 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. 11.26a), 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 is continued throughout. *Dopants* (alloying elements) may be added to the liquid metal to impart special electrical properties. Single crystals of silicon, germanium, and various other elements are grown with 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 second technique for crystal growing is the **floating-zone method** (Fig. 11.26b). Starting with a rod of polycrystalline silicon resting on a single crystal, an induction coil 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, 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.

11.6 Rapid Solidification

The properties of *amorphous alloys* (also known as *metallic glasses*) were described in Section 6.14. 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

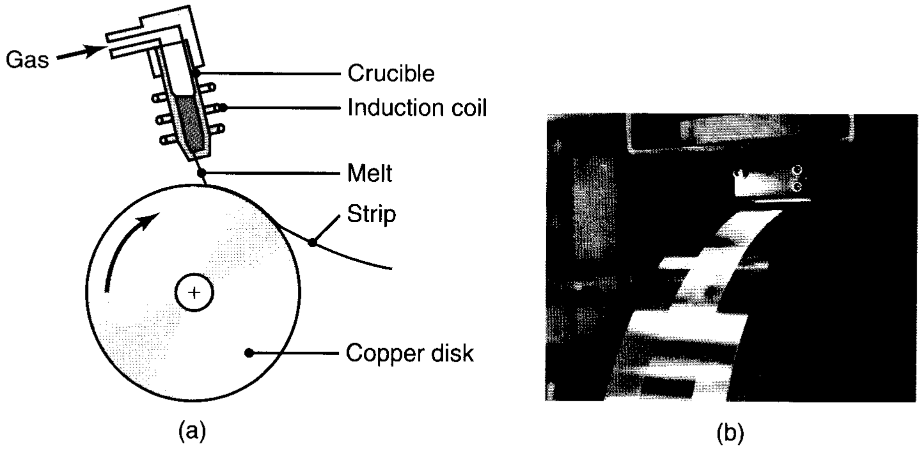


FIGURE 11.27 (a) Schematic illustration of melt spinning to produce thin strips of amorphous metal. (b) Photograph of nickel-alloy production through melt spinning. *Source*: Courtesy of Siemens AG.

sufficient time to crystallize (see also Fig. 1.10). Rapid solidification results in a significant extension of solid solubility, grain refinement, and reduced microsegregation (see Section 10.2.3), among other effects.

In a common method called **melt spinning** (Fig. 11.27), 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 chills the alloy rapidly (**splat cooling**).

11.7 Inspection of Castings

The control of all casting stages—from mold preparation to the removal of castings from molds or dies—is essential to maintaining good quality. Several methods can be used to inspect castings to determine their quality and the presence and types of any possible defects. Castings can be inspected *visually*, or *optically*, for surface defects. Subsurface and internal defects are investigated using various *nondestructive* techniques (Section 36.10). In *destructive* testing (Section 36.11), specimens are removed from various sections of a casting and tested for strength, ductility, and other mechanical properties and to determine the presence, location, and distribution of porosity and any other defects.

Pressure tightness of cast components (valves, pumps, and pipes) usually is determined by sealing the openings in the casting and pressurizing it with water, oil, or air. (Because air is compressible, its use is very dangerous in such tests because of the possibility of a sudden explosion due to a major flaw in the casting.) For extreme leaktightness 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.

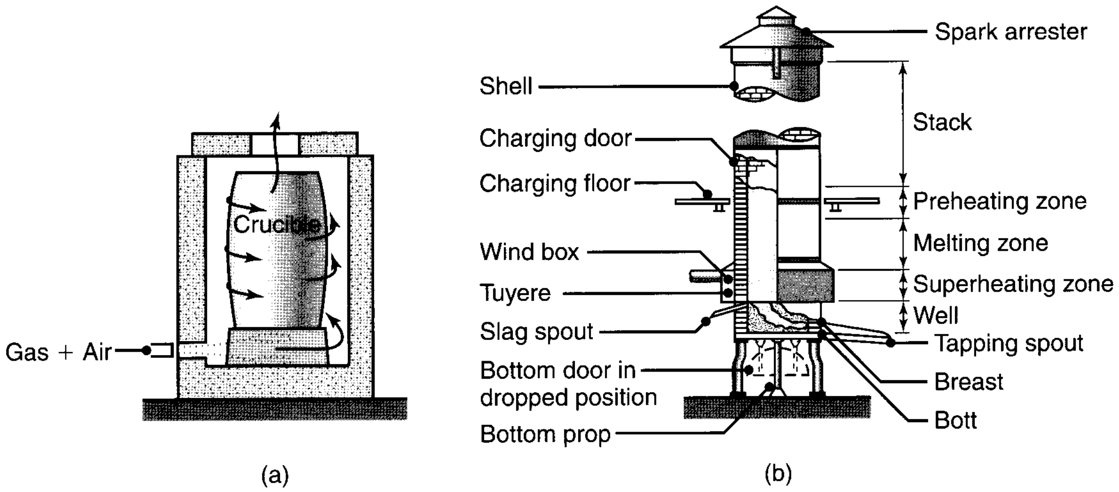
11.8 Melting Practice and Furnaces

The 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.

- **Electric-arc** furnaces, described in Section 5.2.3 and illustrated in Fig. 5.2, are used extensively in foundries and have such advantages as a high rate of melting (and thus high-production rate), much less pollution than other types of furnaces, and the ability to hold the molten metal (keep it at a constant temperature for a period of time) for alloying purposes.
- **Induction** furnaces (Fig. 5.2c) are especially useful in smaller foundries and produce smaller composition-controlled melts. There are two basic types. The *coreless induction furnace* consists of a crucible completely surrounded with a water-cooled copper coil through which a high-frequency current passes. Because there is a strong electromagnetic stirring action during induction heating, this type of furnace has excellent mixing characteristics for alloying and adding a new charge of metal.

  The other type of induction furnace, called a *core* or *channel furnace*, uses a low-frequency current (as low as 60 Hz) and has a coil that surrounds only a small portion of the unit. These furnaces commonly are used in nonferrous foundries and are particularly suitable for (a) superheating (that is, heating above normal casting temperature to improve fluidity), (b) holding (which makes it suitable for die-casting applications), and (c) duplexing (using two furnaces—for instance, melt the metal in one furnace and transfer it to another).
- **Crucible** furnaces (Fig. 11.28a), which have been used extensively throughout history, are heated with various fuels, such as commercial gases, fuel oil, and fossil fuel, as well as electricity. Crucible furnaces may be stationary, tilting, or movable.
- **Cupolas** are basically vertical, refractory-lined steel vessels charged with alternating layers of metal, coke, and flux (Fig. 11.28b). Although they require major investments and increasingly are being replaced by induction furnaces, cupolas operate continuously, have high melting rates, and produce large amounts of molten metal.



(a) (b)

FIGURE 11.28 Two types of melting furnaces used in foundries: (a) crucible and (b) cupola.

- **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 be a source of contamination with oxide inclusions). The molten metal flows downward into an investment-casting mold placed directly below the coil. Investment castings made with 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 usually 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 automated and computer-integrated facilities for all aspects of their operations. They produce a wide variety and sizes of castings at high production rates, with good quality control and at low cost.

As outlined in Fig. 11.2, foundry operations initially involve two separate groups of activities. The first group is pattern and mold making. Computer-aided design and manufacturing (Chapter 38) and rapid-prototyping techniques (Chapter 20) are now used to minimize trial and error and thus improve efficiency. 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 is melting the metals, controlling their composition and impurities, and pouring them into molds.

The rest of the operations, such as pouring into molds 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 now used extensively in foundry operations, such as cleaning, riser cutting, mold venting, mold spraying, pouring, sorting, and inspection. Automatic storage and retrieval systems for cores and patterns using automated guided vehicles (Section 37.5) are other operations.

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 a large number of castings. 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 fundamentally have different effects on casting 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 detail can be achieved.
- Other 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 integrated-circuit manufacture).
- Melting processes and their control also are important factors in casting operations. They include proper melting of the metals; preparation for alloying; removal of slag and dross; and pouring the molten metal into the molds. Inspection of castings for possible internal or external defects also is important.
- Castings are generally subjected to additional processing (such as heat treatment and machining operations) to produce the final desired shapes and surface characteristics and to achieve 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
Fluxes
Foundry
Green molding sand
Insert molding
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 prototyping
Rapid solidification
Rheocasting
Sand casting
Semisolid-metal forming
Shell-mold casting
Slush casting
Sodium-silicate process
Squeeze casting
Thixotropy
Vacuum casting

BIBLIOGRAPHY

Allsop, D.F., and Kennedy, D., **Pressure Die Casting—Part II: The Technology of the Casting and the Die**, Pergamon, 1983.

Ammen, C.W., **Metalcasting**, McGraw-Hill, 2000.

An Introduction to Die Casting, American Die Casting Institute, 1981.

Andresen, W., **Die Cast Engineering: A Hydraulic, Thermal and Mechanical Process**, CRC Press, 2005.

ASM Handbook, Vol. 15: **Casting**, ASM International, 2008.

Beeley, P., **Foundry Technology**, Butterworth-Heinemann, 2002.

Brooks, N., **Mouldmaking and Casting**, Crowood Press, 2005.

Campbell, J., **Castings**, 2nd ed., Butterworth-Heinemann, 2003.

Chastain, S., **Metal Casting: A Sand Casting Manual for the Small Foundry** (2 vols), Chastain Publishing, 2004.

Clegg, A.J., **Precision Casting Processes**, Pergamon, 1991.

Investment Casting Handbook, Investment Casting Institute, 1997.

Martin, A., **The Essential Guide to Mold Making & Slip Casting**, Lark Books, 2007.

Otooni, M.A. (ed.), **Elements of Rapid Solidification: Fundamentals and Applications**, Springer, 1999.

Product Design for Die Casting, Diecasting Development Council, 1988.

Sias, F.R., **Lost-Wax Casting**, Woodsmere Press, 2006.

Steel Castings Handbook, 6th ed., ASM International, 1995.

The Metallurgy of Die Castings, Society of Die Casting Engineers, 1986.

Viarcik, 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? A vent?

11.13. What is squeeze casting? What are its advantages?

11.14. What are the advantages of the lost-foam casting process?

QUALITATIVE PROBLEMS

11.15. What are the reasons for the large variety of casting processes that have been developed over the years? Explain with specific examples.

11.16. Why are risers not as useful in die casting as they are in sand casting?

11.17. Describe the drawbacks to having a riser that is (a) too large and (b) too small.

11.18. Why can blind risers be smaller than open-top risers?

11.19. Why does die casting produce the smallest cast parts?

11.20. Why is the investment-casting process capable of producing fine surface detail on castings?

11.21. What differences, if any, would you expect in the properties of castings made by permanent-mold versus sand-casting processes?

11.22. Would you recommend preheating the molds used in permanent-mold casting? Would you remove the casting soon after it has solidified? Explain your reasons.

11.23. Give reasons for, and examples of, using die inserts.

11.24. 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.

11.25. Explain why squeeze casting produces parts with better mechanical properties, dimensional accuracy, and surface finish than do expendable-mold processes.

11.26. How are the individual wax patterns attached on a "tree" in investment casting?

11.27. Describe the measures that you would take to reduce core shifting in sand casting.

11.28. 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?

11.29. How are hollow parts with various cavities made by die casting? Are cores used? If so, how? Explain.

11.30. It was stated that the strength-to-weight ratio of die-cast parts increases with decreasing wall thickness. Explain why.

11.31. How are risers and sprues placed in sand molds? Explain, with appropriate sketches.

11.32. 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.

11.33. Why does the die-casting machine shown in Fig. 11.20 have such a large mechanism to close the dies? Explain.

11.34. Chocolate forms are available in hollow shapes. What process should be used to make these chocolates?

11.35. What are the benefits to heating the mold in investment casting before pouring in the molten metal? Are there any drawbacks? Explain.

11.36. 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.

11.37. Can a chaplet also act as a chill? Explain.

11.38. 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

11.39. Estimate the clamping force for a die-casting machine in which the casting is rectangular with projected dimensions of 100 mm × 175 mm. Would your answer depend on whether it is a hot-chamber or cold-chamber process? Explain.

11.40. The blank for the spool shown in Fig. P11.40 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.

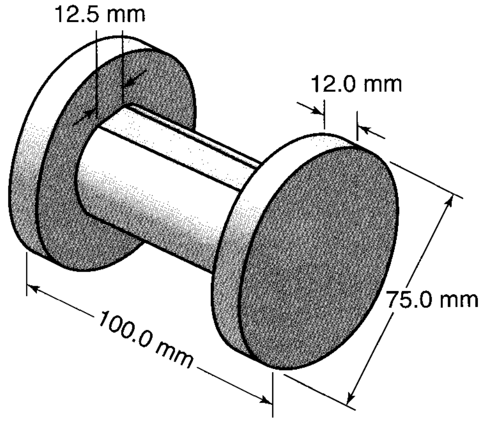


FIGURE P11.40

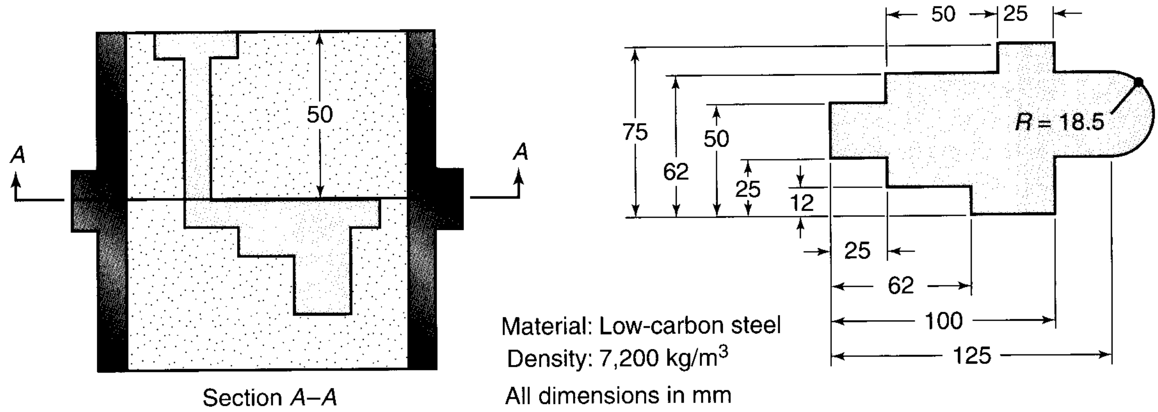
11.41. Repeat Problem 11.40, but assume that the aluminum spool is to be cast by expendable-pattern casting. Explain the important differences between the two patterns.

11.42. 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 in Fig. P11.42, 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.)

▶ **11.43.** If an acceleration of 120 g is necessary to produce a part in true centrifugal casting and the part has an inner diameter of 200 mm, a mean outer diameter of 350 mm, and a length of 6 m, what rotational speed is needed?

▶ **11.44.** A jeweler wishes to produce 24 gold rings in one investment-casting operation, as illustrated in Fig. II.1b. The wax parts are attached to a wax central sprue 12 mm in diameter. The rings are located in four rows, each 12 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. The specific gravity of gold is 19.3.

11.45. 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.



Material: Low-carbon steel
Density: 7,200 kg/m^3
All dimensions in mm

FIGURE P11.42

SYNTHESIS, DESIGN, AND PROJECTS

11.46. Describe the procedures that would be involved in making a large outdoor bronze statue. Which casting process(es) would be suitable? Why?

▶ **11.47.** 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.)

11.48. Sketch and describe a casting line consisting of machinery, conveyors, robots, sensors, etc., that automatically could perform the expendable-pattern casting process.

11.49. Outline the casting processes that would be most suitable for making small toys. Explain your choices.

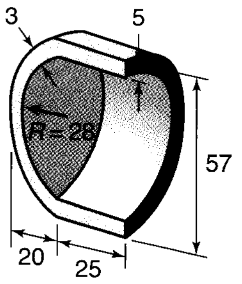
11.50. Make a list of the mold and die materials used in the casting processes described in this chapter. Under each type of material, list the casting processes that are employed and explain why these processes are suitable for that particular mold or die material.

11.51. Write a brief report on the permeability of molds and the techniques that are used to determine permeability.

11.52. Light metals commonly are cast in vulcanized rubber molds. Conduct a literature search and describe the mechanics of this process.

11.53. 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.

11.54. The part shown in Fig. P11.54 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.



Dimensions in mm

FIGURE P11.54

11.55. Porosity that has developed in the boss of a casting is illustrated in Fig. P11.55. Show that the porosity can be eliminated simply by repositioning the parting line of this casting.

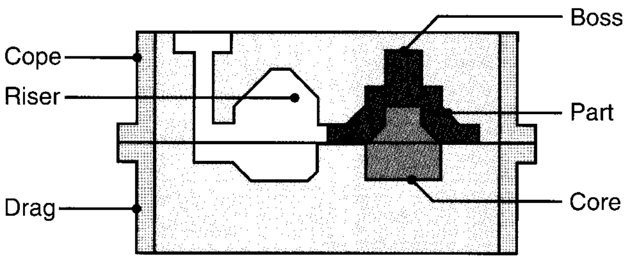


FIGURE P11.55

CHAPTER

12

Metal Casting: Design, Materials, and Economics

- This final chapter on metal casting serves as a general guide to important considerations regarding the interrelationships among product design, material, and process selection, as well as economical methods of casting for specific applications.
- The chapter describes in detail the design considerations for expendable-mold and permanent-mold casting, and discusses the general guidelines for successful casting operations.
- The characteristics and applications of the most common ferrous and nonferrous alloys are then described.
- The chapter ends with a discussion 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 various process parameters. 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 and guidelines for metal casting and presents suggestions for avoiding defects. It also describes the characteristics of the alloys that are commonly cast, together with their typical applications. Because the economics of casting operations are just as important as the technical aspects, this chapter also briefly outlines the basic economic factors relevant to casting operations.

12.2 Design Considerations in Casting

As in all manufacturing operations, certain guidelines and design principles pertaining to casting have been developed over many years. Although these principles have been established primarily through experience, analytical methods, process simulation and modeling, and computer-aided design and manufacturing techniques have all come into wide use as well, thus improving productivity and the quality of castings and resulting in significant cost savings.

All casting operations share some characteristics, such as phase change and thermal shrinkage during the casting cycle. Consequently, a number of design considerations apply equally to, for example, sand casting and die casting. However, each process will have its own particular design considerations: Sand casting will require the consideration of mold erosion and associated sand inclusions in the casting, whereas die casting will not have this concern (although it has its own problems, such as heat checking of dies which reduces die life).

Troubleshooting the causes of defects is often complicated, and the considerations presented in this chapter are by no means an exhaustive list. Also, defects frequently are random and difficult to reproduce, further complicating the implementation of corrective measures. In most cases, a given mold design will produce mostly good parts and some defective ones. It is very difficult for a mold to produce *no* defective parts. For these reasons, quality control procedures must be implemented for critical applications of castings (see Chapter 36).

12.2.1 General Design Considerations for Castings

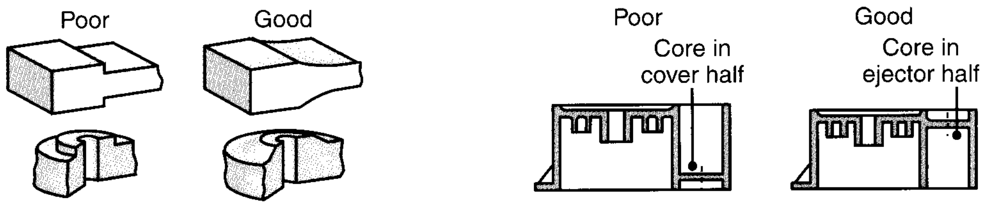
There are two types of design issues in casting: (a) geometric features, tolerances, etc., that should be incorporated into the part and (b) mold features that are needed to produce the desired casting. Robust design of castings usually involves the following steps:

1. Design the part so that the shape is cast easily. A number of important design considerations are given in this chapter to assist in such efforts.
2. Select a casting process and a material suitable for the part, size, required production volume, mechanical properties, and so on. Often, steps 1 and 2 in this list have to be specified simultaneously and can be a demanding design challenge.
3. Locate the parting line of the mold in the part.
4. Locate and design 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 sprue, screens, and risers, as appropriate.
7. Make sure proper controls and good practices are in place.

We will now examine these general rules regarding casting conditions and then discuss specific rules applicable to particular casting operations.

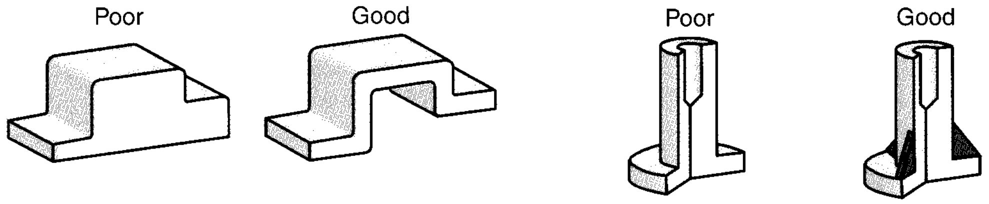
Design of Cast Parts. The following considerations are important in designing castings, as outlined in Fig. 12.1:

- **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 well as of the dies) during solidification. Fillet radii should be selected to reduce stress concentrations and to ensure proper liquid-metal flow during pouring. Fillet radii usually range from 3 to 25 mm, although smaller radii may be permissible in small castings and in specific applications. However, if the fillet radii are too large, the volume of the material in those regions also is large, and consequently, the rate of cooling is lower.



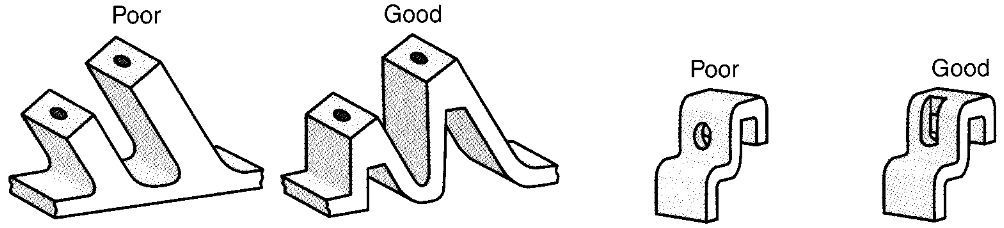
Use radii or fillets to avoid corners and provide uniform cross-section.

Deep cavities should be on one side of the casting where possible.



Wall sections should be uniform.

Ribs and/or fillets improve bosses.



Sloping bosses can be designed for straight die parting to simplify die design.

Side cores can be eliminated with this hole design.

FIGURE 12.1 Suggested design modifications to avoid defects in castings. *Source*: Courtesy of the American Die Casting Institute.

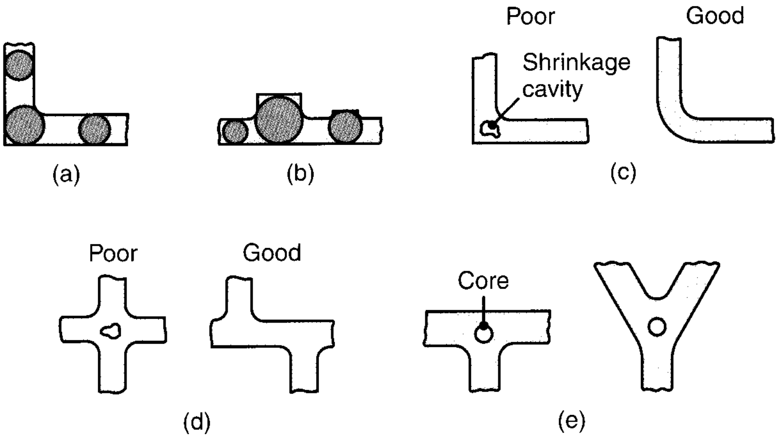


FIGURE 12.2 Examples of designs showing the importance of maintaining uniform cross sections in castings to avoid hot spots and shrinkage cavities.

Section changes in castings should be blended smoothly into each other. The location of the largest circle that can be inscribed in a particular region is critical so far as shrinkage cavities are concerned (Figs. 12.2a and b). Because the cooling rate in regions with larger circles is lower, these regions are called **hot spots**. They can develop **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 affect its strength significantly. It is important to maintain (as much as possible) uniform cross sections and wall thicknesses throughout the casting to avoid or minimize shrinkage cavities. Although they increase the cost of production, *metal paddings* or *chills* in the mold can eliminate or minimize hot spots (see Fig. 10.14).

- **Flat areas.** Large flat areas (plane surfaces) should be avoided, since they may warp during cooling because of temperature gradients, or they develop poor sur-

TABLE 12.1

Normal Shrinkage Allowance for Some Metals Cast in Sand Molds

Metal	Shrinkage allowance (%)
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
Yellow brass	1.3–1.6
Phosphor bronze	1.0–1.6
Aluminum bronze	2.1
High-manganese steel	2.6

face finish because of an uneven flow of metal during pouring. One of the common techniques for avoiding either of these problems is to break up flat surfaces with staggered ribs and serrations.

- **Shrinkage.** To avoid cracking of the casting during cooling, there should be allowances for shrinkage during solidification. In castings with intersecting ribs, the tensile stresses can be reduced by staggering the ribs or by changing the intersection 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 that are commonly sand cast.
- **Draft.** A small draft (taper) typically is provided in sand-mold patterns to enable 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.
- **Dimensional tolerances.** Dimensional tolerances depend on the particular casting process, size of the casting, and type of pattern used. Tolerances should be as wide as possible, within the limits of good part performance; otherwise, the cost of the casting increases. In commercial practice, tolerances are usually in the range of ±0.8 mm for small castings and increase with the size of the castings. Tolerances for large castings, for instance, may be as much as ±6 mm.
- **Lettering and markings.** It is common practice to include some form of part identification (such as lettering or corporate logos) in castings. These features can be sunk into the casting or can protrude from the surface; which one is most desirable 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 mill (Section 24.2), and it is simpler to machine letters into the pattern plate. On the other hand, in die casting, it is simpler to machine letters into the mold.
- **Finishing operations.** In designing a casting, 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 the hole on a flat surface than on a curved surface in order to prevent the drill from wandering. An even better design would incorporate a small dimple as a starting point for the drilling operation. Castings should include features that allow them to be clamped easily into machine tools if secondary machining operations are necessary.

Selecting the Casting Process. Casting processes cannot be selected separately from economic considerations (see Section 12.4). However, Table 11.1 lists some of the advantages and limitations of casting processes that have an impact on casting design.

Locating the Parting Line. A part 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. 10.15). Thus, hydrogen bubbles can form during the casting of aluminum, float upwards due to buoyancy, and cause a higher porosity in the top parts of castings. Therefore, critical surfaces should be oriented so that they face downwards.

A properly oriented part then can have the parting line specified. The parting line is the line or plane separating the upper (cope) and lower (drag) halves of molds (see Fig. 11.4). In general, the parting line should be along a flat plane rather than be contoured. Whenever possible, the parting 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* at the parting line (material squeezing out between the two halves of the mold) will not be as visible. The location of the parting line is important because it influences mold design, ease of molding, number and shape of cores required, method of support, and the gating system.

The parting line should be placed as low as possible (relative to the casting) for less dense metals (such as aluminum alloys) and located at around mid-height for denser metals (such as steels). However, the 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 typical that the runners, gates, and sprue well are 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 needed, and 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. Some of the considerations in designing gating systems are as follows:

- Multiple gates often are preferable and are necessary for large parts. Multiple gates have the benefits of allowing lower pouring temperature and reducing the temperature gradients in the casting.
- Gates should feed into thick 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 far 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 parts.
- 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, but when they are 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 accepts molten metal from the sprue and delivers it to the gates. One runner is used for simple parts, but two-runner systems can be specified for more complicated castings.

Runners 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.

Designing Other Mold Features. The main goal in designing a *sprue* (described in Section 10.3) is to achieve the required metal flow rates while preventing aspiration or excessive dross formation. Flow rates are determined such that turbulence is avoided, but the mold is filled quickly 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, then the dross will float and will not enter the mold cavity. *Filters* are used to trap large contaminants, and these also serve 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 observed widely that a given mold design can produce acceptable parts as well as defective ones and rarely will produce only good or only defective parts. To check for defective castings, 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 metal should not be interrupted, since this 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 cause residual stresses. Stress relieving (Section 4.11) thus 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 considerations, attributable mainly to the mold material, size of parts, and manufacturing method. Clearly, a casting in an expendable-mold process (such an investment casting) will cool much more slowly than it would in, say, die casting, which has important implications in the layout of molds.

Important design considerations for expendable-mold casting are the following:

Mold Layout. The features in the 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 in a uniform front across the casting, with the risers solidifying last. Traditionally, mold layout has been based on experience and on considerations of fluid flow and heat transfer. More recently, commercial computer programs have become available that assist in fluid flow and heat transfer analyses. These programs simulate mold filling and allow the rapid evaluation of mold layouts.

Riser Design. A major concern in the design of castings is the size and placement of risers. Risers are extremely useful in affecting the solidification-front progression across a casting and are an essential feature in the mold layout described previously.

Blind risers are good design features and maintain heat longer than open risers do. Risers are designed according to six basic rules:

1. The riser must not solidify before the casting. This rule usually is satisfied by avoiding the use of small risers and by using cylindrical risers with small aspect ratios (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 large enough to provide a sufficient amount of liquid metal to compensate for shrinkage in the casting.
3. Junctions between the casting and the riser should not develop a hot spot where shrinkage porosity can occur.
4. Risers must be placed so that the liquid metal can be delivered to locations where it is most needed.
5. There must be sufficient pressure to drive the liquid metal into locations in the mold where it is needed. Risers therefore are not as useful for metals with low density (such as aluminum alloys) as they are for those with a higher density (such as steel and cast iron).
6. The pressure head from the riser should suppress cavity formation and encourage complete cavity filling.

Machining Allowance. Because most expendable-mold castings require some additional finishing operations, such as machining and grinding, allowances should be made in casting design for these operations. Machining allowances, which are included in pattern dimensions, depend on the type of casting and increase with the size and section thickness of castings. 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

Typical design guidelines and examples for permanent-mold casting are discussed in Example 12.1. Special considerations are involved in designing tooling for die casting. Although designs may be modified to eliminate the draft for better dimensional accuracy, a draft angle of 0.5° or even 0.25° usually is required; otherwise, galling (localized seizure or sticking of material) may take place between the part and the dies and cause distortion.

Die-cast parts are nearly net shaped, 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 11.2), and in general, they do not require a machining allowance.

EXAMPLE 12.1 Illustrations of Poor and Good Casting Designs

Several examples of poor and good designs in permanent-mold and die casting are shown in Fig. 12.3. The significant differences in design are outlined here for each example:

a. The lower portion of the design on the left has a thin wall with no apparent function. This location of the part thus may fracture if subjected to high forces or impact. The good design eliminates this problem and also may simplify die and mold manufacturing.

b. Large flat surfaces always present difficulties in casting metals (as well as nonmetallic materials), as they tend to warp and develop uneven surfaces. A common practice to avoid this situation is to break up the surface with ribs and serrations on the reverse side of the casting.

This approach greatly reduces distortion and, furthermore, does not adversely affect the appearance and function of the flat surface.

c. This example of poor and good design is relevant not only to castings, but also to parts that are 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, being incompressible, will prevent full insertion of a part from being fully inserted into the cavity. The placement of a small radius at the corners or periphery at the bottom of the part eliminates this problem.

d. A 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 casting die for the good design is easier to manufacture.

e. Note that the poor design has sharp fillets at the base of the longitudinal grooves, which means that the die has sharp (knife-edge) protrusions. Because of their sharpness, it is possible that, with overextended use of the die, these edges will chip off.

f. The poor design on the left has threads reaching the right face of the casting. It then is possible that, during casting, some molten metal will penetrate this region, thus forming a flash and interfering with the function of the threaded insert, such as when a nut is used. The good design uses an offset on the threaded rod, eliminating this problem. This design consideration also is applicable for the injection molding of plastics, an example of which is shown in Fig. 19.9.

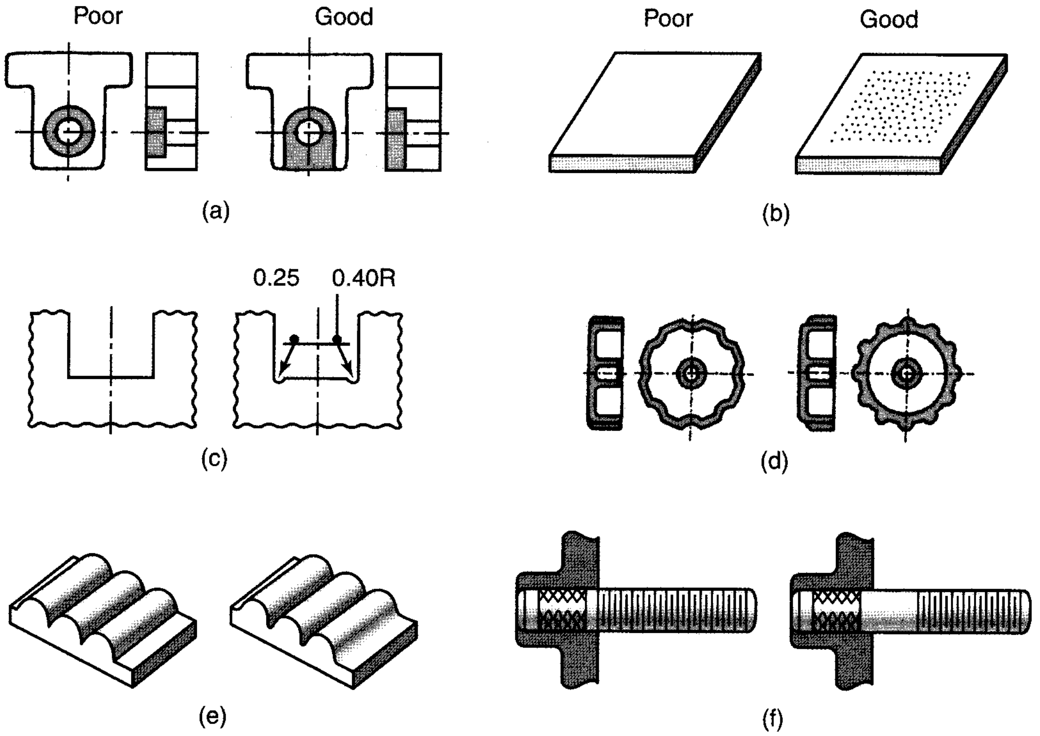


FIGURE 12.3 Examples of undesirable (poor) and desirable (good) casting designs. *Source*: Courtesy of American Die Casting Institute.

12.2.4 Computer Modeling of Casting Processes

Because casting involves complex interactions among material and process variables, a quantitative study of these interactions is essential to the proper design and production of high-quality castings. Although in the past such studies have presented major difficulties because of the large number of independent variables involved, rapid advances in computers and modeling techniques have led to important innovations in modeling various aspects of casting—including fluid flow, heat transfer, and the microstructures developed during solidification—under various casting conditions.

Modeling of *fluid flow in molds* is based on Bernoulli's and the continuity equations (Section 10.3). A model predicts the behavior of the metal during pouring into the gating system and its travel into the mold cavity, as well as the velocity and pressure distributions in the system. Progress also is being made in the modeling of *heat transfer* in casting. Modern software can couple fluid flow and heat transfer and the effects of surface conditions, thermal properties of the materials involved, and natural and forced convection on cooling. Note that the surface conditions vary during solidification, as a layer of air develops between the casting and the mold wall due to shrinkage. Similar studies are being conducted on modeling the development of *microstructures* in casting. These studies encompass heat flow, temperature gradients, nucleation and growth of crystals, formation of dendritic and equiaxed structures, impingement of grains on each other, and movement of the liquid–solid interface during solidification.

Such models now 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) gives 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, modeling techniques are becoming easier to implement. The benefits are increased productivity, improved quality, easier planning and cost estimating, and quicker response to design changes. Several commercial software programs, such as Magmasoft, ProCast, Solidia, and AFSsolid, are now available for modeling casting processes.

12.3 Casting Alloys

The general properties and 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. 12.4 and Tables 12.2 through 12.5. In addition to their casting characteristics, some other important considerations in casting alloys are their machinability and weldability, as alloys typically are assembled with other components to produce the entire part.

The most commonly used casting alloy (in tonnage) is gray iron, followed by ductile iron, aluminum, zinc, lead, copper, malleable iron, and magnesium. Shipments of castings in the United States are around 14 million metric tons per year.

12.3.1 Nonferrous Casting Alloys

Common nonferrous casting alloys are the following:

Aluminum-based Alloys. Aluminum alloys have a wide range of mechanical properties, mainly because of various hardening mechanisms and heat treatments that can

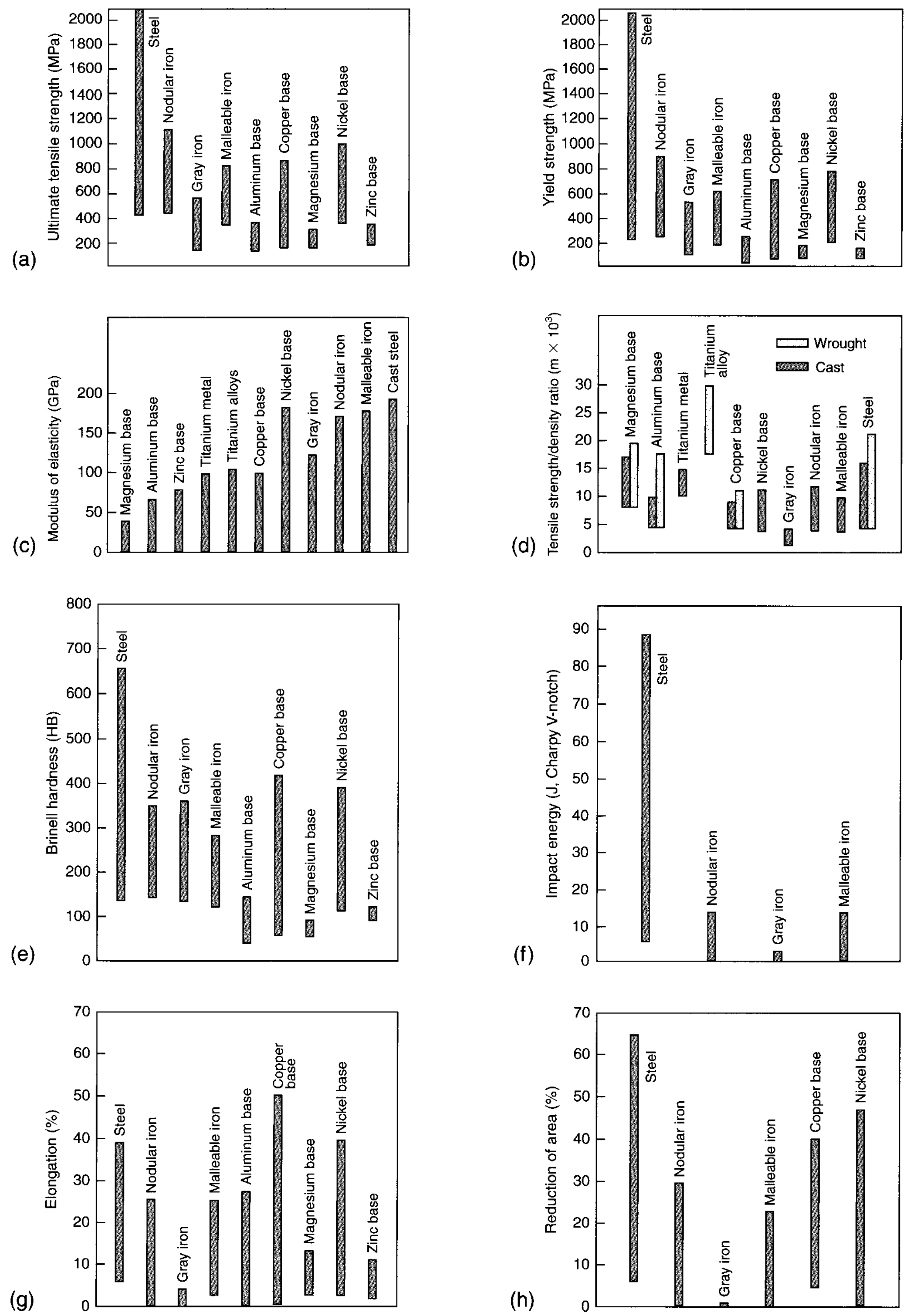


FIGURE 12.4 Mechanical properties for various groups of cast alloys. Note that even within the same group, the properties vary over a wide range, particularly for cast steels. *Source*: Courtesy of Steel Founders' Society of America.

TABLE 12.2

Typical Applications for Castings and Casting Characteristics

Type of alloy	Typical applications	Castability*	Weldability*	Machinability*
Aluminum	Pistons, clutch housings, intake manifolds	E	F	G–E
Copper	Pumps, valves, gear blanks, marine propellers	F–G	F	F–G
Ductile iron	Crankshafts, heavy-duty gears	G	D	G
Gray iron	Engine blocks, gears, brake disks and drums, machine bases	E	D	G
Magnesium	Crankcase, transmission housings	G–E	G	E
Malleable iron	Farm and construction machinery, heavy-duty bearings, railroad rolling stock	G	D	G
Nickel	Gas turbine blades, pump and valve components for chemical plants	F	F	F
Steel (carbon and low-alloy)	Die blocks, heavy-duty gear blanks, aircraft undercarriage members, railroad wheels	F	E	F
Steel (high-alloy)	Gas-turbine housings, pump and valve components, rock-crusher jaws	F	E	F
White iron	Mill liners, shot-blasting nozzles, railroad brake shoes, crushers, and pulverizers	G	VP	VP
Zinc	Door handles, radiator grills	E	D	E

*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

be used with them (Section 4.9). These alloys have high electrical conductivity and generally good atmospheric corrosion resistance. However, their resistance to some acids and all alkalis is poor, and care must be taken to prevent galvanic corrosion. They are nontoxic, lightweight, and have good machinability. Except for alloys with silicon, they generally have low resistance to wear and abrasion. Aluminum-based alloys have numerous applications, including architectural and decorative uses. An increasing trend is their use in automobiles, for components such as engine blocks, cylinder heads, intake manifolds, transmission cases, suspension components, wheels

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
Leaded tin bronze (C92500)	Annealed	260	105	35	Gears, bearings, valves
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

and brakes. Parts made of aluminum and magnesium alloys are known as **light-metal** castings.

Magnesium-based Alloys. The lowest density of all commercial casting alloys are those in the magnesium-based group. 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.

Copper-based Alloys. Although somewhat expensive, copper-based alloys have the advantages of good electrical and thermal conductivity, corrosion resistance, and nontoxicity, as well as wear properties suitable for 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 commonly are used in die casting, particularly for parts with thin walls and intricate shapes.

Tin-based Alloys. Although low in strength, these alloys have good corrosion resistance and typically are used for bearing surfaces.

Lead-based Alloys. These alloys have applications similar to tin-based alloys, but the toxicity of lead is a major drawback to their wider application.

High-temperature Alloys. High-temperature 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.

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, and they can be cast easily into intricate shapes. They generally possess several desirable properties, such as wear resistance, high hardness, 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.

a. **Gray cast iron.** Castings of gray cast iron have relatively few shrinkage cavities and low porosity. Various forms of gray cast iron are *ferritic, pearlitic*, and *martensitic*. Because of differences in their structures, each type has different properties. The mechanical properties for several classes of gray cast iron are given in Table 12.4. Typical uses of gray cast iron are in engine blocks, electric-motor housings, pipes, and wear surfaces for machines. Also, its high damping capacity has made gray iron a common material for machine-tool bases. Gray cast irons are specified by a two-digit ASTM designation. For example, class 20 specifies that the material must have a minimum tensile strength of 140 MPa.

b. **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 the material has a minimum tensile strength of 80 ksi (550 MPa), a minimum yield strength of 55 ksi (380 MPa), and 6% elongation in 2 in. (50 mm).

c. **White cast iron.** Because of its extreme hardness and wear resistance, white cast iron is used mainly for rolls for rolling mills, railroad-car brake shoes, and liners in machinery for processing abrasive materials.

d. **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 of the material is 35 ksi (240 MPa) and its elongation is 18% in 2 in. (50 mm).

e. **Compacted-graphite iron.** First produced commercially in 1976, compacted-graphite iron (CGI) has properties that are between those of gray 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 comparable to those of ductile iron. Because of its strength, parts made of CGI can be smaller and thus lighter. It is easy to cast and has consistent properties throughout the casting, and its machinability is better than that of ductile iron (which is an important consideration, since compacted-graphite iron is used for automotive engine blocks and cylinder heads). New casting techniques are being developed to further improve the machinability of CGI.

Cast Steels. Because of the high temperatures required to melt steels (up to about 1650°C, casting them requires considerable experience. 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). Cast steels can be welded; however, welding alters the cast microstructure in the heat-affected zone (see Fig. 30.17), thus influencing the strength, ductility, and toughness of the base metal. Subsequent heat treatment must be performed to restore the mechanical properties of the casting. Cast weldments have gained importance for assembling large machines and structures in which complex configurations or the size of the casting may prevent casting of the part economically in one location. 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 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 products have high heat and corrosion resistance, especially in the chemical and food industries. Nickel-based casting alloys are used for severely corrosive environments and for very high temperature service.

12.4 Economics of Casting

As is the case with all manufacturing processes, the cost of each cast part (**unit cost**) depends on several factors, including materials, equipment, and labor. Of the various casting processes discussed in Chapter 11, some require more labor than others, some require expensive dies and machinery, and some require a great deal of time to produce the castings (Table 12.6). Each of these individual factors thus affects the overall cost of a casting operation to varying degrees. As described in greater detail in Section 40.9, the cost of a product includes the costs of materials, labor, tooling, and equipment. Preparations for casting a product include the production of molds and dies that require raw materials, time, and effort—all of which also influence product cost.

As can be seen in Table 12.6, relatively little cost is involved in molds for sand casting. On the other hand, molds for various processes and die-casting dies require

TABLE 12.6

General Cost Characteristics of Casting Processes

	Cost*			Production rate
Casting process	Die	Equipment	Labor	(pieces/hr)
Sand	L	L	L–M	<20
Shell mold	L–M	M–H	L–M	<50
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.

expensive materials and a great deal of preparation. There are also major costs involved in making patterns for casting, although (as stated in Section 11.2.1) much progress is being made in utilizing rapid prototyping techniques to reduce costs and time.

Costs also are involved 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 to produce improved mechanical properties. However, heat treating also introduces 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 skills required for these operations can vary considerably, depending on the particular process and level of automation in the foundry. Investment casting, for example, requires much labor because of the many steps involved in the operation, although some automation is possible, such as in the use of robots (Fig. 11.16c), whereas operations such as a highly automated die-casting process can maintain high production rates with little labor required.

Note that the equipment cost per casting will decrease as the number of parts cast increases. Sustained high production rates, therefore, can justify the high cost of dies and machinery. However, if demand is relatively small, the cost per casting increases rapidly, and it then becomes more economical to manufacture the parts by sand casting or other casting processes described in this chapter or by other manufacturing processes described in detail in Parts III and IV.

SUMMARY

- General guidelines have been established to aid in the production of castings that are free from defects and that meet dimensional tolerances, service requirements, and various specifications and standards. These guidelines concern the shape of the casting and 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 the 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), various other considerations (such as weldability, machinability, and surface characteristics) also are important.

- Within the limits of good performance, the economic aspects of casting are 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. An important parameter is the cost per casting, which can justify large expenditures for large production runs using automated machinery and computer controls.

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.

Ammen, C.W., **Metalcasting**, McGraw-Hill, 2000.

Analysis of Casting Defects, American Foundrymen's Society, 1974.

ASM Handbook, Vol. 15: **Casting**, ASM International, 2008.

ASM Specialty Handbook: **Cast Irons**, ASM International, 1996.

Bradley, E.F., **High-Performance Castings: A Technical Guide**, Edison Welding Institute, 1989.

Bralla, J.G., **Design for Manufacturability Handbook**, McGraw-Hill, 1999.

Campbell, J., **Castings**, 2nd ed., Butterworth-Heinemann, 2003.

________, **Castings Practice: The Ten Rules of Castings**, Butterworth-Heinemann, 2004.

Cardelli, F., **Materials Handbook: A Concise Desk Reference**, 2nd ed., Springer, 2008.

Casting Design and Performance, ASM International, 2009.

Clegg, A.J., **Precision Casting Processes**, Pergamon, 1991.

Davis, J.R. (ed.), **Cast Irons**, ASM International, 1996.

Krauss, G., **Steels: Heat Treatment and Processing Principles**, ASM International, 1990.

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.

Rowley, M.T. (ed.), **International Atlas of Casting Defects**, American Foundrymen's Society, 1974.

Steel Castings Handbook, 6th ed., ASM International, 1995.

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. What are light castings and where are they used most commonly?

12.6. Name the types of cast irons generally available, and list their major characteristics and applications.

12.7. Comment on your observations regarding Fig. 12.4.

12.8. Describe the difference between a runner and a gate.

12.9. What is the difference between machining allowance and dimensional tolerance?

QUALITATIVE PROBLEMS

12.10. Describe your observation concerning the design changes shown in Fig. 12.1.

12.11. If you need only a few castings of the same design, which three processes would be the most expensive per piece cast?

12.12. Do you generally agree with the cost ratings in Table 12.6? If so, why?

12.13. Describe the nature of the design differences shown in Fig. 12.3. What general principles do you observe in this figure?

12.14. Note in Fig. 12.4 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.

12.15. Do you think that there will be fewer defects in a casting made by gravity pouring versus one made by pouring under pressure? Explain.

12.16. Explain the difference in the importance of drafts in green-sand casting versus permanent-mold casting.

12.17. What type of cast iron would be suitable for heavy-machine bases, such as presses and machine tools? Why?

12.18. Explain the advantages and limitations of sharp and rounded fillets, respectively, in casting design.

12.19. Explain why the elastic modulus, *E*, of gray cast iron varies so widely, as shown in Table 12.4.

12.20. 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.

12.21. 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.

12.22. The heavy regions of parts typically are placed in the drag in sand casting and not in the cope. Explain why.

QUANTITATIVE PROBLEMS

▶**12.23.** 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?

12.24. 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.

SYNTHESIS, DESIGN, AND PROJECTS

12.25. Describe the general design considerations pertaining to metal casting.

12.26. Add more examples to those shown in Fig. 12.2.

12.27. Explain how ribs and serrations are helpful in casting flat surfaces that otherwise may warp. Give a specific illustration.

12.28. 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.

12.29. 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.

▶**12.30.** If you need only a few units of a particular casting, which process(es) would you use? Why?

▶**12.31.** For the cast metal wheel illustrated in Fig. P12.31, 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.

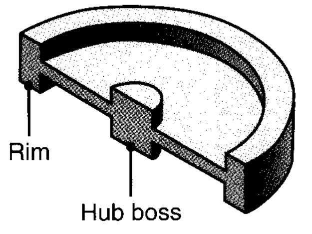


FIGURE P12.31

12.32. Assume that the introduction to this chapter is missing. Write a brief introduction to highlight the importance of the topics covered in it.

▶**12.33.** In Fig. P12.33, 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?

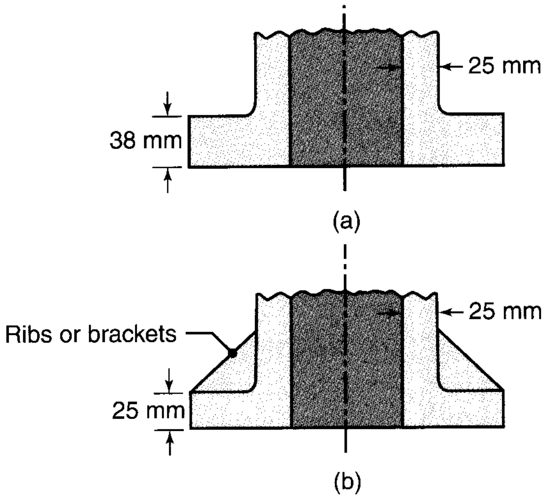


FIGURE P12.33

▶ 12.34. An incorrect and a correct design for casting are shown in Fig. P12.34. Review the changes made and comment on their advantages.

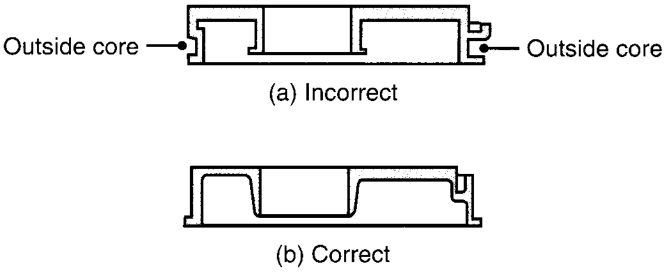


FIGURE P12.34

▶ 12.35. Three sets of designs for die casting are shown in Fig. P12.35. Note the changes made to design 1 and comment on the reasons for then or them.

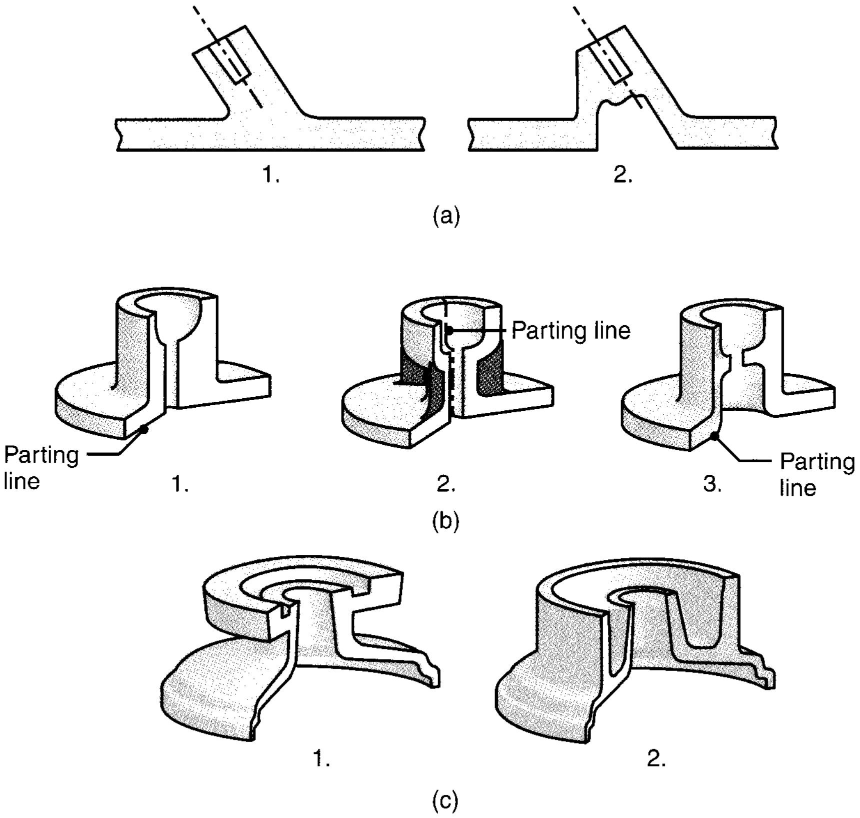


FIGURE P12.35

Forming and Shaping Processes and Equipment

PART III

We generally tend to take for granted many of the products that we use today and the materials and components from which they are made. However, when we inspect these products, we soon realize that a wide variety of materials and processes has been used in making them (Fig. III.1). Note also that some products consist of a few parts (mechanical pencils, light bulbs), while others consist of thousands of parts (automobiles, computers) or even millions of parts (airplanes, ships). Some products have simple shapes with smooth curvatures (ball bearings, bicycle handles), but others have complex configurations (engine blocks, pumps) and detailed surface features (coins, silverware). Some products are used in critical applications (elevator cables, turbine blades), whereas others are used in routine applications (paper clips, forks, knives). Some products are very thin (aluminum foil, plastic film), whereas others are very thick (ship hulls, boiler plates).

Note that the words *forming* and *shaping* are both used in the title of this part of the 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

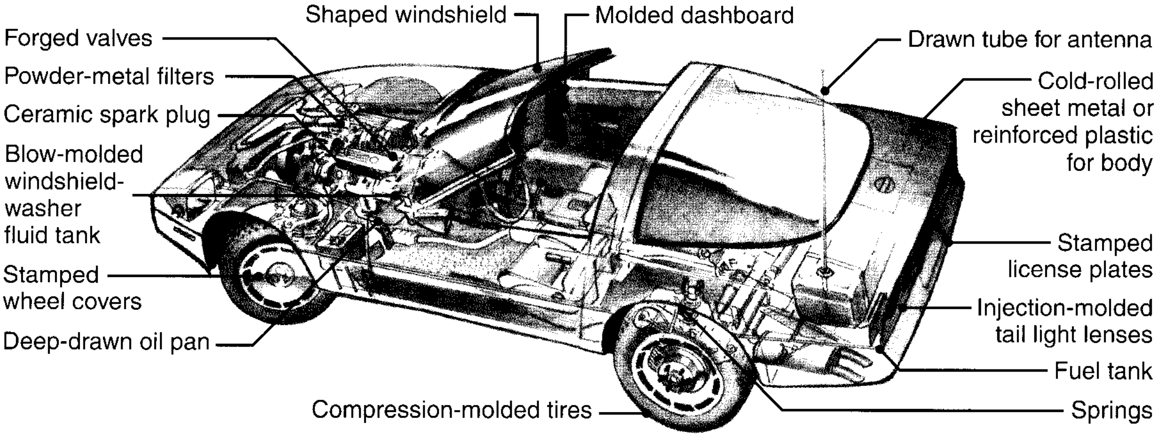


FIGURE III.1 Formed and shaped parts in a typical automobile.

forming processes, the starting material (usually called the workpiece, stock, or blank) may be in the shape of a plate, sheet, bar, rod, wire, or tubing of various cross sections. For example, an ordinary wire coat hanger is made by forming a straight piece of wire by bending and twisting it into the shape of a hanger. As another example, the metal body for an automobile typically is made of cold-rolled, flat steel sheet which is then formed into various shapes (hood, roof, trunk, door panels) using a pair of large dies.

Shaping processes typically involve the molding and casting of soft or molten materials, and the finished product is usually at or near the final desired shape. It may require little or no further finishing. 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. Telephone receivers, refrigerator-door liners, computer housings, and countless other plastic products likewise are shaped by forcing the molten polymer into a mold and letting it solidify. Some of the forming and shaping operations produce long *continuous products*, such as plates, sheets, tubing, wire, and bars with various cross sections. Rolling, extrusion, and drawing processes (Chapters 13 and 15) are capable of making such 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 compaction (Chapter 17), ceramic slip casting and glass pressing (Chapter 18), and processes involving plastics and reinforced plastics (Chapter 19) typically produce *discrete products*.

The initial 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 products. For plastics, the starting material is usually pellets, flakes, or powder, and for ceramics, it is clays and oxides obtained from ores or produced synthetically.

The important factors involved in each forming and shaping process are described in this part of the text, along with how material properties and processes affect product quality (Table III.1). We also explain why some materials can be processed only by certain manufacturing methods and why parts with particular shapes can be processed only by certain techniques and not by others. The characteristics of the machinery and the equipment used in these processes also significantly affect product quality, production rate, and the economics of a particular manufacturing operation.

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

CHAPTER

13

Metal-Rolling Processes and Equipment

- This first chapter of Part III on the forming and shaping of metallic and nonmetallic materials describes the rolling of metals, perhaps the most important metal-forming operation based on volume of metals rolled.
- The chapter begins with a description of the flat-rolling process, analyzing the force, torque, and power required in terms of relevant material and process parameters, as well as a review of defects and their causes in rolled products.
- Shape-rolling processes are then described, where workpieces are passed through a series of shaped rolls.
- Special rolling processes such as cross rolling, ring rolling, thread rolling, tube rolling, and tube piercing are also discussed.
- The chapter ends with a description of the characteristics of rolling mills and roll arrangements for specific products.

Typical products made by various rolling processes: Plates for ships, bridges, structures, 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). This process is similar to rolling dough with a rolling pin to reduce its thickness. Rolling, which accounts for about 90% of all metals 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 involve combining continuous casting with rolling processes. This 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. Typical applications are in the rolling of plastics, powder metals, ceramic slurry, and hot glass.

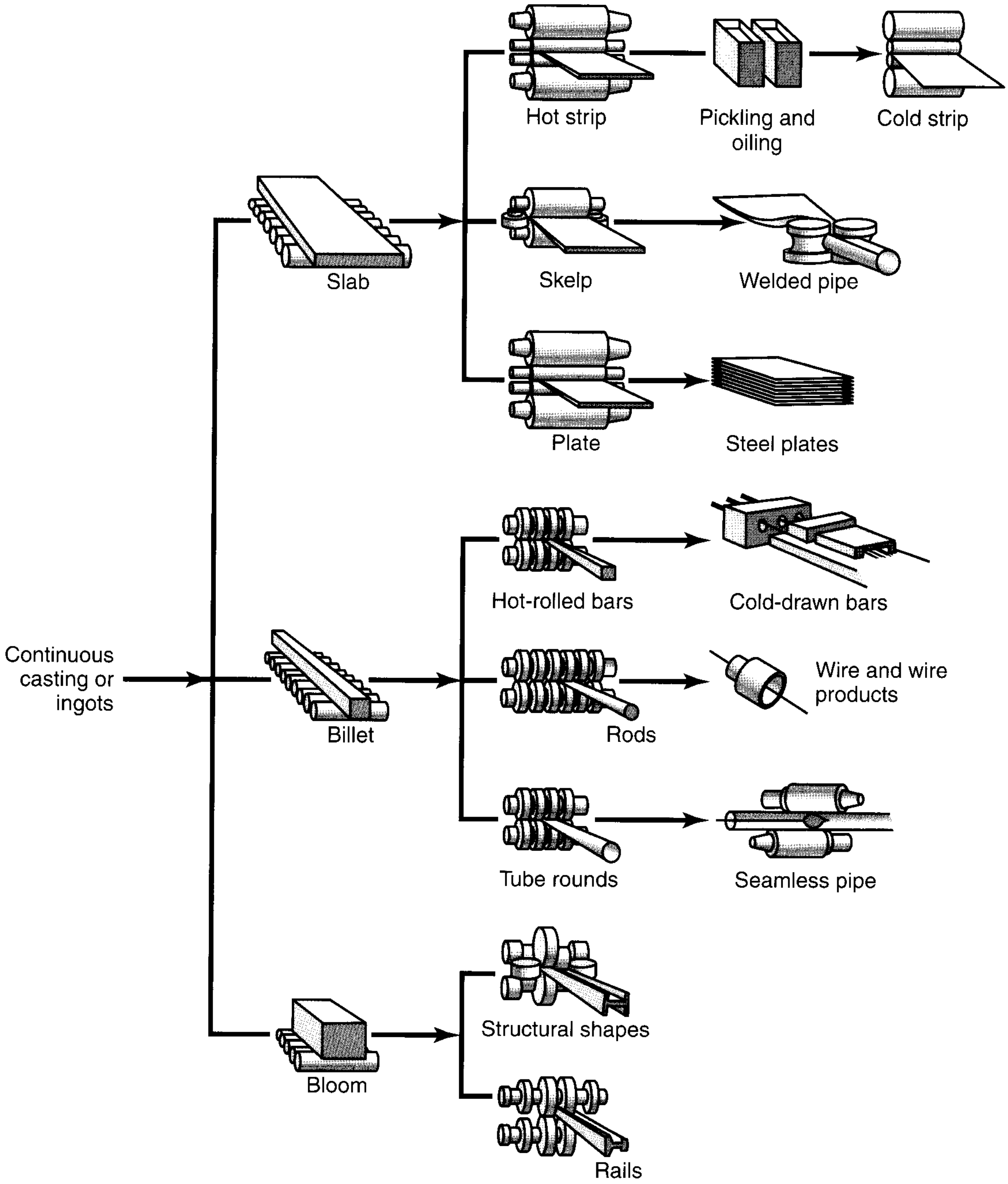


FIGURE 13.1 Schematic outline of various flat-rolling and shape-rolling processes. *Source*: After American Iron and Steel Institute.

Rolling first is carried out at elevated temperatures (hot rolling). During this phase, 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. Subsequently, rolling typically is carried out at room temperature (cold rolling), whereby the rolled product has higher strength and hardness and a better surface finish. However, it requires more energy (because of the higher strength of the material at room temperature) and will result in a product with anisotropic properties (due to preferred orientation or mechanical fibering; see Section 1.6).

Plates generally have a thickness of more than 6 mm and are used for structural applications, such as ship hulls, boilers, bridges, machinery, and nuclear vessels. 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 generally are less than 6 mm thick and typically are provided to manufacturing facilities as coils—weighing as much as 30,000 kg—or as flat sheets for further processing into various products. Sheets typically are used for automobile and aircraft bodies, appliances, food and beverage containers, and kitchen and office equipment. Commercial aircraft fuselages and trailer bodies usually are made of a minimum of 1-mm thick aluminum-alloy sheets. For example, the skin thickness of a Boeing 747 fuselage is 1.8 mm and of a Lockheed L1011 is 1.9 mm. Steel sheets used for automobile and appliance bodies are typically about 0.7 mm thick. Aluminum beverage cans are made from sheets 0.28 mm thick. After processing into a can, this sheet metal becomes a cylindrical body with a wall thickness of 0.1 mm. Aluminum **foil** (typically used for wrapping candy and chewing gum) has a thickness of 0.008 mm, although thinner foils down to 0.003 mm also can be produced with a variety of metals.

This chapter describes the fundamentals of **flat-rolling** and various **shape-rolling** operations, examines the production of seamless tubing and pipe, and discusses the important factors involved in rolling practices.

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_0 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; the velocity of the strip 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 flowing 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 arc of contact in the roll gap, L. At one point along the contact length (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 right of this point, the strip moves faster than the roll. Consequently, the frictional forces—which oppose motion between two sliding bodies—act on the strip as shown in Fig. 13.2b.

The rolls pull the material into the roll gap through a *net frictional force* on the material. Thus, the net frictional force must 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 necessary for rolling materials (just as it is in driving a car on a road), energy is dissipated in overcoming friction. Thus, increasing friction also increases rolling forces and power requirements. Furthermore, high friction could damage the surface of the rolled product (or cause sticking, as can occur in rolling dough). Thus, a compromise is made in practice: Low and controlled friction is induced in rolling through the use of effective lubricants.

The maximum possible **draft** is defined as the difference between the initial and final strip thicknesses, or $(h_o - h_f)$. It can be shown that this quantity is a function

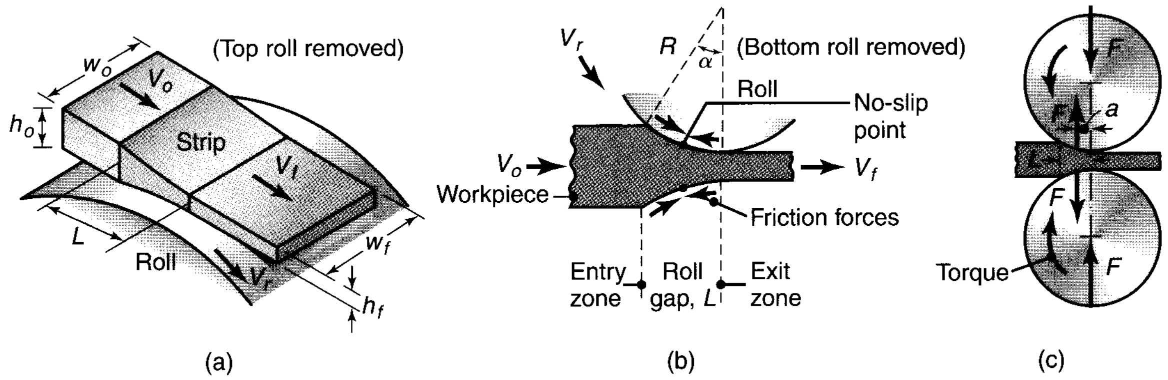


FIGURE 13.2 (a) Schematic illustration of the flat-rolling process. (b) Friction forces acting on strip surfaces. (c) Roll force, F, and torque, T, acting on the rolls. The width of the strip, w, usually increases during rolling, as shown later in Fig. 13.5.

of the roll radius, R, and the coefficient of friction, μ, between the strip and the roll by the following relationship:

$$h_o - h_f = \mu^2 R. \tag{13.1}$$

Thus, as expected, the higher the friction and the larger the roll radius, the greater the maximum possible draft becomes. Note that this situation is similar to the use of large tires (high R) and rough treads (high μ) on farm tractors and off-road earthmoving equipment, thus permitting 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 in order to reduce its thickness, resulting in a *roll force*, F, as shown in Fig. 13.2c. Note that this force appears in the figure as perpendicular to the plane of the strip, rather than at an angle. This is because, in practice, the arc of contact is very small compared with the roll radius, so we can assume that the roll force is perpendicular to the strip without causing significant error in calculations. The roll force in flat rolling can be estimated from the formula

$$F = LwY_{\text{avg}}, \tag{13.2}$$

where L is the roll-strip contact length, w is the width of the strip, and Y_{avg} is the average true stress (see Section 2.2) of the strip in the roll gap. Equation (13.2) is for a *frictionless* situation; 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. The power required per roll can be estimated by assuming that F acts in the middle of the arc of contact; thus, in Fig. 13.2c, $a = L/2$. Therefore, the *total power* (for two rolls), in S.I. units, is

$$\text{Power (in kW)} = \frac{2\pi FLN}{60{,}000} \tag{13.3}$$

where F is in newtons, L is in meters, and N is the revolutions per minute of the roll. In traditional English units, the total power can be expressed as

$$\text{Power (in hp)} = \frac{2\pi FLN}{33{,}000} \tag{13.4}$$

where F is in pounds and L is in feet.

EXAMPLE 13.1 Calculation of Roll Force and Torque in Flat-rolling

An annealed copper strip 228 mm wide and 25 mm thick is rolled to a thickness of 20 mm in one pass. The roll radius is 300 mm, and the rolls rotate at 100 rpm. Calculate the roll force and the power required in this operation.

Solution The roll force is determined from Eq. (13.2), in which L is the roll-strip contact length. It can be shown from simple geometry that this length is given approximately by

$$L = \sqrt{R(h_o - h_f)} = \sqrt{300(25 - 20)} = 38.7\ \text{mm}.$$

The average true stress, Y_{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

$$\varepsilon = \ln\left(\frac{25}{20}\right) = 0.223.$$

Referring to Fig. 2.6, note that 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 280 MPa. Thus, the average true stress is $(80 + 280)/2 = 180$ MPa. We can now define the roll force as

$$F = LwY_{\text{avg}} = \frac{38.7}{1000} \times \frac{250}{1000} \times 180\ \text{MPa}$$
$$= 1.74\ \text{MN}.$$

The total power is calculated from Eq. (13.4), with $N = 100$ rpm. Thus,

$$\text{Power} = \frac{2\pi FLN}{66{,}000} = 2\pi \times 1.74 \times 10^6$$
$$\times \frac{38.7}{1000} \times \frac{100}{66{,}000} = 705\ \text{W}.$$

Exact calculation of the force and the power requirements in rolling is 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, particularly for hot rolling because of the sensitivity of the strength of the material to 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 in turn will affect the rolling operation. Also, the columns of the **roll stand** (including the housing, chocks, and bearings, as shown in Fig. 13.3) may deflect under high roll forces to such an extent that the roll gap can open up significantly. Consequently, the rolls have to be set closer than originally calculated in order to compensate for this deflection and to obtain the desired final thickness.

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 front and/or back tensions to the strip

Among these strategies, the last requires some elaboration. An effective method of reducing roll forces is to apply longitudinal **tension** to the strip during rolling (as a result of which the compressive stresses required to plastically deform the material become smaller). Because they require high roll forces, tensions are important particularly in rolling 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 applying a braking action to the reel that supplies the sheet into the roll gap (*pay-off reel*) by some suitable means. 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—a process known as **Steckel rolling**.

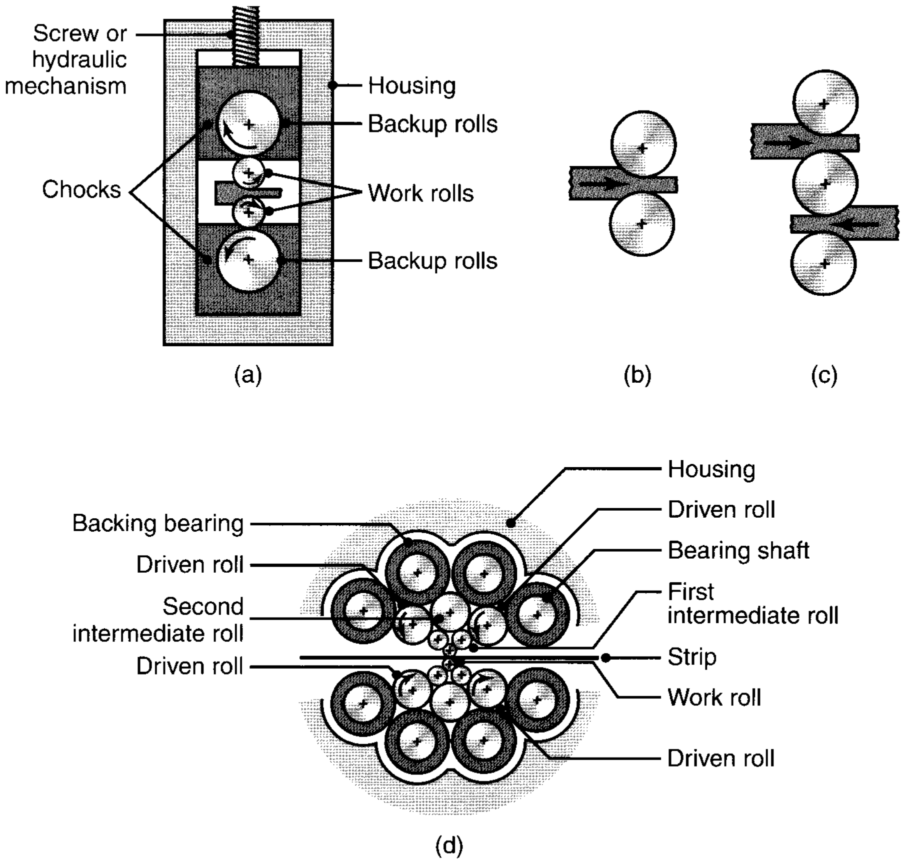


FIGURE 13.3 Schematic illustration of various roll arrangements: (a) four-high rolling mill showing various features. The stiffness of the housing, the rolls, and the roll bearings are all important in controlling and maintaining the thickness of the rolled strip; (b) two-high mill; (c) three-high mill; and (d) cluster (or *Sendzimir*) mill.

13.2.2 Geometric Considerations

Because of the forces acting on them, rolls undergo changes in shape during rolling. Just as a straight beam deflects under a transverse load, roll forces tend to bend the rolls *elastically* during rolling (Fig. 13.4a). As expected, the higher the elastic modulus of the roll material, the smaller the roll deflection.

As a result of roll bending, the rolled strip tends to be thicker at its center than at its edges (**crown**). The usual 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 (**camber**). Thus, when the roll bends, the strip being rolled now has a constant thickness along its width (Fig. 13.4b). For rolling sheet metals, the radius of the maximum camber point is generally 0.25 mm greater than that at the edges of the roll. However, as expected, a particular camber is correct only for a certain load and strip width. To reduce the effects of deflection, the rolls also can be subjected to external bending by applying moments at their bearings

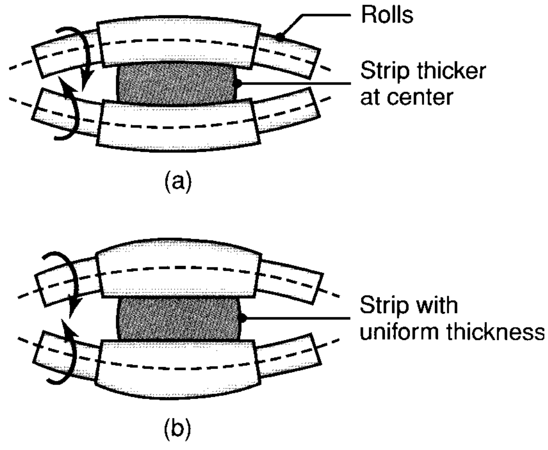


FIGURE 13.4 (a) Bending of straight cylindrical rolls caused by roll forces. (b) Bending of rolls ground with camber, producing a strip with uniform thickness through the strip width. Deflections have been exaggerated for clarity.

(a technique demonstrated simply by bending a wooden stick at its ends, a manipulation that simulates camber).

Because of the heat generated by plastic deformation during rolling, rolls can become slightly barrel shaped (**thermal camber**). Unless compensated for by some means, this condition can produce strips that are thinner at the center than at the edges. Consequently, the total (or final) camber can be controlled by adjusting the location and the flow rate of the coolant along the length of the rolls during hot rolling.

Roll forces also tend to *flatten* the rolls elastically, producing an effect much like the flattening of automobile tires under load. Flattening of the rolls is undesirable, as it results, in effect, in a larger roll radius. This, in turn, means a larger contact area for the same draft, and the roll force increases because of the now larger contact area.

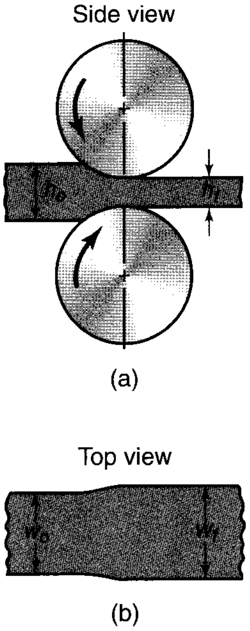


FIGURE 13.5 Spreading in flat rolling. Note that similar spreading can be observed when dough is rolled with a rolling pin.

Spreading. 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 strip with a square cross section), its width increases significantly as it passes through the rolls (an effect commonly observed in the rolling of dough with a rolling pin). This increase in width is called *spreading* (Fig. 13.5). In the calculation of the roll force, the width w in Eq. (13.2) then is taken as an average width.

It can be shown that spreading increases with (a) decreasing width-to-thickness ratio of the entering strip (because of reduction in the width constraint), (b) increasing friction, and (c) decreasing ratio of the roll radius to the strip thickness. The last two effects are due to the increased longitudinal constraint of the material flow in the roll gap. Spreading can be prevented also by using additional rolls (with vertical axes) in contact with the edges of the rolled product in the roll gap (*edger mills*), thus providing a physical constraint to spreading.

13.2.3 Vibration and Chatter

Vibration and *chatter* can have significant effects on product quality and the productivity of metalworking operations. Chatter, generally defined as *self-excited vibration*, can occur in rolling as well as in extrusion, drawing, machining, and grinding operations. In rolling, it leads to periodic variations in the thickness of the rolled sheet and in its surface finish and, consequently, can lead to excessive scrap (see Table 40.3). Chatter in rolling has been found to occur predominantly in *tandem mills*. Chatter is very detrimental to productivity; it has been estimated, for example, that modern rolling mills could operate at up to 50% higher speeds were it not for chatter.

Chatter is a very complex phenomenon (see also Section 25.4) and results from interactions between the structural dynamics of the mill stand and the dynamics of the rolling operation. Rolling speed and lubrication are found to be the two most important parameters. Although not always practical to implement, it also has been suggested that chatter can be reduced by increasing the distance between the stands of the rolling mill, increasing the strip width, decreasing the reduction per pass (draft), increasing the roll radius, increasing the strip-roll friction, and incorporating dampers in the roll supports.

13.3 Flat-rolling Practice

The initial rolling steps (*breaking down*) of the material typically is done by **hot rolling** (above the recrystallization temperature of the metal; see Section 1.7). As described in Section 10.2 and illustrated in Fig. 10.2, a **cast structure** typically is dendritic, and it

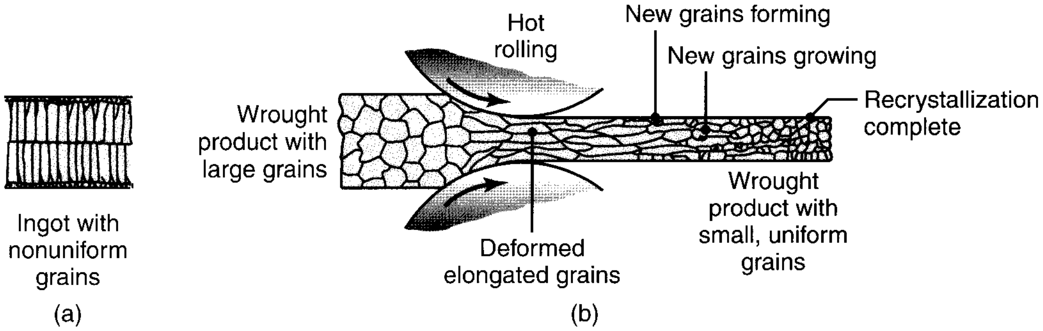


FIGURE 13.6 Changes in the grain structure of cast or of large-grain wrought metals during hot rolling. Hot rolling is an effective way to reduce grain size in metals for improved strength and ductility. Cast structures of ingots or continuous castings are converted to a wrought structure by hot working.

includes coarse and nonuniform grains; this structure usually is brittle and may 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 (especially porosity). 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 product of the first hot-rolling operation is called a **bloom**, a **slab**, or a **billet** (see Fig. 13.1). A bloom usually has a square cross section, at least 150 mm on the side; a slab usually is rectangular in cross section. Blooms are processed further 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 blooms) and later are rolled into various shapes, such as round rods and bars, using shaped rolls. Hot-rolled round rods (**wire rods**) are used as the starting material for rod- and wire-drawing operations (Chapter 15).

In the hot rolling of blooms, billets, and slabs, the surface of the material usually is **conditioned** (prepared for a subsequent operation) prior to rolling them. Conditioning is often done by means of a torch (*scarfing*) to remove heavy scale or by rough grinding to smoothen 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 to remove other defects as well.

Cold rolling is carried out at room temperature and, compared with hot rolling, 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 metal are rolled together, thus improving productivity. *Aluminum foil*, for example, is pack rolled in two layers, so only the top and bottom outer layers have been in contact with the rolls. 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, but the foil-to-roll side is shiny and bright because it has been in contact under high contact stresses 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 correct this situation, the

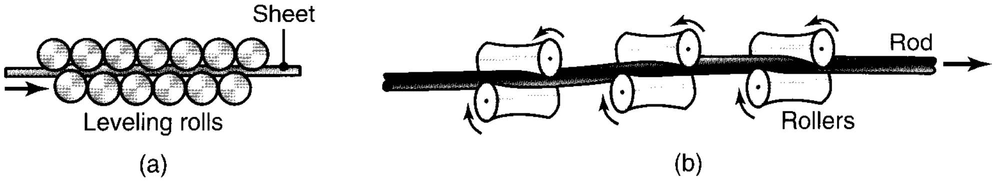


FIGURE 13.7 (a) A method of roller leveling to flatten rolled sheets. (b) Roller leveling to straighten drawn bars.

sheet metal is subjected to a final, light pass of 0.5 to 1.5% reduction known as **temper rolling** or **skin pass** shortly before stretching.

A rolled sheet may not be sufficiently flat as it leaves 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 goes through a series of **leveling rolls**. Several roller arrangements are used, as shown in Fig. 13.7. The workpiece is flexed in opposite directions as it passes through the sets of rollers. Each roll usually is driven separately by an individual electric motor.

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 compromise surface appearance, but also because they may adversely affect strength, formability, and other manufacturing characteristics. Several surface defects (such as scale, rust, scratches, gouges, pits, and cracks) have been identified in sheet metals. These defects may be caused by inclusions and impurities in the original cast material or by various other conditions related to material preparation and to the rolling operation.

Wavy edges on sheets (Fig. 13.8a) are the result of roll bending. The strip is thinner along its edges than at its center (see Fig. 13.4a); thus, the edges elongate more than 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 Figs. 13.8b and c are usually the result of poor material ductility at the rolling temperature. Because the quality of the edges of the sheet may affect sheet-metal-forming operations, edge defects in rolled sheets often are removed by shearing and slitting operations (Section 16.2). **Alligatoring** (Fig. 13.8d) is a complex phenomenon and typically is caused by nonuniform bulk deformation of the billet during rolling or by the presence of defects in the original cast material.

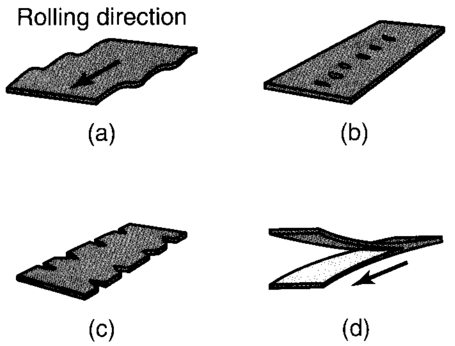


FIGURE 13.8 Schematic illustration of typical defects in flat rolling: (a) wavy edges; (b) zipper cracks in the center of the strip; (c) edge cracks; and (d) alligatoring.

13.3.2 Other Characteristics of Rolled Metals

Residual Stresses. Because of nonuniform deformation of the material in the roll gap, residual stresses can develop in rolled plates and sheets, especially during cold rolling. Small-diameter rolls or small thickness reductions per pass tend to plastically deform the metal more at its surfaces than in the bulk (Fig. 13.9a). This situation results in 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 the surfaces (Fig. 13.9b). This is due to the higher frictional constraint at the surfaces along the arc of contact—a situation that produces residual stress distributions that are the opposite of those with small-diameter rolls.

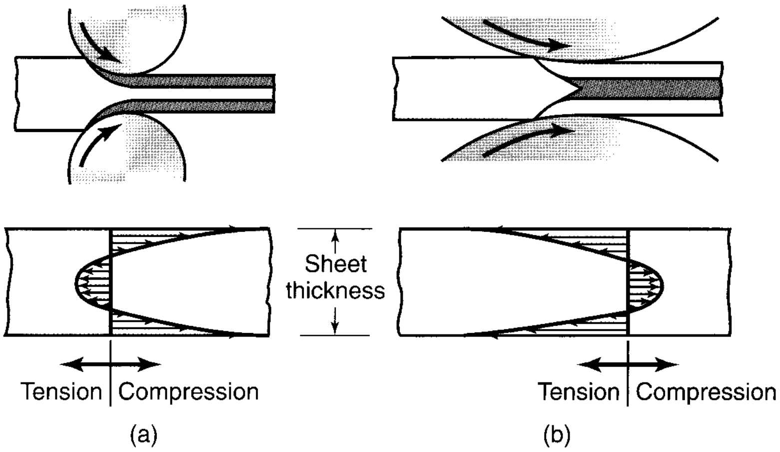


FIGURE 13.9 (a) Residual stresses developed in rolling with small-diameter rolls or at small reductions in thickness per pass. (b) Residual stresses developed in rolling with large-diameter rolls or at high reductions per pass. Note the reversal of the residual stress patterns.

Dimensional Tolerances. Thickness tolerances for cold-rolled sheets usually range from ±0.1 to 0.35 mm, depending on the thickness. Tolerances are much greater 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. 23.13, which, for comparison, includes other manufacturing processes. Note that cold rolling can produce a very fine surface finish; hence, products made of cold-rolled sheets may not require additional finishing operations, depending on the application. Note also that hot rolling and sand casting produce the same range of surface roughness.

Gage Numbers. The thickness of a sheet usually is identified by a *gage number:* the smaller the number, the thicker the sheet. Several numbering systems are used in industry, depending on the type of sheet metal being classified. Rolled sheets of copper and of brass also are identified by thickness changes during rolling, such as ¼ hard, ½ hard, and so on.

13.4 Rolling Mills

Several types of *rolling mills* and equipment are available with diverse roll arrangements. Although the equipment for hot and cold rolling is essentially the same, there are important differences in the roll materials, process parameters, lubricants, and cooling systems. The design, construction, and operation of rolling mills (Fig. 13.10) require major investments. Highly automated mills produce close-tolerance, high-quality plates and sheets at high production rates and low cost per unit weight, particularly when integrated with continuous casting. Rolling speeds may range up to 40 m/s. The width of rolled products may range up to 5 m.

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,

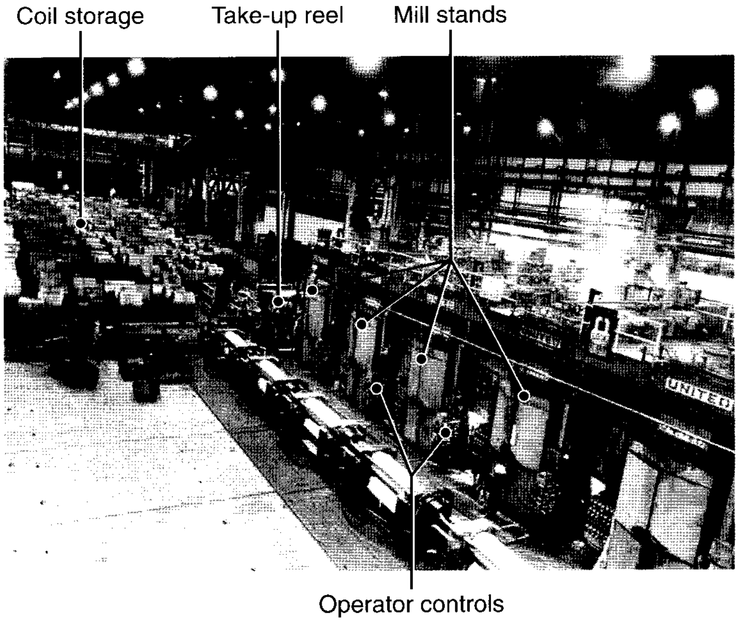


FIGURE 13.10 View of a rolling mill. *Source*: Courtesy of Ispat Inland.

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 is reversed after each pass, using elevator mechanisms and various manipulators. The plate being rolled, which may weigh as much as 145 metric tons, is raised repeatedly to the upper roll gap, rolled, then lowered to the lower roll gap, rolled, and so on.

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 lower roll forces (because of small roll-strip contact area) and power requirements and reduce spreading. Moreover, when worn or broken, small rolls can be replaced at lower cost than can large ones. On the other hand, small rolls deflect more under roll forces and 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 can be very high, it is particularly suitable for cold rolling thin sheets of high-strength metals. 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 13.11). Each stand consists of a set of rolls with its own housing and controls; a group of stands is called a *train*. The control of the strip thickness and the speed at which the strip travels through each roll gap

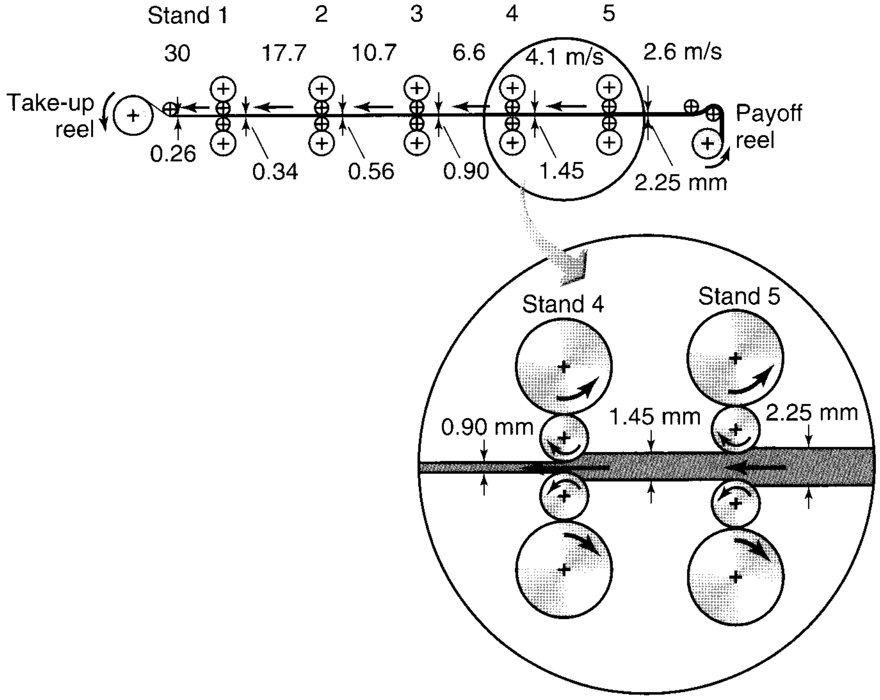


FIGURE 13.11 An example of a tandem-rolling operation.

is critical. Extensive electronic and computer controls are used in these operations, particularly in precision rolling.

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 also are polished. Rolls made for cold rolling should not be used for hot rolling, because they may crack from thermal cycling (*heat checking*) and *spalling* (cracking or flaking of surface layers). Recall also from earlier discussions that the elastic modulus of the roll influences roll deflection and flattening.

Note that the bottom surface of an aluminum beverage can, for example, has what appear to be longitudinal scratches on it. This is explained by the fact that the surface is a replica of the surface finish of the roll, which is produced by grinding (see Fig. 26.2a). In this way, the rolling direction of the original aluminum sheet also can be determined easily.

Lubricants. Hot rolling of ferrous alloys usually is carried out without lubricants, although graphite may be used. Water-based solutions are used to cool the rolls and to break up the scale on the rolled material. Nonferrous alloys are hot rolled with 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.

13.5 Various Rolling Processes and Mills

Several rolling processes and mills have been developed 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 at elevated temperatures by *shape rolling* (*profile rolling*), in which the stock goes through a set of specially designed rolls (Fig. 13.12; see also Fig. 13.1). *Cold shape rolling* also can be done with the starting materials in the shape of wire with various cross sections. Because the material's cross section is reduced nonuniformly, the design of a series of rolls (**roll-pass design**) requires considerable experience in order to avoid external and internal defects, hold dimensional tolerances, and reduce roll wear.

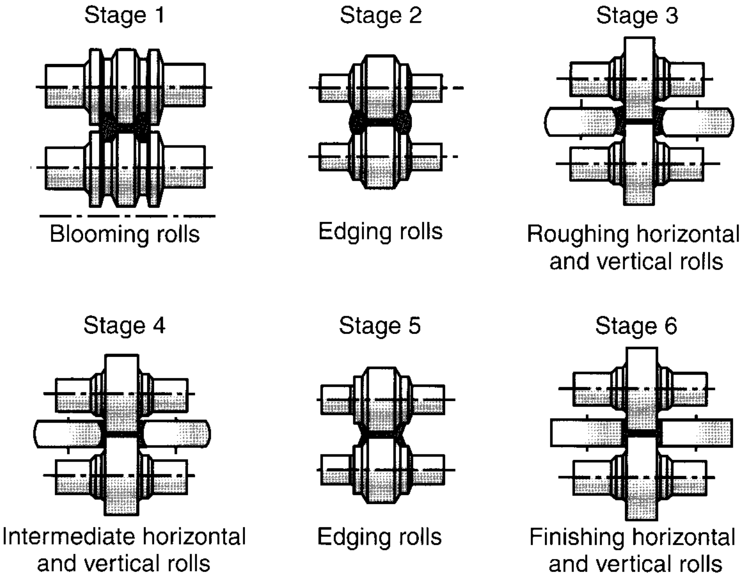


FIGURE 13.12 Steps in the shape rolling of an I-beam part. Various other structural sections, such as channels and rails, also are rolled by this kind of process.

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 profiled grooves (Fig. 13.13). Roll forging typically is used to produce tapered shafts and leaf springs, table knives, and

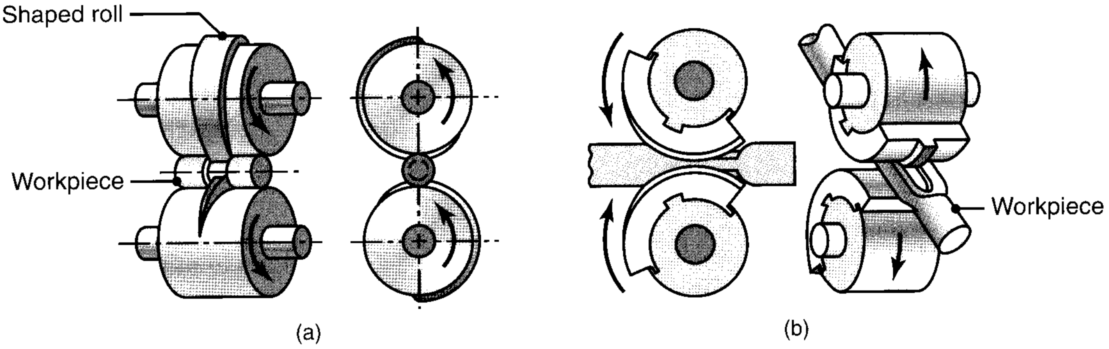


FIGURE 13.13 Two examples of the roll-forging operation, also known as *cross-rolling*. Tapered leaf springs and knives can be made by this process. *Source*: After J. Holub.

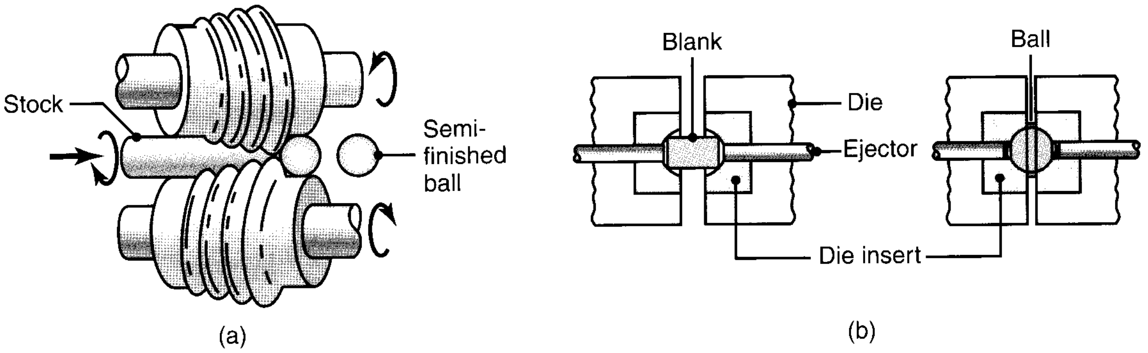


FIGURE 13.14 (a) Production of steel balls by the skew-rolling process. (b) Production of steel balls by upsetting a cylindrical blank. Note the formation of flash. The balls made by these processes subsequently are ground and polished for use in ball bearings.

hand tools; it also may be used as a preliminary forming operation, to be followed by other forging processes.

Skew Rolling. A process similar to roll forging is *skew rolling*, typically used for making ball bearings (Fig. 13.14a). Round wire or rod is fed into the roll gap, and roughly spherical blanks are formed continuously by the action of the rotating rolls. Another method of forming near-spherical blanks for ball bearings is to shear pieces from a round bar and then upset them in headers (see also Fig. 14.11) between two dies with hemispherical cavities (Fig. 13.14b). The balls subsequently are ground and polished in special machinery (see Fig. 26.15).

Ring Rolling. In *ring rolling*, a thick ring is expanded into a large-diameter thinner one. The ring is placed between two rolls, one of which is driven while the other is idle (Fig. 13.15a). Its thickness is reduced by bringing the rolls closer together as they rotate. Since the volume of the ring material remains constant during plastic deformation (volume constancy), the reduction in ring thickness results in an increase in its diameter. Depending on its size, the ring-shaped blank may be produced by such means as cutting from a plate, piercing, or cutting a thick-walled pipe. Various shapes can be ring rolled using shaped rolls (Fig. 13.15). Note that 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.

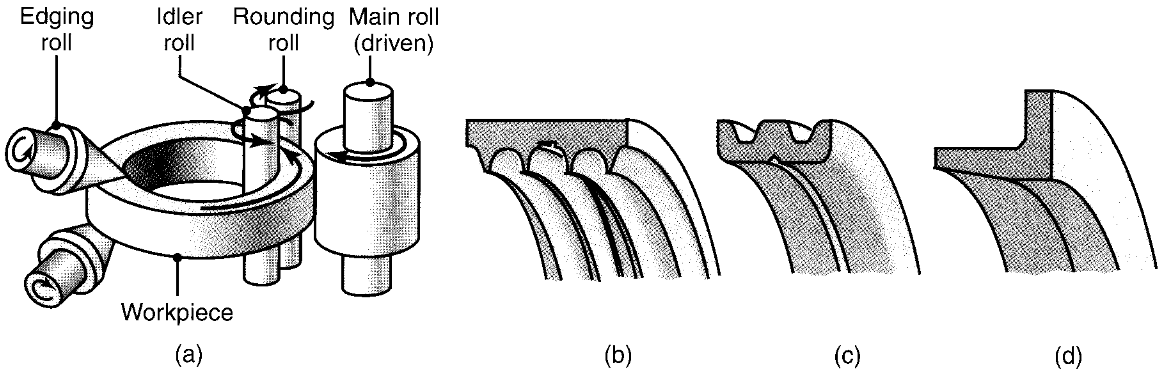


FIGURE 13.15 (a) Schematic illustration of a ring-rolling operation. Thickness reduction results in an increase in the part diameter. (b) through (d) Examples of cross sections that can be formed by ring rolling.

Typical applications of ring rolling are large rings for rockets and turbines, jet-engine cases, gearwheel rims, ball-bearing and roller-bearing races, flanges, and reinforcing rings for pipes. The process can be carried out at room temperature or at an elevated temperature, depending on the size (which can be up to 3 m, in diameter), strength, and ductility of the workpiece material. Compared with other manufacturing processes 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. The design of the profile rolls requires considerable experience. Analytical techniques are being developed to rely less on established practice and help minimize defects in rolled products.

Thread Rolling. *Thread rolling* is a cold-forming process by which straight or tapered threads are formed on round rods or wire. The threads are formed on the rod or wire with each stroke of a pair of flat reciprocating dies (Fig. 13.16a). In another method, threads are formed with *rotary dies* (Fig. 13.16c), at production rates as high as 80 pieces per second. Typical products are screws, bolts, and similar threaded parts. Depending on die design, the major diameter of a rolled thread may or may not be larger than a machined thread (Fig. 13.17a)—that is, the same as the blank diameter. The thread-rolling process is capable of generating other shapes as well, such as grooves and various gear forms, and it is used to produce almost all threaded fasteners at high production rates.

The thread-rolling process has the advantages of generating threads with good strength (due to cold working) and without any loss of material (scrap). The surface finish produced is very smooth, and the process induces compressive residual stresses on the workpiece surfaces, thus improving fatigue life. Thread rolling is superior to other methods of thread manufacture, because machining the threads cuts through the grain-flow lines of the material, whereas rolling the threads results in a grain-flow pattern that improves the strength of the thread (Fig. 13.17).

Spur and helical gears can be produced by cold-rolling processes similar to thread rolling (see also Section 24.7). The process 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**. This operation is similar to external thread rolling, and it produces accurate internal threads with good strength.

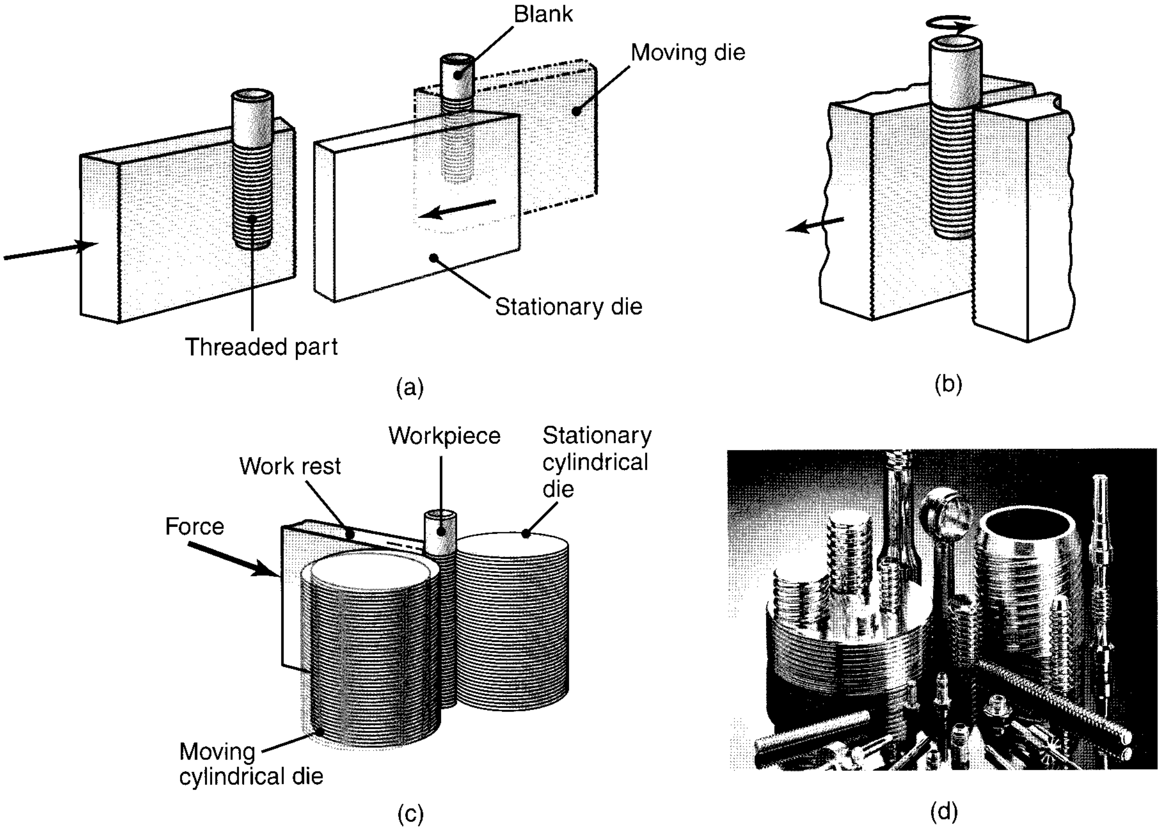


FIGURE 13.16 Thread-rolling processes: (a) and (b) reciprocating flat dies; (c) two-roller dies; (d) A collection of thread-rolled parts made economically at high production rates. *Source*: Courtesy of Tesker Manufacturing Corp.

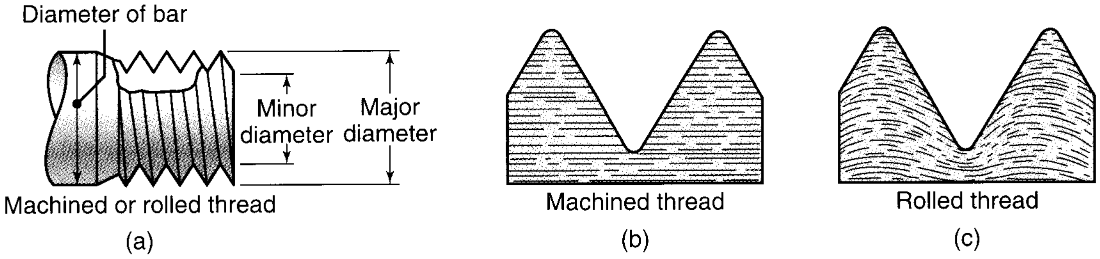


FIGURE 13.17 (a) Features of a machined or rolled thread. Grain flow in (b) machined and (c) rolled threads. Unlike machining, which cuts through the grains of the metal, the rolling of threads imparts improved strength because of cold working and favorable grain flow.

Lubrication is important in thread-rolling operations in order to obtain a good surface finish and surface integrity and to minimize defects. Lubrication affects the manner in which the material deforms during deformation, which is an important consideration because of the possibility of internal defects being developed (for example, see Fig. 14.16). Typically made of hardened steel, rolling dies are expensive because of their complex shape. They usually cannot be reground after they are worn. With proper die materials and preparation, however, die life may range up to millions of pieces.

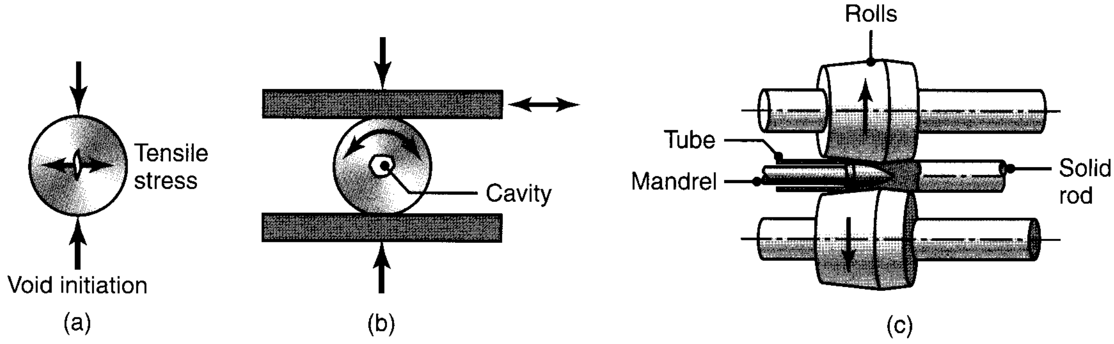


FIGURE 13.18 Cavity formation in a solid, round bar and its utilization in the rotary tube-piercing process for making seamless pipe and tubing. (See also Fig. 2.9.)

Rotary Tube Piercing. Also known as the **Mannesmann process**, this is a hot-working operation for making long, thick-walled *seamless pipe and tubing* (Fig. 13.18). Developed in the 1880s, this process is based on the principle that when a round bar is subjected to radial compressive forces, tensile stresses develop at the center of the bar (see Fig. 2.9). When it is subjected continuously to these cyclic compressive stresses (Fig. 13.18b), the bar begins to 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. 13.18b.)

Rotary tube piercing is carried out using an arrangement of rotating rolls (Fig. 13.18c). 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. Because of the severe deformation that the bar undergoes, the material must be high in quality and free from defects (since internal defects may propagate rapidly and cause premature failure of the part during forming).

Tube Rolling. The diameter and thickness of pipes and tubing can be reduced by *tube rolling*, which utilizes shaped rolls (Fig. 13.19). Some of these operations can be carried out either with or without an internal mandrel. In the *pilger mill*, the tube and an internal mandrel undergo a reciprocating motion; the rolls are specially shaped and are rotated continuously. During the gap cycle on the roll, the tube is advanced and rotated, starting another cycle of tube reduction. As a result, the tube undergoes a reduction in both diameter and wall thickness. Steel tubing of 265 mm in diameter have been produced by this process. Other operations for tube manufacturing are described in Chapter 15.

13.5.1 Various Mills

Integrated Mills. These mills are large facilities that involve complete integration of the 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. Competition in the steel industry has led to the development of *minimills*, in which 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, which is obtained

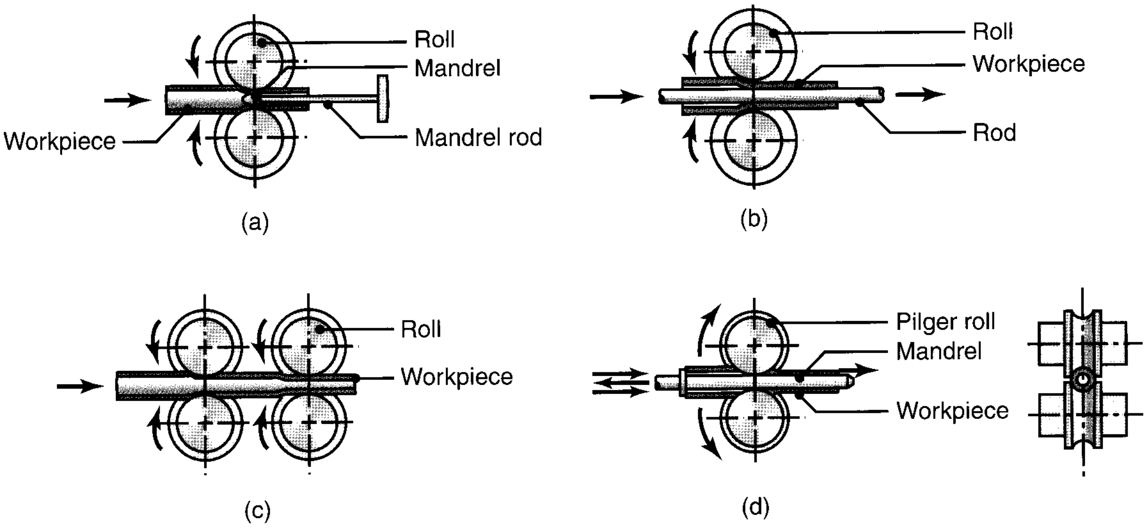


FIGURE 13.19 Schematic illustration of various tube-rolling processes: (a) with a fixed mandrel; (b) with a floating mandrel; (c) without a mandrel; and (d) pilger rolling over a mandrel and a pair of shaped rolls. Tube diameters and thicknesses also can be changed by other processes, such as drawing, extrusion, and spinning.

locally (to reduce transportation costs), is typically old machinery, cars, and farm equipment. Minimills have the economic advantages of low-investment optimal operations for each type of metal and product line and of low labor and energy costs. The products usually are aimed at markets in the mill's particular geographic area.

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. In addition to flat rolling, shape rolling is used to make products with various cross sections. Products made by rolling include: (a) plate, sheet, foil, rod, seamless pipe, and tubing; (b) shape-rolled products, such as I-beams and structural shapes; and (c) bars of various cross section. Other rolling operations include ring rolling and thread rolling.
- Rolling may be carried out at room temperature (cold rolling) or at elevated temperatures (hot rolling). The process 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 semifinished products have become a common practice because of their economic benefits.
- Integrated mills are large facilities involving the complete sequence of activities, from the production of hot metal in a blast furnace to the casting and the rolling of finished products 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
Thread rolling
Tube rolling
Wrought structure

BIBLIOGRAPHY

Ginzburg, V.B., **High-Quality Steel Rolling: Theory and Practice**, Marcel Dekker, 1993.

Ginzburg, V.B., and Ballas, R., **Flat Rolling Fundamentals**, CRC Press, 2001.

Hosford, W.F., and Caddell, R.M., **Metal Forming: Mechanics and Metallurgy**, 3rd ed., Prentice Hall, 2008.

Lange, K. (ed.), **Handbook of Metal Forming**, McGraw-Hill, 1985.

Lee, Y., **Rod and Bar Rolling: Theory and Applications**. CRC Press, 2004.

Lenard, J.G., **Primer on Flat Rolling**, Elsevier, 2007.

Tschaetch, H., **Metal Forming Practice: Processes, Machines, Tools**, Springer, 2007.

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. Explain the types of deflections that rolls undergo.

13.5. Describe the difference between a bloom, a slab, and a billet.

13.6. Why may roller leveling be a necessary operation?

13.7. List the defects commonly observed in flat rolling.

13.8. What are the advantages of tandem rolling? Pack rolling?

13.9. How are seamless tubes produced?

13.10. Why is the surface finish of a rolled product better in cold rolling than in hot rolling?

13.11. What is a *Sendzimir* mill? What are its important features?

QUALITATIVE PROBLEMS

13.12. Explain why the rolling process was invented and developed.

13.13. 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.

13.14. Explain how the residual stress patterns shown in Fig. 13.9 become reversed when the roll radius or reduction-per-pass is changed.

13.15. Explain whether it would be practical to apply the roller-leveling technique shown in Fig. 13.7a to thick plates.

13.16. Describe the factors that influence the magnitude of the roll force, *F*, in Fig. 13.2c.

13.17. Explain how you would go about applying front and back tensions to sheet metals during rolling. How would you go about controlling these tensions?

13.18. What typically is done to make sure that the product in flat rolling is not crowned?

13.19. Make a list of some parts that can be made by (a) shape rolling and (b) thread rolling.

13.20. 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?

13.21. 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.

13.22. 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?

13.23. Explain the consequence of applying too high a back tension in rolling.

13.24. 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.

13.25. Describe the importance of controlling roll speeds, roll gaps, temperature, and other process variables in a tandem-rolling operation, as shown in Fig. 13.11. Explain how you would go about determining the distance between the stands.

13.26. In Fig. 13.9a, if you remove the top compressive layer by, say, grinding, will the strip remain flat? If not, which way will it curve and why?

13.27. Name several products that can be made by each of the operations shown in Fig. 13.1.

13.28. List the possible consequences of rolling at (a) too high of a speed and (b) too low of a speed.

QUANTITATIVE PROBLEMS

▶ **13.29.** 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 200 mm.

▶ **13.30.** Calculate the individual drafts in each of the stands in the tandem-rolling operation shown in Fig. 13.11.

▶ **13.31.** Estimate the roll force, *F*, and the torque for an AISI 1020 carbon-steel strip that is 200 mm wide, 10 mm thick, and rolled to a thickness of 7 mm. The roll radius is 200 mm, and it rotates at 200 rpm.

13.32. 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

13.33. A simple sketch of a four-high mill stand is shown in Fig. 13.3a. Make a survey of the technical literature and present a more detailed sketch for such a stand, showing the major components.

13.34. Obtain a piece of soft, round rubber eraser, such as that at the end of a pencil, and duplicate the process shown in Fig. 13.18b. Note how the central portion of the eraser will begin to erode, producing a hole.

13.35. If you repeat the experiment in Problem 13.34 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.

13.36. Design a set of rolls to produce cross sections other than those shown in Fig. 13.12.

Metal-Forging Processes and Equipment

CHAPTER 14

- This chapter describes the fundamentals of forging and related processes, including design and economic considerations.
- Open-die forging operations for producing simple shapes are discussed 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 also discussed.
- The chapter ends with a review of the design of forged parts, 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 components for machinery, transportation, and farm equipment.

Alternative processes: Casting, powder metallurgy, machining, and fabrication.

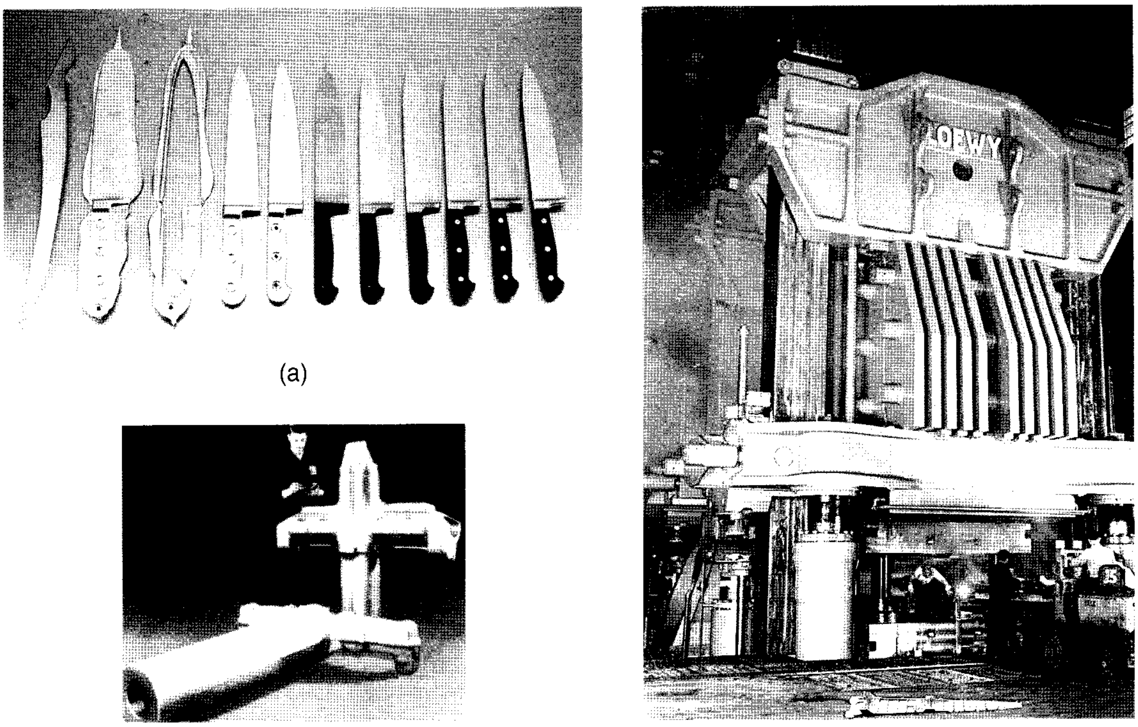
EXAMPLE:

CASE STUDIES:

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 first was 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; bolts and rivets; cutlery (Fig. 14.1a); hand tools; numerous structural components for machinery, aircraft (Fig. 14.1b), and railroads; and a variety of other transportation equipment.

Unlike rolling operations described in Chapter 13 that generally produce continuous plates, sheets, strips, or 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

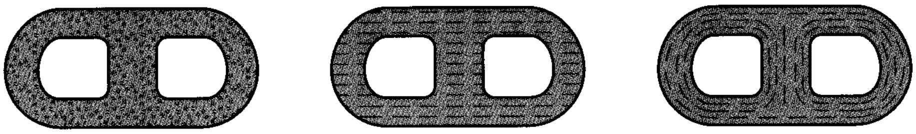


(a)

(b)

(c)

FIGURE 14.1 (a) Illustration of the steps involved in forging a knife. (b) Landing-gear components for the C5A and C5B transport aircraft, made by forging. (c) General view of a 445-MN (50,000-ton) hydraulic press. *Source*: (a) Courtesy of Mundial, LLC. (b) and (c) Courtesy of Wyman-Gordon Company.



(a) (b) (c)

FIGURE 14.2 Schematic illustration of a part 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.

can be performed with a heavy hammer and an anvil, as has been done traditionally by blacksmiths. However, most forgings require a set of dies and such equipment as a press or a powered forging hammer.

Forging may be carried out at room temperature (*cold forging*) or at elevated temperatures (*warm* or *hot forging*) depending on the homologous temperature; (see Section 1.8). Cold forging requires higher forces (because of the higher strength of the workpiece material), and the workpiece material must possess sufficient ductility at room temperature to undergo the necessary deformation without cracking. Cold-forged parts have a 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 in cold forging.

Forgings generally are subjected to additional finishing operations, such as heat treating to modify properties and machining to obtain accurate final dimensions and a good surface finish. These finishing operations can be minimized by *precision forging*, which is an important example of *net-shape* or *near-net-shape* forming processes. As we shall see throughout this book, components that can be forged successfully also may be manufactured economically by other methods, such as casting (Chapter 11), powder metallurgy (Chapter 17), 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 275 metric tons have been made. Part sizes may range from very small (the size of nails, pins, and bolts) to very large (up to 23 m, long shafts for ship propellers). Open-die forging can be depicted by a solid workpiece placed between two flat dies and reduced in height by compressing it (Fig. 14.3a)—a process that is also called **upsetting** or **flat-die forging.** The die surfaces also may have shallow cavities or incorporate features to produce relatively simple forgings.

The deformation of a workpiece under *frictionless conditions* is shown in Fig. 14.3b. Because constancy of volume is maintained, any reduction in height increases the diameter of the forged part. Note that the workpiece is deformed *uniformly*. In actual operations, however, there is friction, and the part develops a *barrel* shape (Fig. 14.3c)—a deformation mode also known as *pancaking*.

Barreling is caused primarily by frictional forces that oppose the outward flow of the workpiece at the die interfaces and thus can be minimized by using an effective

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

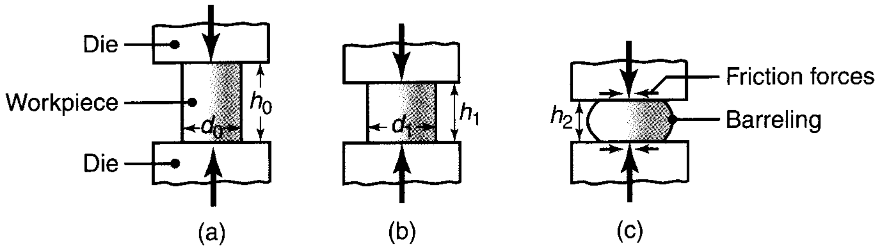


FIGURE 14.3 (a) Solid cylindrical billet upset between two flat dies. (b) Uniform deformation of the billet without friction. (c) Deformation with friction. Note barreling of the billet caused by friction forces at the billet–die interfaces.

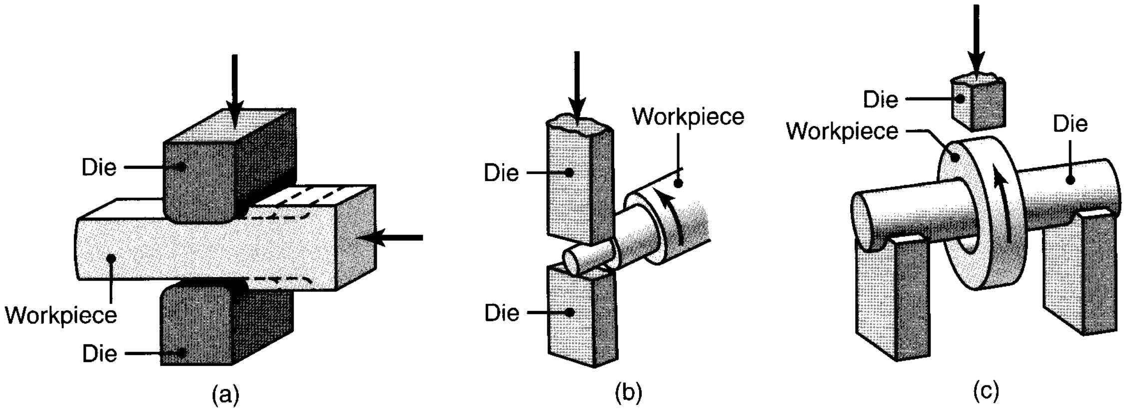


FIGURE 14.4 (a) Schematic illustration of a cogging operation on a rectangular bar. Blacksmiths use this process to reduce the thickness of bars by hammering the part on an anvil. Reduction in thickness is accompanied by barreling, as in Fig. 14.3c. (b) Reducing the diameter of a bar by open-die forging; note the movements of the dies and the workpiece. (c) The thickness of a ring being reduced by open-die forging.

lubricant. Barreling also can develop in upsetting hot workpieces between cold dies. The material at or near the die surfaces cools rapidly, while the rest of the workpiece 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, at the 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 by similar open-die forging techniques, as shown in Figs. 14.4b and c. Because the contact area between the die and the workpiece is small, a long section of a bar can be reduced in thickness without requiring large forces or heavy machinery. Blacksmiths perform such operations with a hammer and an anvil, using hot pieces of metal. Typical products are iron fences of various designs. Note that cogging can be a rough substitute for rolling operations. Cogging of larger workpieces usually is done using mechanized equipment and computer controls in which lateral and vertical movements are coordinated to produce the desired part.

Forging Force. The *forging force*, F, in an *open-die forging* operation on a solid cylindrical workpiece can be estimated from the formula

$$F = Y_f \pi r^2 \left(1 + \frac{2\,\mu r}{3h}\right), \qquad (14.1)$$

where Y_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, respectively, the instantaneous radius and height of the workpiece. (Derivations of this formula and of others for various forging processes are given in references listed in the bibliography at the end of the chapter.)

EXAMPLE 14.1 Calculation of Forging Force in Upsetting

A solid cylindrical slug made of 304 stainless steel is 150 mm in diameter and 100 mm high. It is reduced in height by 50% at room temperature by open-die forging with flat dies. Assuming that the coefficient of friction is 0.2, calculate 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. Thus, 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,

$$(\pi)(75)^2(100) = (\pi)(r)^2(50).$$

Therefore, $r = 106$ mm.

The quantity Y_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

$$\varepsilon = \ln\left(\frac{100}{50}\right) = 0.69.$$

We can determine the flow stress 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 method is to refer to Fig. 2.6 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. 2.6 are from different sources. Taking the latter value, the forging force now can be calculated, noting that in this problem the units in Eq. (14.1) must be in N and m. Thus,

$$F = (1000)(10^6)(\pi)(0.106)^2(1) + \frac{(2)(0.2)(0.106)}{(3)(0.050)}$$

$$= 4.5 \times 107 \text{ N} = 45 \text{ MN}.$$

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. 14.5a through c). This process usually is carried out at elevated temperatures to lower the required forces and attain enhanced ductility in the workpiece. Note in Fig. 14.5c 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 presents a severe constraint on any outward flow of the material in the die. Thus, based on the principle that in plastic deformation the material flows in the direction of least resistance (because it requires less energy), the material flows preferentially into the die cavity, ultimately filling it completely.

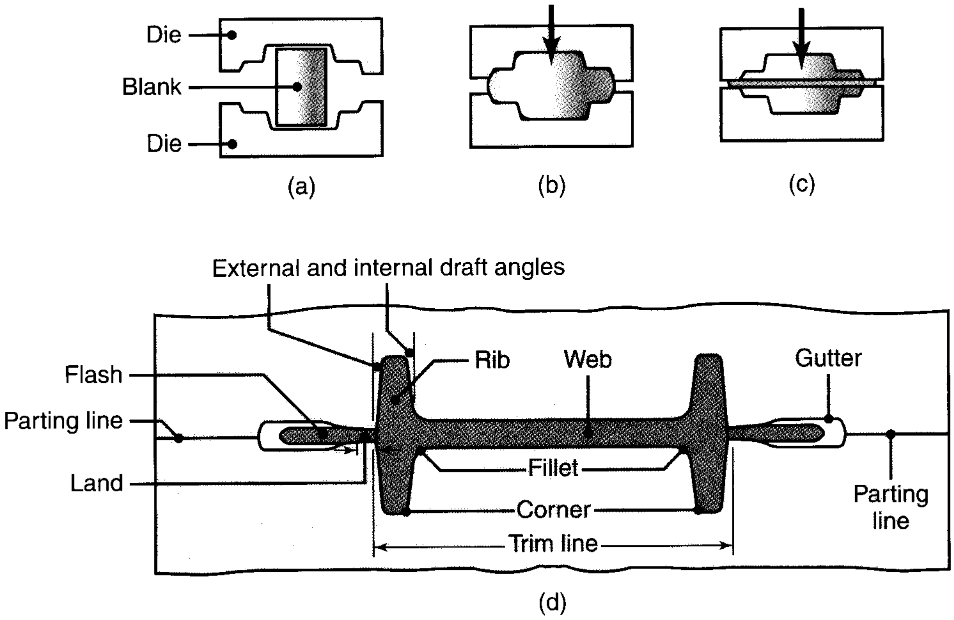


FIGURE 14.5 (a) through (c) Stages in impression-die forging of a solid round billet. Note the formation of flash, which is excess metal that is subsequently trimmed off. (d) Standard terminology for various features of a forging die.

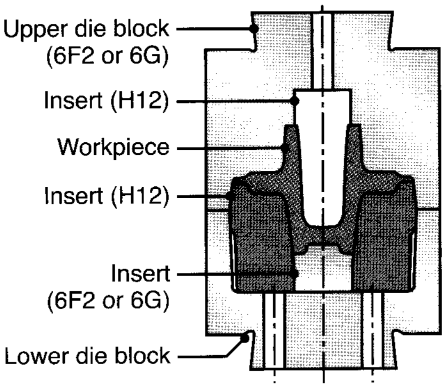


FIGURE 14.6 Die inserts used in forging an automotive axle housing. (See Section 5.7 for die materials.)

The standard terminology for a typical forging die is shown in Fig. 14.5d. Instead of being made as one piece, dies may be made of several pieces (segmented), including **die inserts** (Fig. 14.6) and particularly for complex shapes. The inserts can be replaced easily in the case of wear or failure in a particular section of the die and usually are made of stronger and harder materials.

The blank to be forged is prepared by (a) *cropping* from an extruded or drawn bar stock; (b) preforming from operations such as *powder metallurgy*; (c) *casting*; or (d) using a preformed blank from a prior forging operation. The blank is placed on the lower die, and as the upper die begins to descend, the blank's shape gradually changes—as is shown for the forging of a connecting rod in Fig. 14.7a.

Preforming operations (Figs. 14.7b and c) typically are used to distribute the material properly into various regions of the blank using simple shaped dies of various contours. In **fullering**, material is *distributed away* from an area. In **edging**, it is *gathered into* a localized area. The part then is formed into the rough shape (say, a connecting rod) by a process called **blocking**, using *blocker dies*. The final operation is the finishing of the forging in *impression dies* that give the forging its final shape. The flash is removed later by a trimming operation (Fig. 14.8).

Forging Force. The *forging force, F*, required to carry out an *impression-die forging* operation can be estimated from the formula

$$F = kY_fA, \tag{14.2}$$

where k is a multiplying factor obtained from Table 14.2, Y_f is the flow stress of the material at the forging temperature, and A is the projected area of the forging,

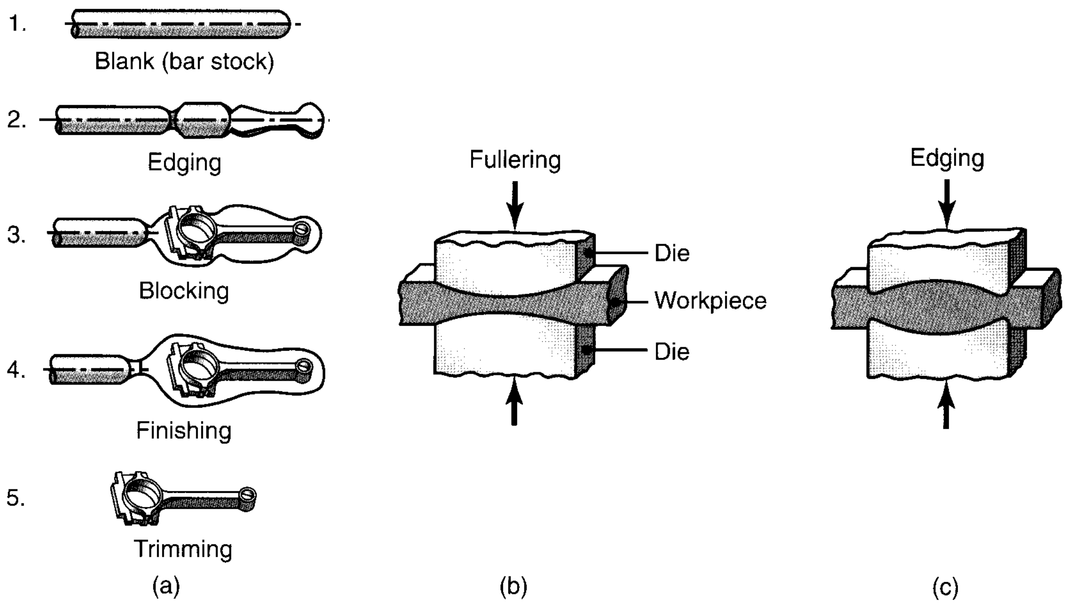


FIGURE 14.7 (a) Stages in forging a connecting rod for an internal combustion engine. Note the amount of flash required to ensure proper filling of the die cavities. (b) Fullering and (c) edging operations to distribute the material properly when preshaping the blank for forging.

including the flash. 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 700 MPa, and a part (such as that shown in Fig. 14.7a) has a projected area (with flash) of 38,000 mm^2. Taking a value of $k = 10$ from Table 14.2, the forging force would be $F = (10)(700)(38{,}000) = 266$ MN.

Closed-die Forging. The process shown in Fig. 14.5 also is referred to as *closed-die forging*. However, in true closed-die forging, flash does not form (hence the term *flashless forging*), and the workpiece completely fills the die cavity (see right side of Fig. 14.9b). Consequently, the forging pressure is very high, and accurate control of the blank volume and proper die design are essential to producing a forging with the desired 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 machine to jam.

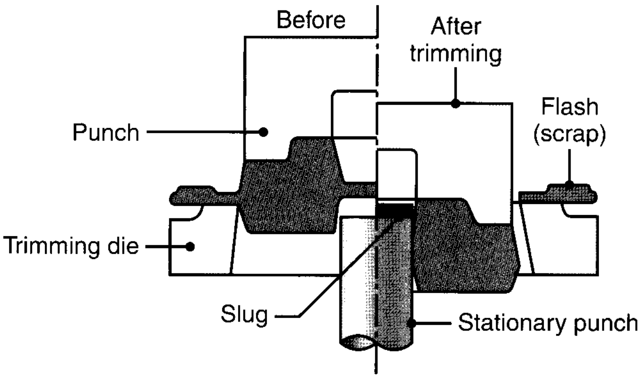


FIGURE 14.8 Trimming flash from a forged part. Note that the thin material at the center is removed by punching.

TABLE 14.2

Range of k Values for Eq. (14.2)

Shape	k
Simple shapes, without flash	3–5
Simple shapes, with flash	5–8
Complex shapes, with flash	8–12

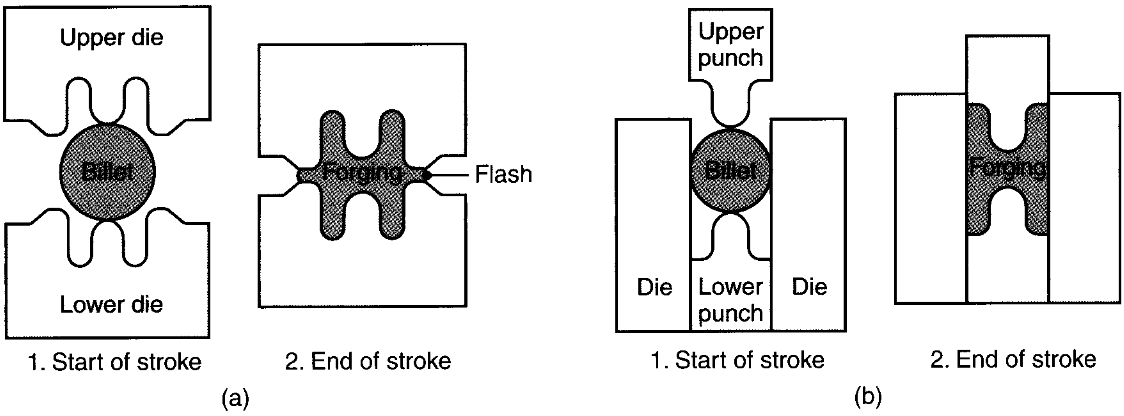


FIGURE 14.9 Comparison of (a) closed-die forging with flash and (b) precision or flashless forging of a round billet. *Source*: After H. Takemasu, V. Vazquez, B. Painter, and T. Altan.

Regardless, the term *closed-die forging* is often applied to impression die forging with flash generation, whereas *open-die forging* generally applies to operations with simple dies and tooling and with large deformations.

Precision Forging. In order to reduce the number of additional finishing operations required—hence the 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 in the die cavity. Also, because of the higher forces required to obtain fine details on the part, this process requires higher capacity equipment. Aluminum and magnesium alloys are particularly suitable for precision forging because of the relatively low forging loads and temperatures that they require; however, steels and titanium also can be precision forged.

Forging Practice and Product Quality. A forging operation typically involves the following sequence of steps:

1. Prepare a slug, billet, or preform by processes such as shearing (cropping), sawing, or cutting off. If necessary, clean surfaces by such means as shot blasting.
2. For hot forging, heat the workpiece in a suitable furnace and 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 scale (which is brittle) falls off during deformation.
3. For hot forging, preheat and lubricate the dies; for cold forging, lubricate the blank.
4. Forge the billet in appropriate dies and in the proper sequence. If necessary, remove any excess material (such as flash) by trimming, machining, or grinding.
5. Clean the forging, check its dimensions, and (if necessary) machine it to final dimensions and specified tolerances.
6. Perform additional operations, such as straightening and heat treating (for improved mechanical properties). Also, perform any finishing operations that may be required, such as machining and grinding.
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 are performed and controlled. Generally, dimensional tolerances range between ±0.5 and ±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, die closure (whether the dies have closed properly), and mismatching of the dies. The surface finish of the forging depends on blank preparation, die surface finish, die wear, and the effectiveness of the lubricant.

14.4 Various Forging Operations

Several other operations related to the basic forging process are carried out in order to impart the desired shape and features to forged products.

Coining. This is essentially a closed-die forging process that is typically used in the minting of coins, medallions, and jewelry (Fig. 14.10). The blank or slug is coined in a completely closed die cavity. In order to produce fine details (for example, the detail on newly minted coins), the pressures required can be as high as five or six times the strength of the material. On some parts, several coining operations may be required. Lubricants cannot be applied in coining, because they can become entrapped in the die cavities and (being incompressible) prevent the full reproduction of die-surface details and surface finish.

Marking parts with letters and numbers also can be done rapidly through coining. In addition, the process is used with forgings and other products to improve surface finish and to impart the desired dimensional accuracy with little or no change in part size. Called **sizing**, this process requires high pressures.

Heading. Also called **upset forging**, *heading* is essentially an upsetting operation, usually performed on the end of a round rod or wire in order to increase the cross section. Typical products are nails, bolt heads, screws, rivets, and various other fasteners (Fig. 14.11a). Heading can be carried out cold, warm, or hot. An important

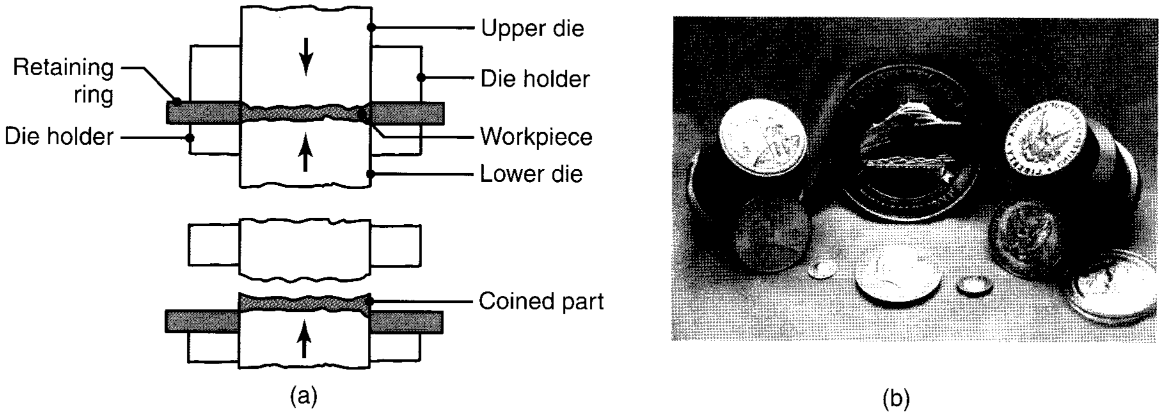


FIGURE 14.10 (a) Schematic illustration of the coining process. The earliest coins were made by open-die forging and lacked precision and sharp details. (b) An example of a modern coining operation, showing the coins and tooling. Note the detail and superior surface finish that can be achieved in this process. *Source*: Courtesy of C & W Steel Stamp Co., Inc.

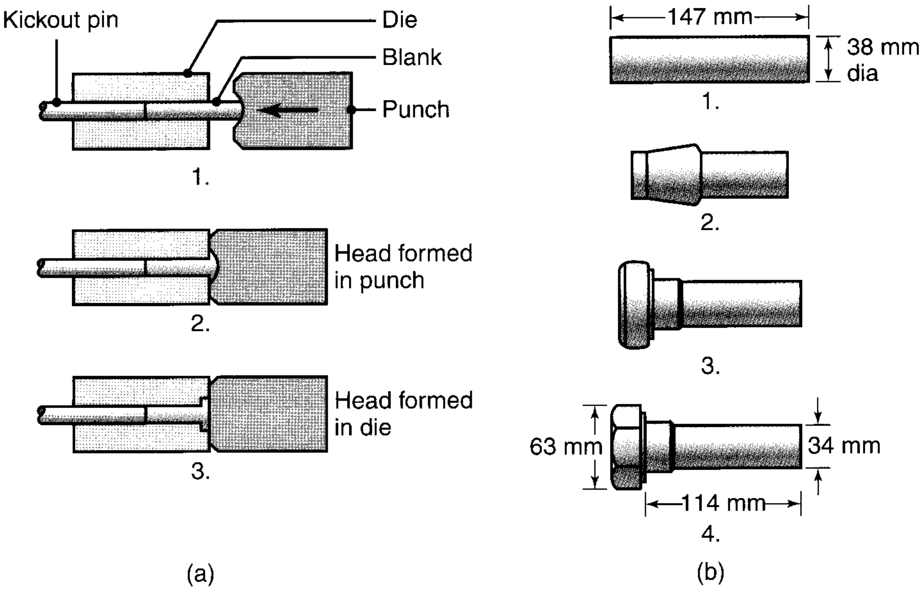


FIGURE 14.11 (a) Heading operation to form heads on fasteners, such as nails and rivets. (b) Sequence of operations used to produce a typical bolt head by heading.

consideration in heading is the tendency for the bar to buckle if its unsupported length-to-diameter ratio is too high. This ratio usually is limited to less than 3:1, but with appropriate dies, it can be higher. For example, higher ratios can be accommodated if the diameter of the die cavity is not more than 1.5 times the bar diameter.

Heading operations are performed on machines called **headers**, which usually 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*. These machines tend to be noisy; a soundproof enclosure or the use of ear protectors is required. Heading operations can be combined with cold-extrusion processes to make various parts, as described in Section 15.4.

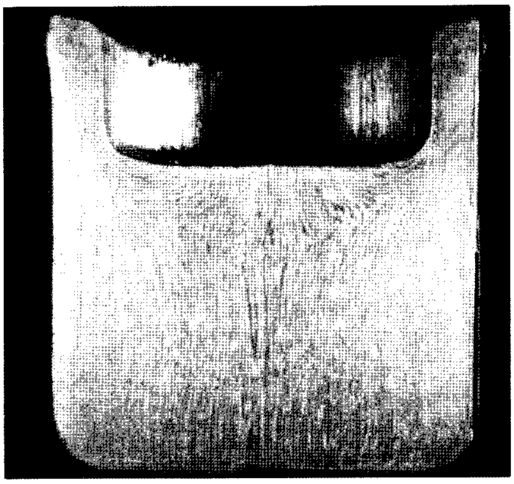


FIGURE 14.12 A pierced round billet showing grain-flow pattern. (See also Fig. 14.2c). *Source*: Courtesy of Ladish Co., Inc.

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. 14.12). The workpiece may be confined in a container (such as a die cavity) or may be unconstrained. The deformation of the workpiece will depend on how much it is constrained from flowing freely as the punch descends. A common example of piercing is the indentation of the hexagonal cavity in bolt heads. Piercing may be followed by punching to produce a hole in the part. (For a similar depiction of this situation, see the *slug* above the stationary punch in the central portion of Fig. 14.8.) Piercing also is performed to produce hollow regions in forgings using side-acting auxiliary equipment.

The *piercing force* depends on (a) the cross-sectional area and the tip geometry of the punch, (b) the strength of the material, and (c) the magnitude of friction at the sliding interfaces. The pressure may range from three to five times the strength of the material, which is approximately the same level of stress required to make an indentation in hardness testing (Section 2.6).

CASE STUDY 14.1 Manufacture of a Stepped Pin by Heading and Piercing Operations

Fig. 14.13a shows a stepped pin made from SAE 1008 steel and used as a part of a roller assembly to adjust the position of a car seat. The part is fairly 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. 14.13b. First, a solid, cylindrical blank is extruded in two operations, followed by an upsetting operation. 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 the impression-die forming, a piercing operation is performed to forms the bore. The part is made to net shape on a cold-forming machine at a rate of 240 parts per minute.

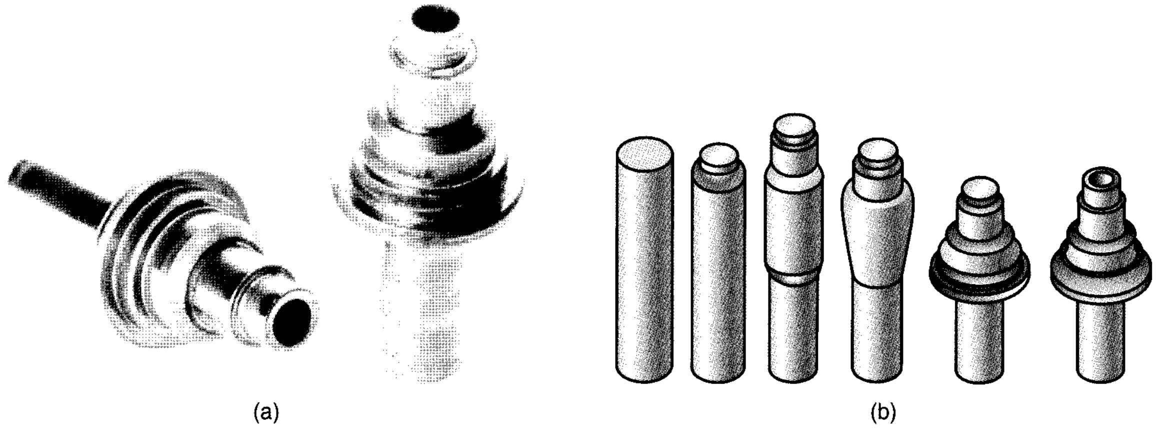


FIGURE 14.13 (a) The stepped pin used in Case Study 14.1. (b) Illustration of the manufacturing steps used to produce the stepped pin. *Source*: Courtesy of National Machinery, LLC.

Hubbing. This process consists of pressing a hardened punch with a particular 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 the making of tableware. The die cavity usually is shallow, but for deeper cavities, some material may be removed from the surface by machining prior to hubbing (see Figs. 24.2c and d). The *hubbing force* can be estimated from the equation

$$\text{Hubbing force} = 3(\text{UTS})(A), \qquad (14.3)$$

where UTS is obtained from Table 2.2 and A is the projected area of the impression. For example, for high-strength steel with UTS = 1500 MPa and a part with a projected area of 400 mm^2, the hubbing force is $(3)(1500\ \text{N/mm}^2)(400\ \text{mm}^2) = 1.8\ \text{MN}$.

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. Although not in common use, typical components that may be forged by this process 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* below). 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 shape in several small steps. The operation is somewhat similar to cogging (see Fig. 14.4a), in which the die penetrates the blank to different depths along the surface. Because of the smaller area of contact with the die, the process requires much lower forces compared with conventional impression-die forging, and the tools are simpler and less costly.

Isothermal Forging. Also known as **hot-die forging**, this process heats the dies to the same temperature as that of the hot workpiece. Because the workpiece remains hot, its flow strength and high ductility are maintained during forging. Also, the forging load is low, and material flow within the die cavity is improved. Complex parts with good dimensional accuracy can be isothermally forged to near-net shape by one stroke in a hydraulic press. The dies for hot forging of high-temperature alloys usually are made of nickel or molybdenum alloys (because of their resistance to high temperature), but steel dies can be used for aluminum alloys. Isothermal forging is expensive and the production rate is low. However, it can be economical for specialized, intricate forgings made of materials such as aluminum, titanium, and superalloys, provided that the quantity required is sufficiently high to justify the die costs.

Rotary Swaging. In this process (also known as *radial forging, rotary forging*, or simply *swaging*), a solid rod or tube is subjected to radial impact forces by a set of reciprocating dies of the machine (Figs. 14.14a and b). The die movements are obtained by means of a set of rollers in 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 obtained through the reciprocating motion of wedges (Fig. 14.14c). The dies can be opened wider than those in rotary swagers, thereby accommodating large-diameter or variable-diameter parts. In another type of machine, the dies do not rotate, but move radially in and out. Typical products made are screwdriver blades and soldering-iron tips.

Swaging also can be used to *assemble* fittings over cables and wire; in such cases, the tubular fitting is swaged directly onto the cable. The process is also used for operations such as *pointing* (tapering the tip of a cylindrical part) and *sizing* (finalizing the dimensions of a part).

Swaging generally is limited to a maximum workpiece diameter of about 150 mm; 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, with rates as high as 50 parts per minute possible, 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).

Tube Swaging. In this process, the internal diameter and/or the thickness of the tube is reduced with or without the use of *internal mandrels* (Figs. 14.15a 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. 14.15c). 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 has been built to swage gun barrels and other parts with starting diameters as large as 350 mm.

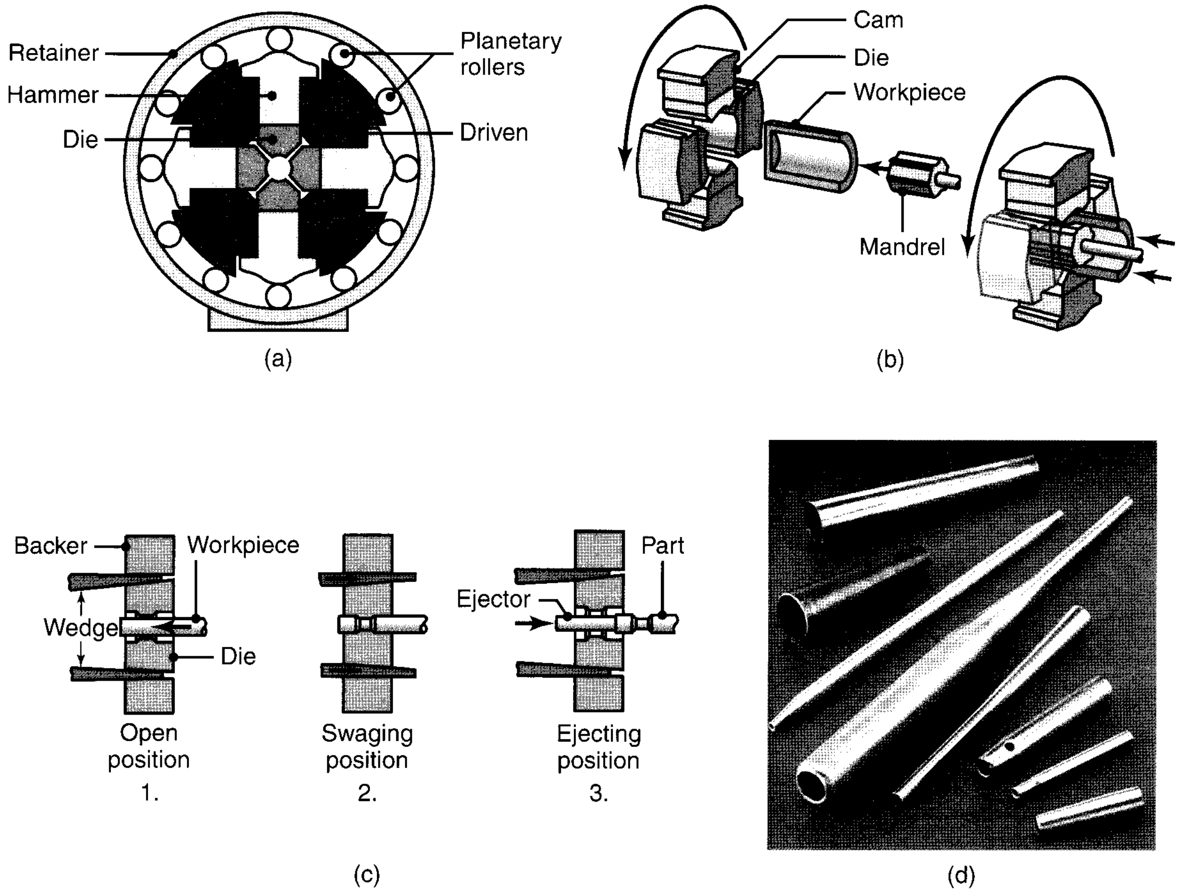


FIGURE 14.14 (a) Schematic illustration of the rotary-swaging process. (b) Forming internal profiles on a tubular workpiece by swaging. (c) A die-closing swaging machine, showing forming of a stepped shaft. (d) Typical parts made by swaging. *Source*: (d) Courtesy of J. Richard Industries.

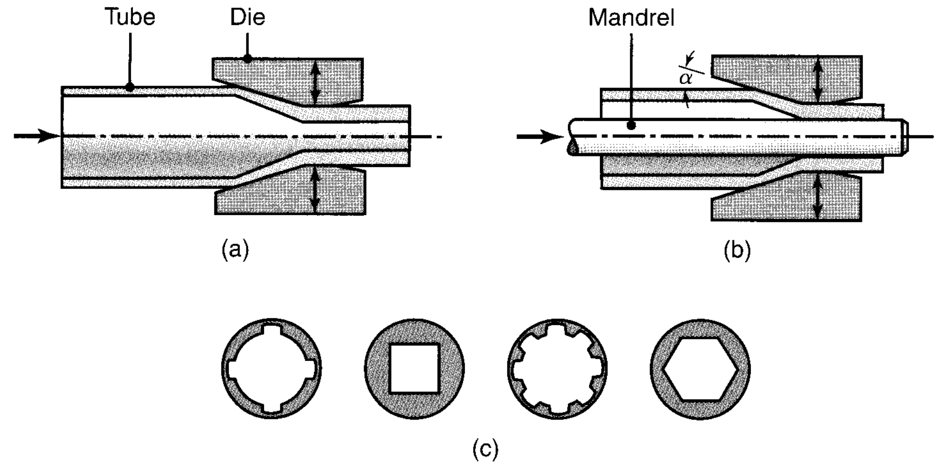


FIGURE 14.15 (a) Swaging of tubes without a mandrel; note the increase in wall thickness in the die gap. (b) Swaging with a mandrel; note that the final wall thickness of the tube depends on the mandrel diameter. (c) Examples of cross sections of tubes produced by swaging on shaped mandrels. Rifling (internal spiral grooves) in small gun barrels can be made by this process.

14.5 Forgeability of Metals; Forging Defects

Forgeability is generally defined as the capability of a material to undergo deformation without cracking. Various tests have been developed 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, and the reduction in height at which cracking on the barreled surfaces begins is noted (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 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. The hot-twist 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. Forgeability is based on considerations such as ductility and strength of the material, forging temperature required, frictional behavior, and the quality of the forgings produced. These ratings should be regarded only as general guidelines. Typical *hot-forging temperature* ranges for various metals and alloys are included in Table 14.3. Note that higher forging temperature does not necessarily indicate greater difficulty in forging that material. For *warm* forging, temperatures range from 200° to 300°C for aluminum alloys and from 550° to 750°C for steels.

Forging Defects. In addition to surface cracking, other defects can develop during forging as a result of the material flow pattern in the die, as described next in Section 14.6 regarding die design. For example, if there is an insufficient volume of material to fill the die cavity completely, the web may buckle during forging and develop laps (Fig. 14.16a). On the other hand, if the web is too thick, the excess material flows past the already formed portions of the forging and develops internal cracks (Fig. 14.16b).

TABLE 14.3

Forgeability of Metals, in Decreasing Order

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

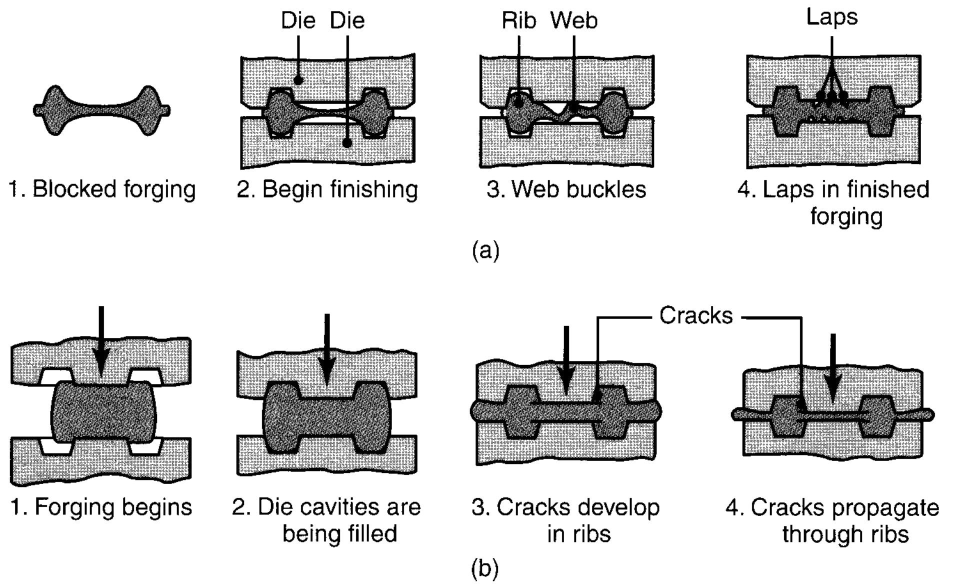


FIGURE 14.16 Examples of defects in forged parts. (a) Laps formed by web buckling during forging; web thickness should be increased to avoid this problem. (b) Internal defects caused by an oversized billet. Die cavities are filled prematurely, and the material at the center flows past the filled regions as the dies close.

The various radii in the forging-die cavity can significantly influence the formation of such defects. Internal defects also may develop from (a) nonuniform deformation of the material in the die cavity, (b) temperature gradients 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. 14.12. In this condition, known as **end grains**, the grain boundaries become exposed directly to the environment and can be attacked by it, developing a rough surface and also acting as stress raisers.

Forging defects can cause fatigue failures, and they also may lead to such problems as corrosion and wear during the service life of the forged component. The importance of inspecting forgings prior to their placement in service, particularly in critical applications, such as aircraft, 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 considerable knowledge and experience regarding the shape and complexity of the workpiece, its ductility, its strength and sensitivity to deformation rate and temperature, and its frictional characteristics. Die distortion under high forging loads is also an important design consideration, particularly if close dimensional tolerances are required. The most important rule in die design is the fact that the part will flow in the direction of least resistance. Thus, the workpiece *intermediate shapes* should be planned so that they properly fill the die cavities. An example of the intermediate shapes for a connecting rod is given in Fig. 14.7a.

With continuing advances in developing reliable simulation of all types of metalworking operations, software is available to help predict material flow in forging-die cavities. The simulation incorporates various conditions, such as workpiece temperature and heat transfer, to tooling, frictional conditions at die–workpiece contact surfaces, and forging speed. Such software can be very helpful in die design and in eliminating future problems with defective forgings (see also Section 38.7).

Preshaping. In a properly preshaped workpiece, the material should not flow easily into the flash (otherwise die filling will be incomplete), the grain flow pattern should be favorable for the products' strength and reliability, and sliding at the workpiece–die interfaces should be minimized in order to reduce die wear. The selection of preshapes requires considerable experience and involves calculations of cross-sectional areas at each location in the forging. Computer modeling and simulation techniques are useful in such calculations.

Die Design Features. The terminology for forging dies is shown in Fig. 14.5d, and the significance of various features is described next. Some of these considerations are similar to those for casting (Section 12.2). For most forgings, the **parting line** is located at the largest cross section of the part. For simple symmetrical shapes, the parting line is normally a straight line at the center of the forging, but for more complex shapes, the line may not lie in a single plane. The dies are then designed in such a way that they lock during engagement, in order to avoid side thrust, balance forces, and maintain die alignment during forging.

After sufficiently 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 of the forging. The length of the *land* is usually two to five times the flash thickness.

Draft angles are necessary in almost all forging dies in order to facilitate removal of the part from the die. Upon cooling, the forging shrinks both radially and longitudinally, so internal draft angles (about 7° to 10°) are made larger than external ones (about 3° to 5°).

Selection of the proper radii for corners and fillets is important in ensuring smooth flow of the metal into the die cavity and improving die life. Small radii generally are undesirable because of their adverse effect on metal flow and their tendency to wear rapidly (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, *allowances* are provided in forging-die design when machining the forging is necessary to obtain final desired dimensions and surface finish. Machining allowance should be provided at flanges, at holes, and at mating surfaces.

Die Materials. Most forging operations (particularly for large parts) are carried out at elevated temperatures. General requirements for die materials therefore are

- Strength and toughness at elevated temperatures
- Hardenability and ability to harden uniformly
- Resistance to mechanical and thermal shock
- Wear resistance, particularly resistance to abrasive wear, because of the presence of scale in hot forging.

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 and surface finish. Die-manufacturing methods are described in Section 14.7.

Lubrication. A wide variety of metalworking fluids can be used in forging, as described in Section 33.7. Lubricants greatly influence friction and wear. Consequently, they affect the forces required [see Eq. (14.1)], die life, and the manner in which the material flows into the 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 improving metal flow. Another important role of the lubricant is to act as a *parting agent*, preventing the forging from sticking to the dies and helping release it from the die.

14.7 Die-manufacturing Methods and Die Failures

From the topics described thus far, it should be evident that dies, their quality, and die life are highly significant aspects of the total manufacturing operation, including the quality of the parts produced. This is particularly noteworthy in view of the fact that the vast majority of discrete parts that are produced in large quantities (such as gears, shafts, bolts, etc.), as well as castings of all types of products, are made in individual dies and molds. Dies also have an impact on the overall economics of manufacturing, because of their cost and the lead time needed to produce them, as some dies require months to manufacture. Equally important considerations are the maintenance of dies and their modifications as parts are first produced.

Several manufacturing methods, either singly or in combination, can be used to make dies for forging, as well as for other metalworking processes. These methods include casting, forging, machining, grinding, electrical and electrochemical methods—particularly electrical-discharge machining (EDM) and wire EDM—and the use of lasers for small dies. An important and continuing development is the production of tools and dies by **rapid tooling** using rapid protoyping techniques, described in Section 20.5.

The process of producing a die 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 small dies with shallow cavities. Dies are usually heat treated for higher hardness and wear resistance (Chapter 33). If necessary, their surface profile and finish are improved further by finish grinding and polishing, either by hand or by programmable industrial robots.

The choice of a die-manufacturing method depends on its size and shape and the particular operation in which the die is to be used, such as casting, forging, extrusion, powder metallurgy, or plastics molding. As in all manufacturing operations, cost often dictates the process selected, because tool and die costs can be significant in manufacturing operations. Dies of various sizes and shapes 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 casting small dies). Cast steels generally are preferred for large dies because of their strength and toughness, as well as the ease with which the steel composition, grain size, and other properties can be controlled and modified.

Most commonly, dies are *machined* from *forged die blocks* by processes such as high-speed milling, turning, grinding, and electrical discharge and electrochemical machining. 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 (Section 25.6) and electrical and elctrochemical machining are a common practice. Typically, a die is machined by milling on computer-controlled machine tools with various software packages (see Fig. I.11) that have the capability (economically) of optimizing the cutting-tool path. Thus, the best surface finish can be obtained in the least possible machining time. Equally important is the setup for machining, because dies should be machined as much as possible 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 processes, to obtain the desired surface finish and dimensional accuracy. This also may include *laser surface treatments* and *coatings* (Chapter 34) to improve die life. Lasers are sometimes used for die repair and reconfiguration of the worn regions of dies (see also Fig. 33.11).

Die Costs. From the preceding discussion, it is evident that the cost of a die depends greatly on its size, shape complexity, application, and surface finish required, as well as the die material and manufacturing, heat treating, and finishing methods employed. Consequently, specific die costs cannot be categorized easily. 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 to make, and the cost of a set of dies for automotive body panels can be as much as $2 million. On the other hand, because a large number of parts usually are made from one set of dies, the *die cost per piece made* is generally a small portion of a part's manufacturing cost (see also Section 40.9). The lead time required to produce dies also can have a significant impact on the overall manufacturing cost, particularly in a global and competitive marketplace.

Die Failures. Failure of dies in manufacturing operations generally results from one or more of the following causes:

- Improper die design
- Defective or improper selection of die material
- Improper manufacturing and improper heat-treatment and finishing operations
- Overheating and heat checking (i.e., cracking caused by temperature cycling)
- Excessive wear
- Overloading (i.e., excessive force on the die)
- Improper alignment of the die components with respect to their movements
- Misuse
- Improper handling of the die.

Although these factors typically apply to dies made of tool and die steels, many also apply to other die materials, such as carbides, ceramics, and diamond.

The proper design of dies is as important as the proper selection of die materials. In order to withstand the forces involved, a die must have sufficiently large cross sections and clearances (to prevent jamming). 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 can have detrimental effects on 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 failure in cold-extrusion dies is that of the operator (or of a programmable robot) to fail to remove a formed part from the die before loading it with another blank.

14.8 Forging Machines

A variety of forging machines is available with a range of capacities (tonnage), speeds, and speed–stroke characteristics (Table 14.4).

Hydraulic Presses. These presses operate at constant speeds and are *load limited*, or load restricted. In other words, a press stops if the load required exceeds its capacity. Large amounts of energy can be transmitted to a workpiece by a constant load throughout a stroke—the speed of which can be controlled. Because forging in a hydraulic press takes longer than in the other types of forging machines described next, 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. 14.17), rams, and hydraulic pumps driven by electric motors. The ram speed can be varied during the stroke. Press capacities range up to 125 MN for open-die forging and up to 450 MN in North America, 640 MN in France, and 730 MN in Russia for closed-die forging. The main landing-gear support beam for the Boeing 747 aircraft is forged in a 450-MN hydraulic press, shown in Fig. 14.1c

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

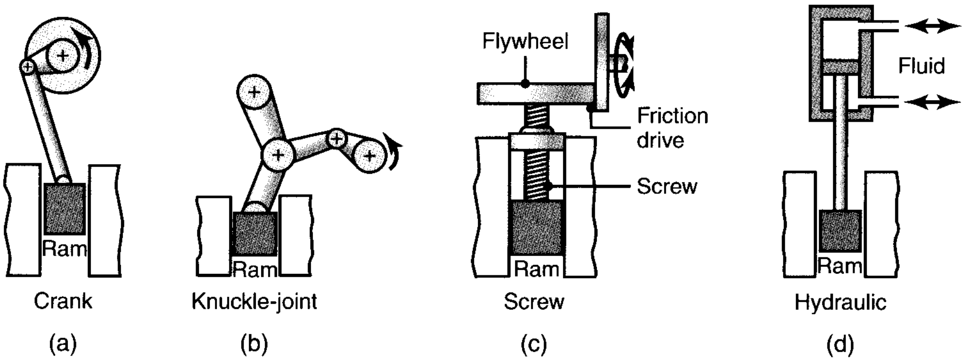


FIGURE 14.17 Schematic illustration of the principles of various forging machines. (a) Mechanical press with an eccentric drive; the eccentric shaft can be replaced by a crankshaft to give up-and-down motion to the ram. (b) Knuckle-joint press. (c) Screw press. (d) Hydraulic press.

with the part in the forefront. This part is made of a titanium alloy and weighs approximately 1350 kg.

Mechanical Presses. These presses are basically of either the crank or the eccentric type (Fig. 14.17a). The speed varies from a maximum at the center of the stroke to zero at the bottom of the stroke; thus, mechanical presses 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 translates the rotary motion into a reciprocating linear motion. A *knuckle-joint* mechanical press is shown in Fig. 14.17b. Because of the linkage design, very high forces can be applied in this type of press (see also Fig. 11.20).

The force available in a mechanical press depends on the stroke position and 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 107 MN. Mechanical presses are preferred for forging parts with high precision.

Screw Presses. These presses (Fig. 14.17c) 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 is dissipated. If the dies do not close at the end of the cycle, the operation is repeated until the forging is completed. Screw presses are used for various open-die and closed-die forging operations. They are suitable particularly for small production quantities, especially thin parts with high precision, such as turbine blades. Press capacities range from 1.4 to 280 MN.

Hammers. Hammers derive their energy from the potential energy of the ram, which is converted into kinetic energy; hence, they are *energy limited.* Unlike hydraulic presses, hammers (as the name implies) operate at high speeds, and the resulting low forming time minimizes the cooling of a hot forging. Low cooling rates then allow the forging of complex shapes, particularly those with thin and deep recesses. To complete the forging, several successive blows usually are 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 at about 750 kPa. 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 of a drop hammer 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 part may be rotated between blows for proper shaping of the workpiece during forging. Counterblow hammers operate at high speeds and transmit less vibration to their bases. Capacities range up to 1200 kJ.

High-energy-rate Forging Machines. In these machines, the ram is accelerated rapidly by inert gas at high pressure and the part is forged in one blow at a very high speed. Although there are several types of these machines, various problems associated with their operation and maintenance, as well as die breakage and safety considerations, have greatly limited their use in industry.

14.9 Economics of Forging

Several factors are involved in the cost of forgings. Depending on the complexity of the forging, tool and die costs range from moderate to high. However, as in other manufacturing operations, these costs are spread out over the number of parts forged with that particular die set. Thus, even though the cost of workpiece material per piece made is constant, setup and tooling costs per piece decrease as the number of pieces forged increases (Fig. 14.18).

The ratio of the cost of the die material to the total cost of forging the part increases with the weight of forgings: The more expensive the material, the higher the cost of the material relative to the total cost. 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.

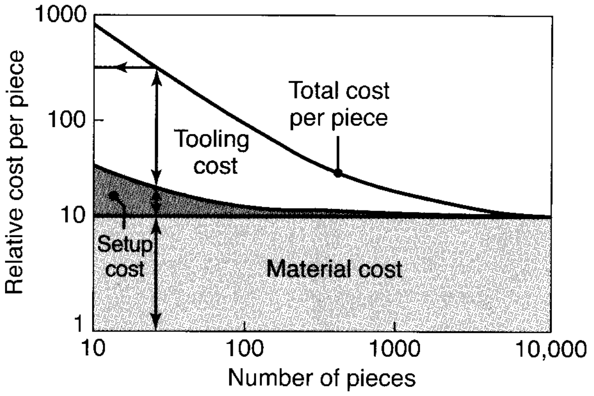


FIGURE 14.18 Typical cost per piece in forging; note how the setup and the tooling costs per piece decrease as the number of pieces forged increases if all pieces use the same die.

The size of forgings also has some effect on cost. Sizes range from small forgings (such as utensils and small automotive components) to large ones (such as gears, crankshafts, and connecting rods for large engines). As forging size increases, the share of material cost in the total cost also increases, but at a lower rate. This occurs 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 piece made is not that much higher.

The total cost involved in a forging operation 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 in forging generally are moderate. Also, die design and manufacturing are now performed by computer-aided design and manufacturing techniques (Chapter 38), which result in major savings in time and effort.

The cost of forging a part compared to that of making it by various casting techniques, powder metallurgy, machining, or other methods is an important consideration in a competitive global marketplace. For example, all other factors being the same, and depending on the number of pieces required, manufacturing a certain part by, say, expendable-mold casting may well be more economical than producing it by forging for shorter production runs (Fig. 14.19). This casting method does not require expensive molds and tooling, whereas forging requires expensive dies. Some competitive aspects of manufacturing and process selection are discussed in greater detail in Chapter 40.

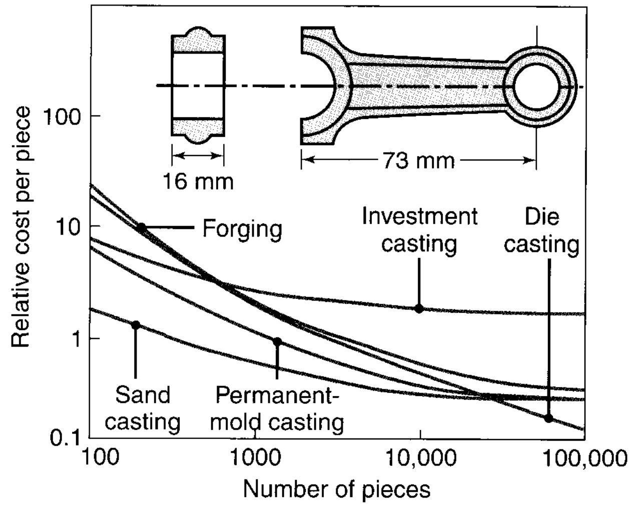


FIGURE 14.19 Relative unit costs of a small connecting rod made by various forging and casting processes. Note that, for large quantities, forging is more economical. Sand casting is the most economical process for fewer than about 20,000 pieces.

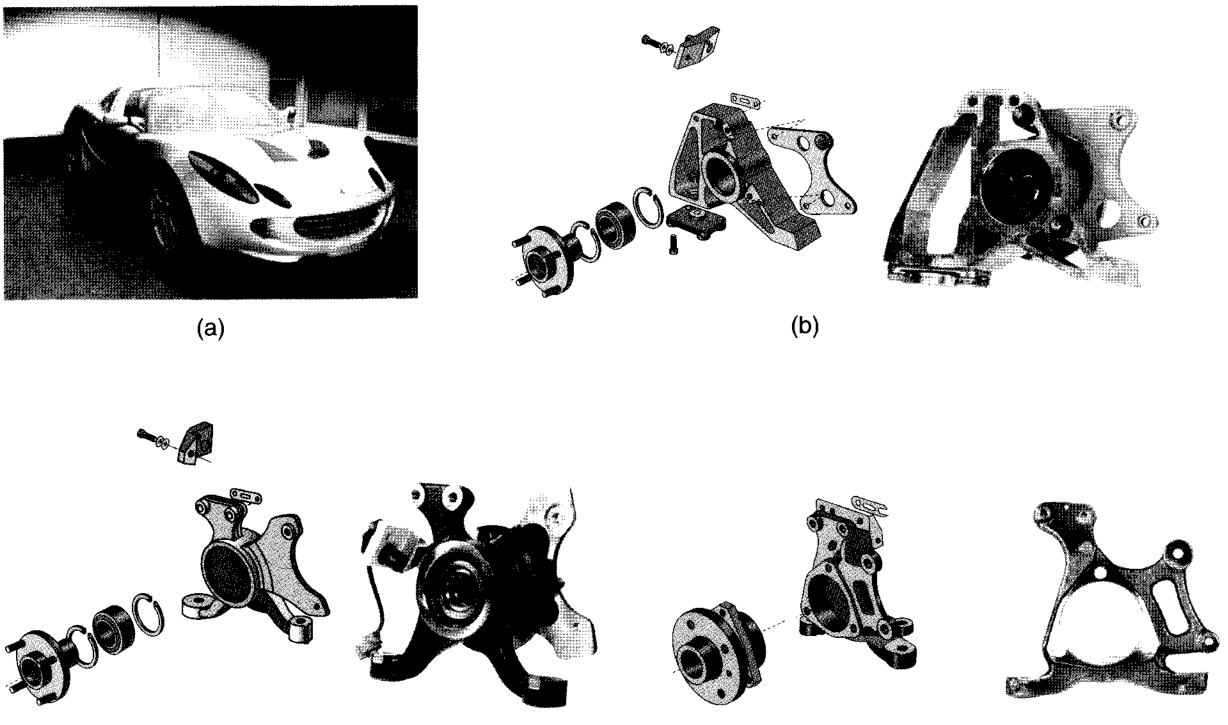
CASE STUDY 14.2 Suspension Components for the Lotus Elise Automobile

The automotive industry increasingly has been subjected to a demanding set of performance, cost, fuel efficiency, and environmental regulations. One of the main strategies in improving vehicle design with respect to all of these possibly conflicting constraints is to reduce vehicle weight while using advanced materials and manufacturing processes to preserve performance and safety. Previous design optimization has shown that weight savings of up to 34% can be realized on suspension system components—a significant savings, since suspensions make up approximately 12% of a car's mass. These weight savings could be achieved largely by developing optimum designs, utilizing advanced analytical tools, and using net-shape or near-net-shape steel forgings instead of cast-iron components. In addition, significant cost savings have been demonstrated in many parts when optimized steel forgings are used, as opposed to aluminum castings and extrusions.

The Lotus Elise is a high-performance sports car designed for superior ride and superior 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. 14.20. 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, which is used for similar applications.

The revised designs are summarized in Table 14.5. As can be seen, the optimized new forging


(a)

(b)

(c)

(d)

FIGURE 14.20 (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) Courtesy of Fox Valley Motorcars. (b) through (d) Courtesy of Lotus Engineering and the American Iron and Steel Institute.

TABLE 14.5

Comparison of Suspension Upright Designs for the Lotus Elise Automobile

Fig. 14.20 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%)

design (Fig. 14.20d) 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, and 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.

Source: Courtesy of Lotus Engineering and the American Iron and Steel Institute.

SUMMARY

- Forging denotes a family of metalworking processes in which deformation of the workpiece is carried out by compressive forces applied through a set of dies. Forging is capable of producing a wide variety of structural parts with favorable characteristics, such as higher strength, improved toughness, dimensional accuracy, and reliability in service.
- The forging process 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.
- Various defects can develop if the forging process is not designed or controlled properly. Defects appear especially in workpiece quality, billet or preform shape, and die geometry. Computer-aided design and manufacturing techniques are now 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 and characteristics. Forging operations are now highly automated and use 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, die design, material selection, and production method are of major 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, 2005.

ASM Handbook, Vol. 14A: **Metalworking: Bulk Forming**, ASM International, 2005.

Blazynski, T.Z., **Plasticity and Modern Metal-forming Technology,** Elsevier, 1989.

Byrer, T.G. (ed.), **Forging Handbook**, Forging Industry Association, 1985.

Dieter, G.E., Kuhn, H.A., and Semiatin, S.L. (eds.), **Handbook of Workability and Process Design,** ASM International, 2003.

Fundamentals of Tool Design, 5th ed., Society of Manufacturing Engineers, 2003.

Hosford, W.F., and Caddell, R.M., **Metal Forming: Mechanics and Metallurgy,** 3rd ed., Cambridge, 2007.

Lange, K. (ed.), **Handbook of Metal Forming,** McGraw-Hill, 1985.

Product Design Guide for Forging, Forging Industry Association, 1997.

Tool and Manufacturing Engineers Handbook, 4th ed., Vol. 2, *Forming*, Society of Manufacturing Engineers, 1984.

Tschaetch, H., **Metal Forming Practice: Processes, Machines,** *Tools*, Springer, 2007.

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?

QUALITATIVE PROBLEMS

14.11. How can you tell whether a certain part is forged or cast? Explain the features that you would investigate.

14.12. Identify casting design rules, described in Section 12.2, that also can be applied to forging.

14.13. Describe the factors involved in precision forging.

14.14. Why is control of the volume of the blank important in closed-die forging?

14.15. Why are there so many types of forging machines available? Describe the capabilities and limitations of each.

14.16. What are the advantages and limitations of (a) a cogging operation and (b) isothermal forging?

14.17. Describe your observations concerning Fig. 14.16.

14.18. What are the advantages and limitations of using die inserts? Give some examples.

14.19. Review Fig. 14.5d and explain why internal draft angles are larger than external draft angles. Is this also true for permanent-mold casting?

14.20. Comment on your observations regarding the grain-flow pattern in Fig. 14.12.

14.21. Describe your observations concerning the control of the final tube thickness in Fig. 14.15.

14.22. 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.

14.23. Describe the difficulties involved in defining the term "forgeability" precisely.

QUANTITATIVE PROBLEMS

14.24. 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.

▶ **14.25.** Calculate the forging force for a solid, cylindrical workpiece made of 1020 steel that is 90 mm high and 125 mm in diameter and is to be reduced in height by 30%. Let the coefficient of friction be 0.15.

▶ **14.26.** Using Eq. (14.2), estimate the forging force for the workpiece in Problem 14.25, 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.

▶ **14.27.** To what thickness can a cylinder of 5052-O aluminum that is 100 mm in diameter and 25 mm high be forged in a press that can generate 450 kN?

▶ **14.28.** 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.

▶ **14.29.** 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, Y_f, of the material is constant. Remember that the volume of the material remains constant during forging; thus, as h decreases, r increases.

14.30. 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 320 mm^2. Explain clearly. (*Hint*: See Section 2.6 on hardness.)

14.31. A mechanical press is powered by a 23-kW motor and operates at 40 strokes per minute. It uses a flywheel, so that the crankshaft speed does not vary appreciably during the stroke. If the stroke is 150 mm, what is the maximum constant force that can be exerted over the entire stroke length?

14.32. 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

14.33. Devise an experimental method whereby you can measure only the force required for forging the flash in impression-die forging.

14.34. Assume that you represent the forging industry and that you are facing a representative of the casting industry. What would you tell that person about the merits of forging processes?

14.35. Figure P14.35 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.

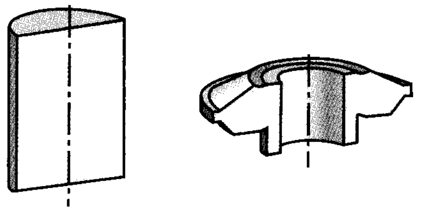


FIGURE P14.35

14.36. 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.

▶ **14.37.** 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 60-ton capacity as a function of the material.

14.38. Review the sequence of operations in the production of the stepped pin shown in Fig. 14.13. If the conical-upsetting step is not performed, how would the final part be affected?

14.39. 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).

14.40. Discuss the possible environmental concerns regarding the operations described in this chapter.

CHAPTER

15 Metal Extrusion and Drawing Processes and Equipment

- Extrusion and drawing involve, respectively, pushing or pulling a material through a die basically for the purpose of reducing or changing its cross-sectional area.
- This chapter examines the fundamentals of these processes and their applications.
- The chapter starts by discussing the basic types of extrusion processes, namely, direct, indirect, and hydrostatic extrusion, and explains how the extrusion force can be estimated from material and processing parameters.
- Hot and cold extrusion are then discussed; cold extrusion is often done in combination with forging to produce specific parts.
- Extrusion practices and die designs that avoid common defects are also presented.
- The drawing of rod, wire, and tubing is then examined in a similar manner, along with die design.
- The 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, and continuous casting.

15.1 Introduction

Extrusion and drawing have numerous applications in the manufacture of continuous as well as discrete products from a wide variety of metals and alloys. Plastics also are extruded extensively, as described in Section 19.2. In **extrusion**, a cylindrical billet is forced through a die (Fig. 15.1) in a manner similar to squeezing toothpaste from a tube or extruding Play-Doh® in various cross sections in a toy press. A wide variety of solid or hollow cross sections may be produced by extrusion, which essentially are semifinished parts. A characteristic of extrusion (from the Latin *extrudere*, meaning "to force out") is that large deformations can take place without fracture (see Section 2.2.8), because the material is under high triaxial compression. Since the die geometry remains unchanged throughout the operation, extruded products typically have a constant cross section.

Typical products made by extrusion are railings for sliding doors, window frames, tubing having various cross sections, aluminum ladder frames, and numerous structural and architectural shapes. Extrusions can be cut into desired lengths, which then become discrete parts, such as brackets, 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, with various levels of difficulty.

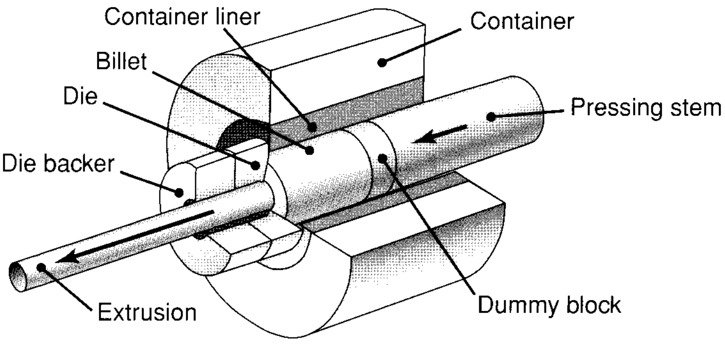


FIGURE 15.1 Schematic illustration of the direct-extrusion process.

Because a chamber is involved, each billet is extruded individually; thus, extrusion is a batch, or semicontinuous, process. Extrusion 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, extrusion is carried out at room or elevated temperatures. Extrusion at room temperature often is combined with forging operations, in which case it generally is known as **cold extrusion** (see also Section 14.4). It has numerous important applications, including fasteners and components for automobiles, bicycles, motorcycles, heavy machinery, and transportation equipment.

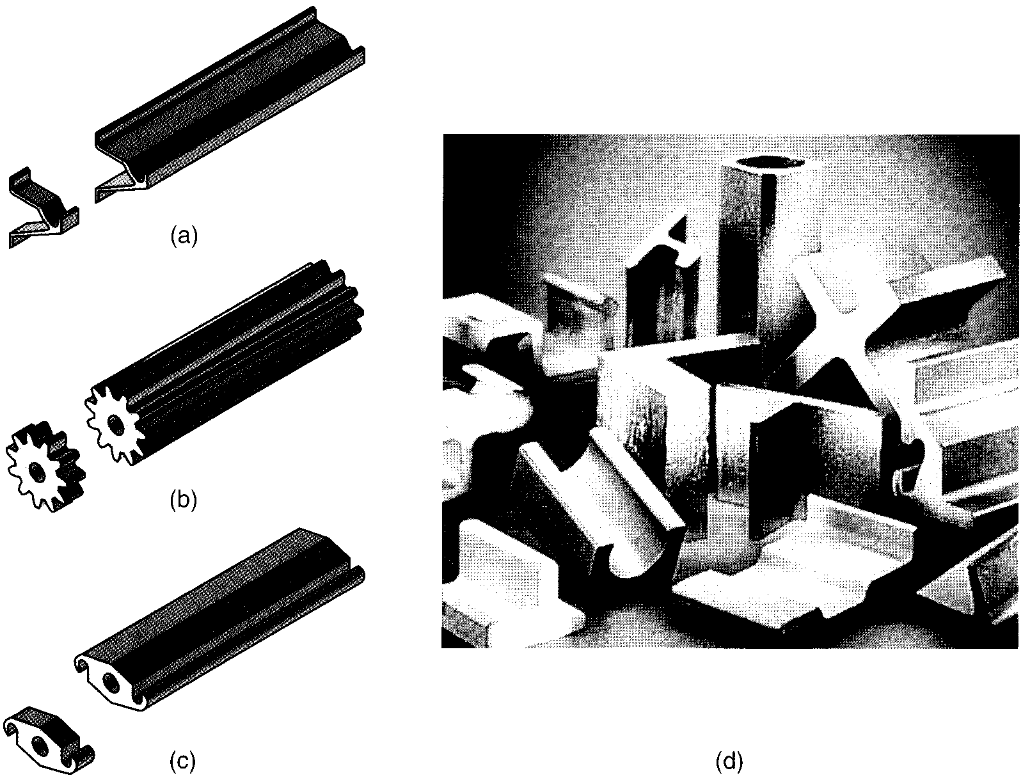


FIGURE 15.2 Extrusions and examples of products made by sectioning off extrusions. *Source*: Courtesy of Plymouth Extruded Shapes.

In **drawing**, an operation that was developed between 1000 and 1500 A.D., the cross section of 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 can be drawn. The term *drawing* also is used to refer to making cup-shaped parts by sheet-metal-forming operations, as described in Section 16.7.

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 generally is defined as a rod that has been drawn through a die at least once, or its diameter is small enough so that it can be coiled. Wire drawing involves smaller diameters than rod drawing, with sizes down to 0.01 mm for magnet wire and even smaller for use in very low current fuses.

15.2 The Extrusion Process

There are three basic types of extrusion. In the most common process (called **direct** or **forward extrusion**), a billet is placed in a *chamber* (container) and forced through a die opening by a hydraulically driven ram (pressing stem or punch), as shown in Fig. 15.1. The die opening may be round, or it may have various shapes, depending on the desired profile. The function of the dummy block shown in the figure is to protect the tip of the pressing stem (punch), particularly in hot extrusion. Other types of extrusion are indirect, hydrostatic, and impact extrusion.

In **indirect** extrusion (also called reverse, inverted, or backward extrusion), the die moves toward the unextruded billet (Fig. 15.3a). Indirect extrusion has the advantage of having no billet–container friction, since there is no relative motion. Thus, indirect extrusion 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 chamber (which is filled with a fluid), and the pressure is transmitted to the fluid by a ram. The fluid pressure results in triaxial compressive stresses acting on the workpiece and thus improved formability; also, 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).

As can be seen in Fig. 15.4, the geometric variables in extrusion are the die angle, α, and the ratio of the cross-sectional area of the billet to that of the extruded product, A_0/A_f, called the **extrusion ratio**, R. Other variables are the temperature of the billet, the speed at which the ram travels, and the type of lubricant used.

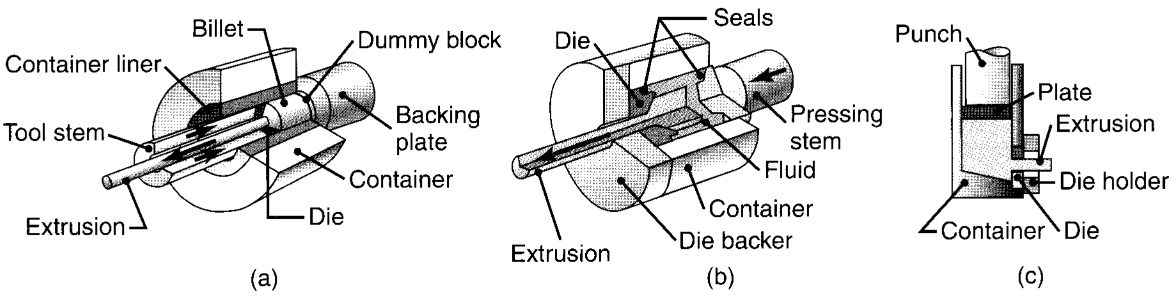


FIGURE 15.3 Types of extrusion: (a) indirect; (b) hydrostatic; (c) lateral.

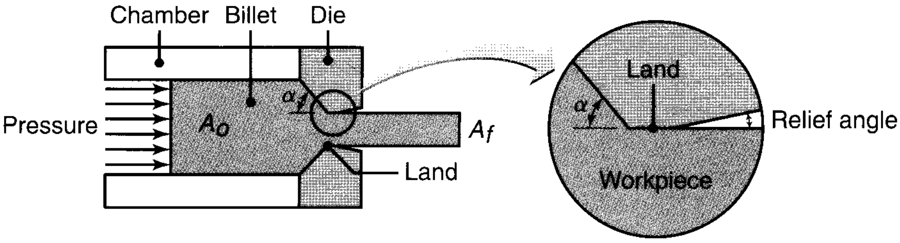


FIGURE 15.4 Process variables in direct extrusion. The die angle, reduction in cross section, extrusion speed, billet temperature, and lubrication all affect the extrusion pressure.

Extrusion Force. The force required for extrusion depends on (a) the strength of the billet material, (b) the extrusion ratio, (c) the friction between the billet and the chamber and die surfaces, and (d) process variables, such as the temperature of the billet and the speed of extrusion. The *extrusion force, F*, can be estimated from the formula

$$F = A_o k \ln\left(\frac{A_o}{A_f}\right), \qquad (15.1)$$

where k is the **extrusion constant** (which is determined experimentally) and A_o and A_f are the billet and extruded product areas, respectively. The k value in Eq. (15.1) is thus a measure of the strength of the material being extruded and the frictional conditions. Figure 15.5 gives the k values of several metals for a range of extrusion temperatures.

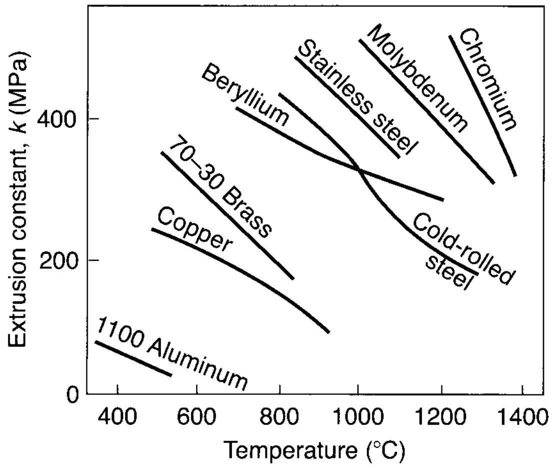


FIGURE 15.5 Extrusion constant k for various metals at different temperatures. *Source*: After P. Loewenstein.

EXAMPLE 15.1 Calculation of Force in Hot Extrusion

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. Calculate the extrusion force required.

Solution The extrusion force is calculated using Eq. (15.1), 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,

$$F = \frac{\pi(125)^2}{4}(250)\ln\left[\frac{\pi(125)^2}{\pi(50)^2}\right] = 4$$
$$= 5.6 \text{ MN}.$$

(See Section 15.6 for capacities of extrusion presses.)

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 mechanical properties of the extruded product. The material flows longitudinally, much like incompressible fluid flows in a channel; thus, extruded products have an elongated grain structure (preferred orientation). Section 15.5 describes how improper metal flow during extrusion can produce various defects in the extruded product.

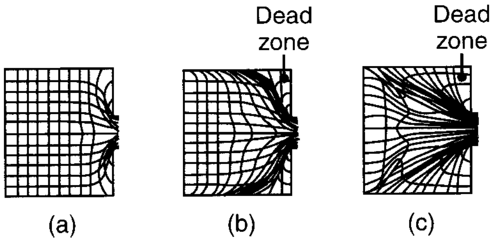


FIGURE 15.6 Types of metal flow in extruding with square dies. (a) Flow pattern obtained at low friction or in indirect extrusion. (b) Pattern obtained with high friction at the billet–chamber interfaces. (c) Pattern obtained at high friction or with cooling of the outer regions of the billet in the chamber. This type of pattern, observed in metals whose strength increases rapidly with decreasing temperature, leads to a defect known as pipe (or extrusion) defect.

A common technique for investigating the flow pattern is to cut the round billet in half lengthwise and then mark one face with a square grid pattern. The two halves may be brazed together, after which they are placed in the chamber together and are extruded. The two pieces are then taken apart (by melting the braze, if used, in a furnace) and studied. Figure 15.6 shows typical flow patterns obtained by this technique for the case of direct extrusion with *square dies* (a 90° die angle). The flow pattern is a function of several variables, including friction.

The conditions under which these different flow patterns occur are described in the caption of Fig. 15.6. Note the **dead-metal zones** in Figs. 15.6b and c, where the metal at the corners essentially is stationary. This situation is similar to the stagnation of fluid flow in channels that have sharp angles or turns.

Process Parameters. Because they have high ductility, wrought aluminum, copper, and magnesium and their alloys, as well as steels and stainless steels, are extruded with relative ease into numerous shapes. Other metals (such as titanium and refractory metals) also can be extruded, but only with some difficulty and considerable die wear.

In practice, extrusion ratios, R, usually range from about 10 to 100. They may be higher for special applications (400 for softer nonferrous metals) or lower for less ductile materials, although the ratio usually has to be at least 4 to deform the material plastically through the bulk of the workpiece. Extruded products usually are less than 7.5 m long because of the difficulty in handling greater lengths, but 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, higher speeds for steels, titanium, and refractory alloys. Dimensional tolerances in extrusion are usually in the range from ± 0.25 to 2.5 mm, and they increase with increasing cross section.

Most extruded products—particularly those with small cross sections—require straightening and twisting. This is accomplished typically by stretching and twisting the extruded product, usually in a hydraulic stretcher equipped with jaws. The presence of a die angle causes a small portion of the end of the billet to remain in the chamber after the operation has been completed. This portion (called scrap or the *butt end*) subsequently is removed by cutting off the extrusion at the die exit and removing the scrap from the chamber. Alternatively, another billet or a graphite block may be placed in the chamber to extrude the piece remaining from the previous extrusion.

In **coaxial extrusion**, or **cladding**, coaxial billets are extruded together—provided that the strength and ductility of the two metals are compatible. An example is copper clad with silver. *Stepped extrusions* are produced by extruding the billet partially in one die and then in one or more 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 hot-working operations, hot extrusion has special requirements because of the high operating temperatures. For example, die wear can

TABLE 15.1

Typical Extrusion Temperature Ranges for Various Metals and Alloys

	°C
Lead	200–250
Aluminum and its alloys	375–475
Copper and its alloys	650–975
Steels	875–1300
Refractory alloys	975–2200

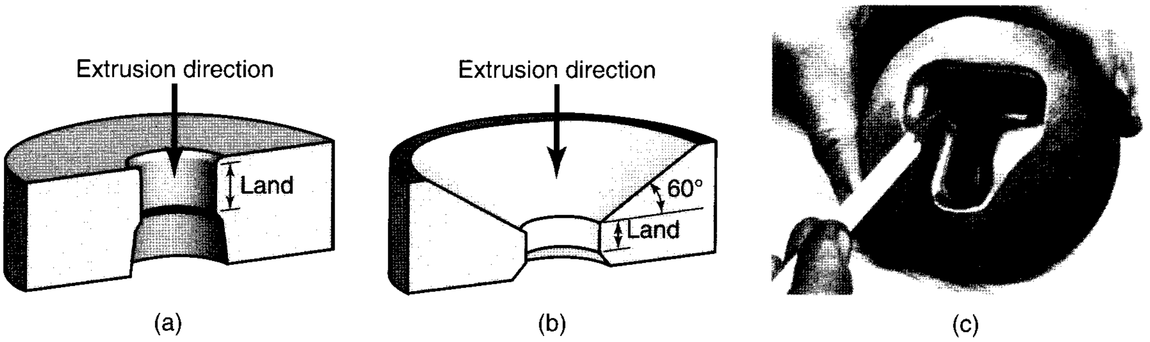


FIGURE 15.7 Typical extrusion–die configurations: (a) die for nonferrous metals; (b) die for ferrous metals; (c) die for a T-shaped extrusion made of hot-work die steel and used with molten glass as a lubricant. *Source*: (c) Courtesy of LTV Steel Company.

be excessive, and cooling of the surfaces of the hot billet (in the cooler chamber) and the die can result in highly nonuniform deformation (Fig. 15.6c). To reduce cooling of the billet and to prolong die life, extrusion dies may be preheated, as is done in hot-forging operations.

Because the billet is hot, it develops an oxide film, unless it is heated in an inert-atmosphere furnace. Oxide films 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 important. In order to avoid the formation 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. The skull is removed later from the chamber.

Die Design. Die design requires considerable experience, as can be appreciated by reviewing Fig. 15.7. *Square dies* (*shear dies*) are used in extruding nonferrous metals, especially aluminum. Square dies develop dead-metal zones, which in turn form a "die angle" (see Figs. 15.6b and c) along which the material flows in the deformation zone. The dead-metal zones produce extrusions with bright finishes because of the burnishing that takes places as the material flows past the "die angle" surface.

Tubing is 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.

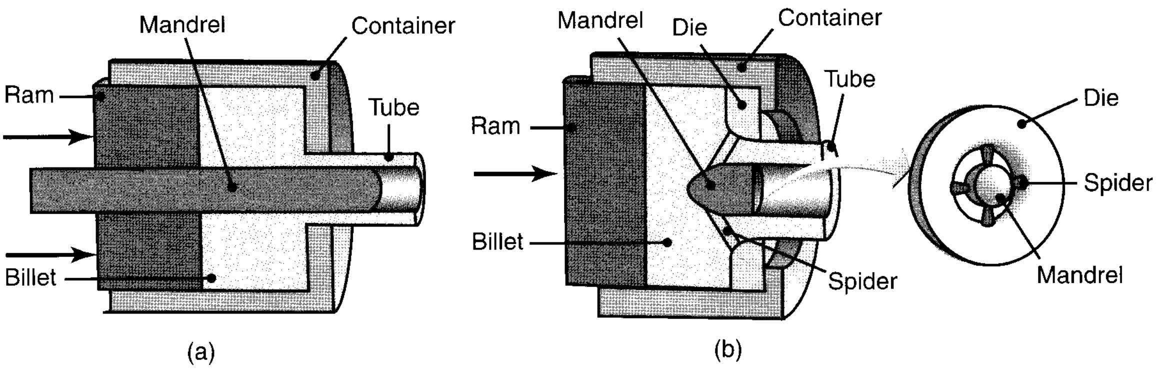
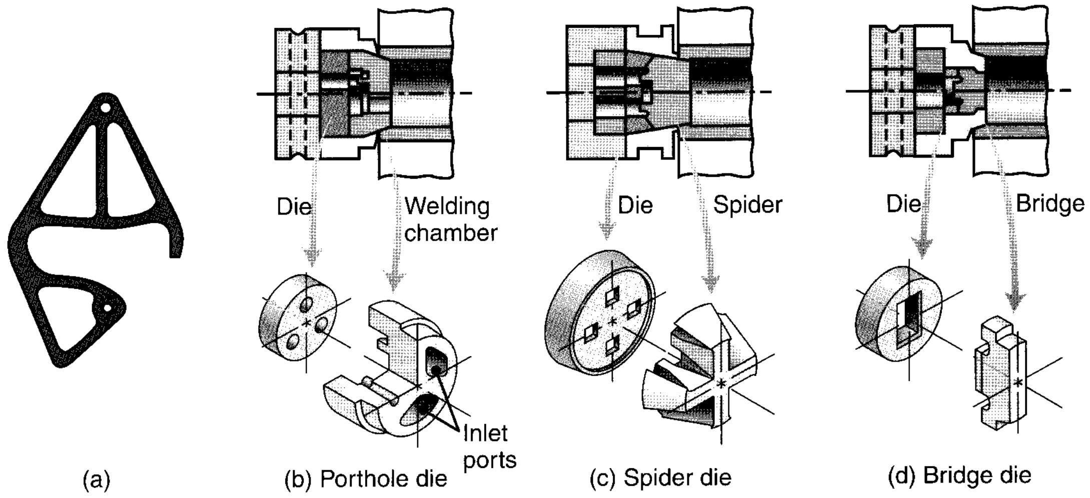


FIGURE 15.8 Extrusion of a seamless tube (a) using an internal mandrel that moves independently of the ram. (An alternative arrangement has the mandrel integral with the ram.) (b) using a spider die (see Fig. 15.9) to produce seamless tubing.



(a) (b) Porthole die (c) Spider die (d) Bridge die

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.

Hollow cross sections (Fig. 15.9a) can be extruded by welding-chamber methods and using various dies known as a **porthole die, spider die,** and **bridge die** (Figs. 15.9b to d). During extrusion, the metal divides and flows around the supports for the internal mandrel into strands. (This condition is much like that of air flowing around a moving car and rejoining downstream or water flowing around large rocks in a river and rejoining.) The strands then become rewelded under the high pressure in the welding chamber before they exit through the die. The rewelded surfaces have good strength because they have not been exposed to

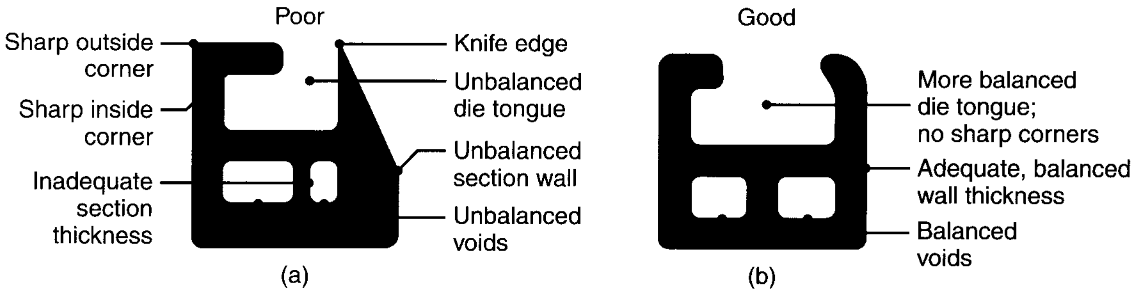


FIGURE 15.10 Poor and good examples of cross sections to be extruded. Note the importance of eliminating sharp corners and of keeping section thicknesses uniform.
Source: J.G. Bralla (ed.), *Handbook of Product Design for Manufacturing*. McGraw-Hill Publishing Company, 1986. Used with permission.

the environment; otherwise they would develop oxides on the surfaces, thereby inhibiting good welding. However, the welding-chamber process is suitable only for aluminum and some of its alloys, because of their capacity for developing a strong weld under pressure (as is described in Section 31.2). Lubricants, of course, cannot be used, because they prevent rewelding of the metal in the die.

Some guidelines for proper die design in extrusion are illustrated in Fig. 15.10. Note the (a) importance of symmetry of cross section, (b) avoidance of sharp corners, and (c) avoidance of extreme changes in die dimensions within the cross section.

Die Materials. Die materials for hot extrusion usually are hot-worked die steels (Section 5.7). Coatings (such as partially stabilized zirconia) may be applied to the dies to extend their life. Partially stabilized zirconia dies (Section 8.2.2) also are used for hot extrusion of tubes and rods. However, they are not suitable for making dies for extruding complex shapes, because of the severe stress gradients developed in the die, which may lead 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 steels, stainless steels, and high-temperature metals and alloys. In a process developed in the 1940s and known as the **Séjournet process** (after J. Séjournet), a circular glass or fiberglass pad is placed in the chamber at the die entrance. The hot billet conducts heat to the pad, whereupon a thin layer of glass begins to melt and acts as a lubricant at the die interface as the extrusion progresses. Before the billet is 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.

For metals that have a tendency to stick to the container and the die, the billet can be enclosed in a thin-walled container, or jacket, 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. Also, if the billet material is toxic or radioactive, the jacket prevents it from contaminating the environment. This technique also can be used for extruding reactive metal powders (Chapter 17).

EXAMPLE 15.2 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 economically as does aluminum.

Hot extrusion of aluminum is preferred for heat-sink applications, such as those in the electronics industry. Figure 15.11a shows an extruded heat sink used to remove heat from a transformer on a printed circuit board. Heat sinks usually are designed with a large number of fins that are intended to maximize the surface area and are evaluated from a thermodynamics standpoint using computer simulations. The fins are very difficult and expensive to machine, forge, or roll form. However, the tooling for hot extrusion can be produced through electrical-discharge machining (Section 27.5), so the process is favorable economically.

Figure 15.11b shows a die and a hot-extruded cross section suitable to serve as a heat sink. The shapes shown could be produced through a casting operation, but extrusion is preferred, since there is no internal porosity and the thermal conductivity is slightly higher.

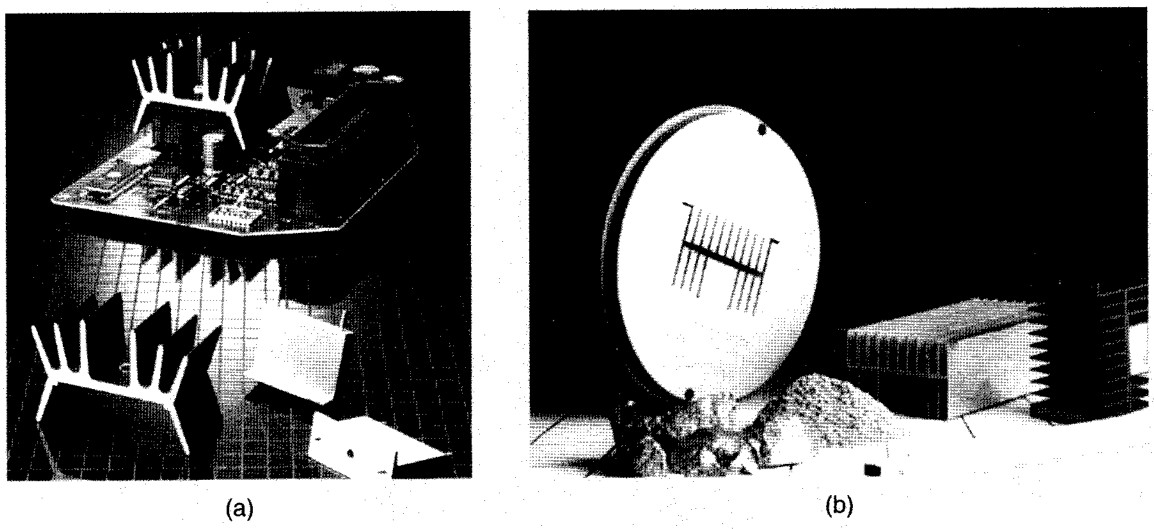

(a) (b)

FIGURE 15.11 (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.

15.4 Cold Extrusion

Developed in the 1940s, *cold extrusion* is a general term that often denotes a *combination* of operations, such as direct and indirect *extrusion and forging* (Fig. 15.12). Cold extrusion is used widely for components in automobiles, motorcycles, bicycles, and appliances and in transportation and farm equipment.

The cold-extrusion process uses slugs cut 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.

The *force, F*, in cold extrusion may be estimated from the formula

$$F = 1100 A_o Y_{avg} \epsilon, \qquad (15.2)$$

where A_o is the cross-sectional area of the blank, Y_{avg} 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; i.e., $\ln(A_o/A_f)$. For example, assume that a round slug 10 mm in diameter and made of a metal with $Y_{avg} = 350$ MPa is reduced to a final diameter of 7 mm by cold extrusion. Then the force would be

$$F = 1100(\pi)(10^2/4)(350)[\ln(10/7)^2] = 21.6 \text{ MN}.$$

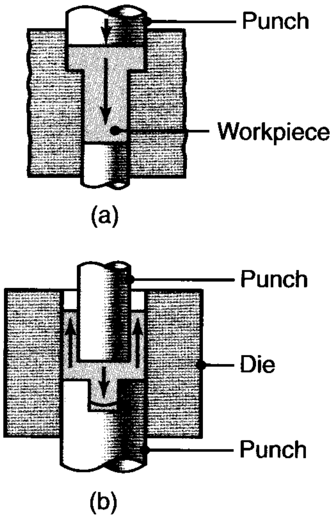


FIGURE 15.12 Two examples of cold extrusion. Thin arrows indicate the direction of metal flow during extrusion.

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, 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 that are competitive with those of other methods of producing the same part, such as machining. Some machines are capable of producing more than 2000 parts per hour.

The magnitude of the stresses on the tooling in cold extrusion, however, is very high (especially with steel and specialty-alloy workpieces), 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, as they must possess not only sufficient strength, but also sufficient toughness and resistance to wear and fatigue failure.

Lubrication is critical, especially with steels, because of the possibility of sticking (*seizure*) between the workpiece and the tooling (in the case of lubricant breakdown). The most effective means of lubrication is the application of a *phosphate-conversion coating* on the workpiece, 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 crucial to the success of cold extrusion. Also important are the selection and control of the workpiece material with regard to its quality, the accuracy of the slug dimensions, and its surface condition. Several specialty alloys have been developed (particularly for critical applications requiring high performance) that are suitable for a variety of cold-extrusion and cold-forming operations with good properties, dimensional tolerances, and at low cost.

EXAMPLE 15.3 Cold-extruded Part

A typical cold-extruded product, similar to the metal component of an automotive spark plug, is shown in Fig. 15.13. First, a slug is sheared off the end of a round rod (Fig. 15.13, left). It then is cold extruded (Fig. 15.13, middle) in an operation similar to those shown in Fig. 15.12 but with a blind hole. Then the material at the bottom of the blind hole 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. Furthermore, the part usually is sectioned in the midplane and then polished and etched to display the grain flow, as shown in Fig. 15.14 (see also Fig. 14.11).

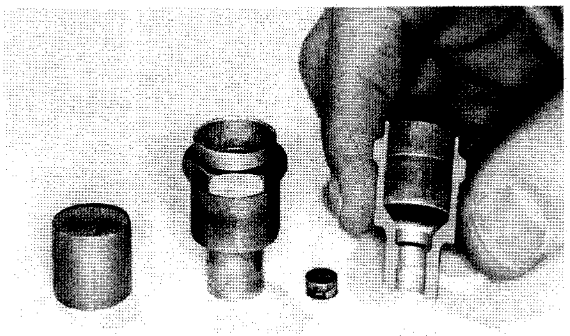


FIGURE 15.13 Production steps for a cold-extruded spark plug. *Source*: Courtesy of National Machinery Company.

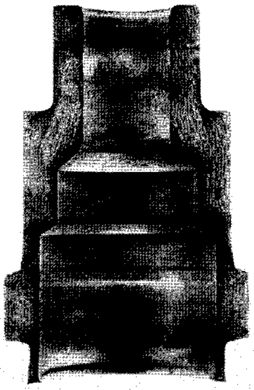


FIGURE 15.14 A cross section of the metal part in Fig. 15.13, showing the grain-flow pattern. *Source*: Courtesy of National Machinery Company.

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. 15.15). Because of volume constancy, the thickness of the tubular extruded section is a function of the clearance between the punch and the die cavity.

Typical products made by this process are shown in Figs. 15.16a to c. Other examples of products made by impact extrusion are collapsible tubes (similar to those used for toothpaste), 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.

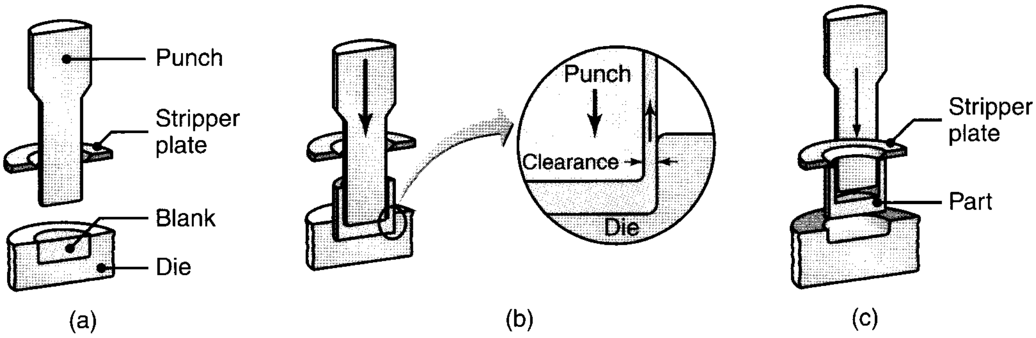


FIGURE 15.15 Schematic illustration of the impact-extrusion process. The extruded parts are stripped by the use of a stripper plate, because they tend to stick to the punch.

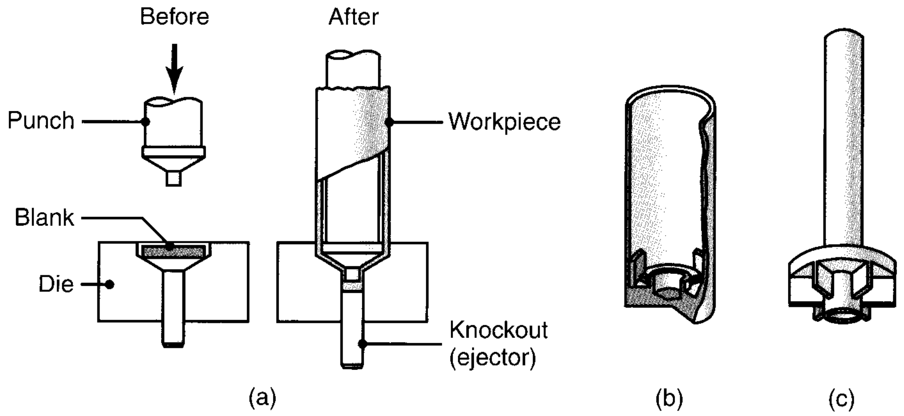


FIGURE 15.16 (a) Impact extrusion of a collapsible tube by the *Hooker process.* (b) and (c) Two examples of products made by impact extrusion. These parts also may be made by casting, forging, or machining. The choice of process depends on the materials involved, part dimensions and wall thickness, and the properties desired. Economic considerations also are important in final process selection.

The maximum diameter of the parts made is about 150 mm. The impact-extrusion process can produce thin-walled tubular sections having thickness-to-diameter ratios as small as 0.005. Consequently, the symmetry of the part and the concentricity of the punch and the blank are important.

15.4.2 Hydrostatic Extrusion

In *hydrostatic extrusion*, the pressure required in the chamber is supplied via a piston 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 usually is carried out at room temperature, typically using vegetable oils as the fluid (particularly castor oil, because it is a good lubricant and its viscosity is not influenced significantly by pressure).

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. Long wires also have been extruded from an aluminum billet at room temperature and at an extrusion ratio of 14,000, which means that a 1-m billet becomes a 14-km-long wire. In spite of the success obtained, hydrostatic extrusion has had limited industrial applications, mainly because of the somewhat complex nature of the tooling, the experience needed with high pressures and the design of specialized equipment, and the long cycle times required—all of which make 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 the naked eye, while others can be detected only by the techniques described in Section 36.10. There are three principal *extrusion defects:* surface cracking, pipe, and internal cracking.

Surface Cracking. If extrusion temperature, friction, or speed is too high, surface temperatures can rise significantly, which may cause surface cracking and tearing (*fir-tree cracking* or *speed cracking*). These cracks are intergranular (i.e., along the grain boundaries; see Fig. 2.27) and usually are caused by **hot shortness** (Section 1.5.2). These defects occur especially in aluminum, magnesium, and zinc alloys, although they may also occur in high-temperature alloys. This situation can be avoided by lowering the billet temperature and the extrusion speed.

Surface cracking also may occur at lower temperatures, where it has been attributed to periodic sticking of the extruded product along the die land. Because of the similarity in appearance to the surface of a bamboo stem, it is known as a **bamboo defect**. When the product being extruded temporarily sticks to the *die land* (see Fig. 15.7), the extrusion pressure increases rapidly. Shortly thereafter, the product 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 thus 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. Another method is to machine the billet's surface prior to extrusion, so that scale and surface impurities are removed. These impurities also can be removed by the chemical etching of the surface oxides prior to extrusion.

Internal Cracking. The center of the extruded product can develop cracks, called *center cracking*, *center-burst*, *arrowhead fracture*, or *chevron cracking* (Fig. 15.17a). These cracks are attributed to a state of hydrostatic tensile stress at the centerline in the deformation zone in the die (Fig. 15.17b), a situation similar to the necked region in a tensile-test specimen (see Fig. 2.23). These cracks also have been observed in tube extrusion and in tube spinning (see Figs. 16.46b and c); they appear on the inside

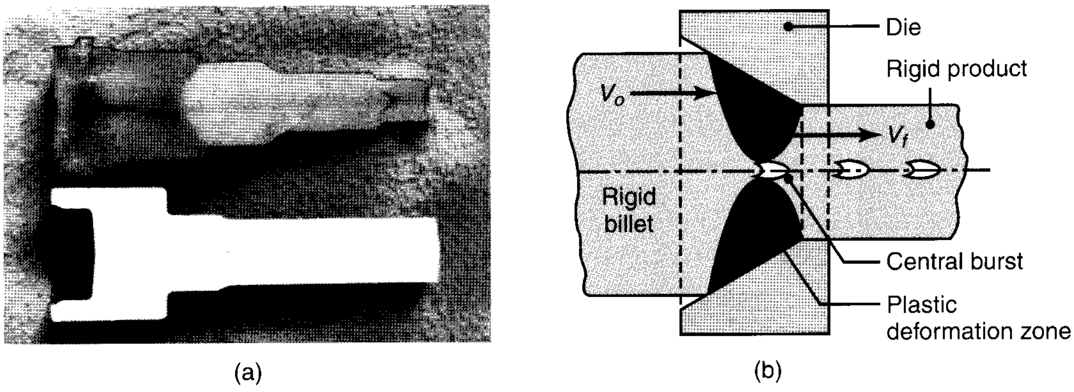


FIGURE 15.17 (a) Chevron cracking (central burst) in extruded round steel bars. Unless the products are inspected, such internal defects may remain undetected and later cause failure of the part in service. This defect can also develop in the drawing of rod, of wire, and of tubes. (b) Schematic illustration 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, by increasing the reduction in cross section, or both. *Source*: After B. Avitzur.

surfaces of tubes. The tendency for center cracking (a) increases with increasing die angle, (b) increases with increasing amount of impurities, and (c) decreases with increasing extrusion ratio and friction.

15.6 Extrusion Equipment

The basic equipment for extrusion is a *horizontal hydraulic press* (Fig. 15.18; see also Fig. 14.17d). These presses are suitable for extrusion because the stroke and speed of the operation can be controlled, depending on the particular application. They are capable of applying a constant force over a long stroke. Consequently, long billets can be used, correspondingly larger extrusions produced per setup, and the production rate thus increased. Hydraulic presses with a ram-force capacity as high as 120 MN have been built, particularly for hot extrusion of large-diameter billets.

Vertical hydraulic presses typically are used for cold extrusion. They generally have less capacity than those used for hot extrusion, but they take up less floor space. In addition to such presses, *crank-joint* and *knuckle-joint* mechanical presses 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 operations, are carried out on specially designed presses.

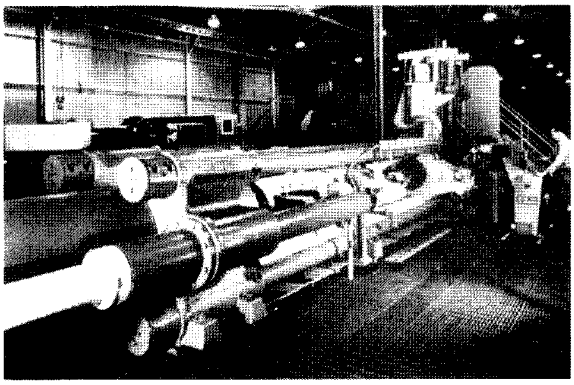


FIGURE 15.18 General view of a 9-MN hydraulic-extrusion press. *Source*: Courtesy of Jones & Laughlin Steel Corporation.

15.7 The Drawing Process

In *drawing*, the cross section of a long rod or wire is reduced or changed by pulling (hence the term drawing) it through a die called a *draw die* (Fig. 15.19). Thus, the difference between drawing and extrusion is that in extrusion the material is pushed through a die, whereas in drawing it is pulled through it. 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.

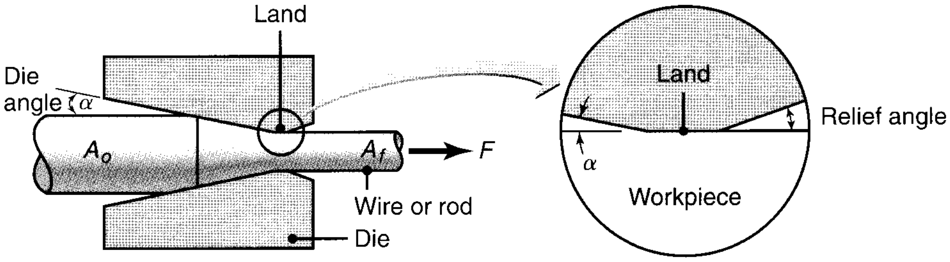


FIGURE 15.19 Process variables in wire drawing. The die angle, the reduction in cross-sectional area per pass, the speed of drawing, the temperature, and the lubrication all affect the drawing force, *F*.

The major processing variables in drawing are similar to those in extrusion—that is, reduction in cross-sectional area, die angle, friction along the die–workpiece interface, and drawing speed. The die angle influences the drawing force and the quality of the drawn product.

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

$$F = Y_{\text{avg}} A_f \ln\left(\frac{A_o}{A_f}\right), \tag{15.3}$$

where Y_{avg} is the average true stress of the material in the die gap. Because more work has to be done to overcome friction, the force increases with increasing friction. Furthermore, because of the nonuniform deformation that occurs within the die zone, additional energy (known as the *redundant work of deformation*) is required. Although various equations have been developed to estimate the force (described in greater detail in advanced texts), a useful formula that includes friction and the redundant work is

$$F = Y_{\text{avg}} A_f \left[\left(1 + \frac{\mu}{\alpha}\right)\ln\left(\frac{A_o}{A_f}\right) + \frac{2}{3}\alpha\right], \tag{15.4}$$

where α is the die angle in radians.

As can be seen from these equations, the drawing force increases as reduction increases. However, there has to be a limit to the magnitude of the force, because when the tensile stress reaches the yield stress of the metal being drawn, the workpiece will simply yield and, eventually, break. It can be shown that, ideally and without friction, the maximum reduction in cross-sectional area per pass is 63%. Thus, a 10-mm-diameter rod can be reduced (at most) to a diameter of 6.1 mm in one pass without failure.

It can be shown that, for a certain reduction in diameter and a certain 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 the actual die angle may deviate from this value.

Drawing of Other Shapes. Various solid cross sections can be produced by drawing through dies with different profiles. Proper die design and the proper selection of

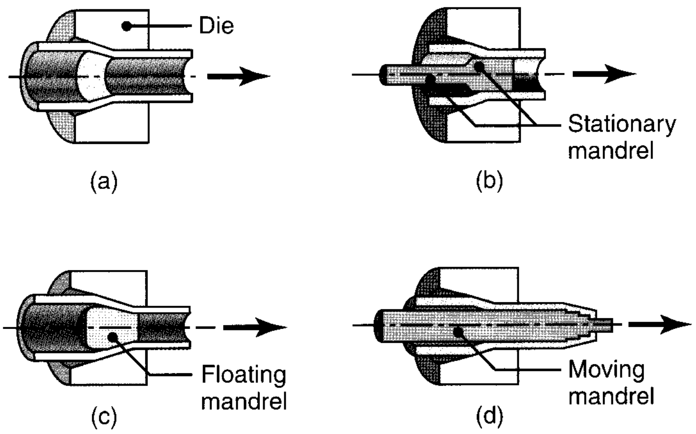


FIGURE 15.20 Examples of tube-drawing operations, with and without an internal mandrel. Note that a variety of diameters and wall thicknesses can be produced from the same initial tube stock (which has been made by other processes).

reduction sequence per pass require considerable experience to ensure proper material flow in the die, reduce 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 reduced further by *tube-drawing* processes (Fig. 15.20). Tubes as large as 0.3 m in diameter can be drawn by these techniques. Mandrels of various profiles are available for tube-drawing operations.

Wedge-shaped dies are used for the drawing of flat strips and are used only in specific applications. However, the principle behind this process is the fundamental deformation mechanism in **ironing**, used extensively in making aluminum beverage cans, as shown in Fig. 16.30.

15.8 Drawing Practice

As in all metalworking processes, successful drawing requires careful selection of process parameters. In drawing, reductions in the cross-sectional area per pass range up to about 45%. Usually, the smaller the initial cross section, the smaller the reduction per pass. Fine wires usually are drawn at 15 to 25% reduction per pass and larger sizes at 20 to 45%. Reductions of higher than 45% may result in lubricant breakdown, leading to surface-finish deterioration. 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 (**sizing pass**) also may be taken on rods to improve their surface finish and dimensional accuracy. However, because they basically deform only the surface layers, light reductions usually produce highly nonuniform deformation of the material and its microstructure. Consequently, the properties of the material will vary with location 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 and be pulled. This typically is done by *swaging* the tip of the rod or wire in a manner similar to that shown in Figs. 14.15a and b; this operation is called *pointing*. Drawing speeds depend on the material and on 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, temperatures can rise substantially at high drawing speeds and can have detrimental effects on product quality.

Drawn copper and brass wires are designated by their *temper* (such as 1/4 hard, 1/2 hard, etc.) because of work hardening. Intermediate annealing between passes may be necessary to maintain sufficient ductility of the material during cold drawing. High-carbon steel wires for springs and for musical instruments are made by heat treating (**patenting**) the drawn wire; the microstructure obtained is fine pearlite (see Fig. 4.11). These wires have ultimate tensile strengths as high as 5 GPa and a tensile reduction of area of about 20%.

Bundle Drawing. Although very fine wire can be produced by drawing, the cost can be high. One method employed to increase productivity is to draw many wires (a hundred or more) simultaneously as a *bundle*. The wires are separated from one another by a suitable metallic material with similar properties, but lower chemical resistance (so that it subsequently can be leached out from the drawn-wire surfaces).

Bundle drawing produces wires that are somewhat polygonal, rather than round, in cross section. In addition to producing continuous lengths, techniques have been developed to produce fine wire that is 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. The wires produced can be as small as 4 μm in diameter and can be made from such materials as stainless steels, titanium, and high-temperature alloys.

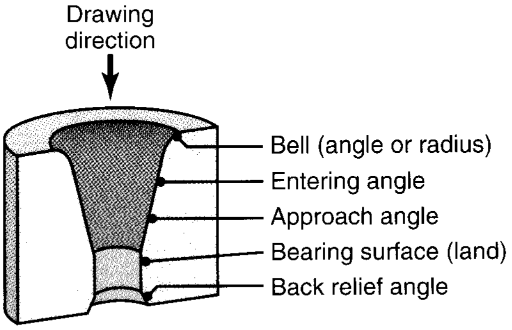


FIGURE 15.21 Terminology pertaining to a typical die used for drawing a round rod or wire.

Die Design. The characteristic features of a typical die for drawing are shown in Fig. 15.21. Die angles usually range from 6° to 15°. Note, however, that there are two angles (entering and approach) in a typical die. The purpose of the bearing surface (*land*) is to set the final diameter of the product (*sizing*) and to maintain this diameter even with wear on the die–workpiece interface.

A set of dies is required for **profile drawing**, which involves 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 held together in a retaining ring. Computer-aided design techniques are being implemented to design dies for smooth material flow, as well as to minimize defects. 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 that in ordinary draw dies, because the rolls can be adjusted to different positions and angles for specific profiles.

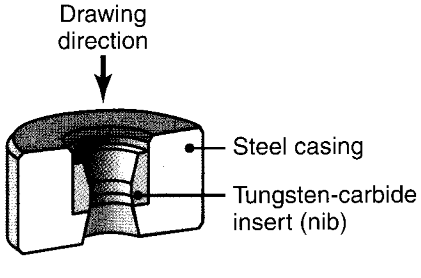


FIGURE 15.22 Tungsten-carbide die insert in a steel casing. Diamond dies used in drawing thin wire are encased in a similar manner.

Die Materials. Die materials for drawing (Table 5.8) typically are tool steels and carbides. For hot drawing, cast-steel dies are 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 made from a single-crystal diamond or in polycrystalline form with diamond particles in a metal matrix (*compacts*). Because of their very low tensile strength and toughness, carbide and diamond dies typically are used as **inserts** or **nibs**, which are supported in a steel casing (Fig. 15.22).

Lubrication. Proper lubrication is essential in drawing in order to improve die life and product surface finish and to reduce drawing forces and temperature. Lubrication is critical, particularly in tube drawing, because of the difficulty of maintaining a sufficiently thick lubricant film at the mandrel–tube interface. In the drawing of rods, a common method of lubrication uses phosphate coatings.

The following are the basic methods of lubrication used in wire drawing (see also Section 33.7):

- **Wet drawing**, in which the dies and the rod are immersed completely in the lubricant
- **Dry drawing**, in which 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**, in which the rod or wire is coated with a soft metal, such as copper or tin, that acts as a solid lubricant
- **Ultrasonic vibration** of the dies and mandrels; in this process, vibrations reduce forces, improve surface finish and die life, and allow larger reductions per pass without failure.

15.9 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. 15.17). Another major type of defect in drawing is **seams**, which are longitudinal scratches or folds in the material. Seams may open up during subsequent forming operations (such as upsetting, heading, thread rolling, or bending of the rod or wire), and they can cause serious quality-control problems. Various other surface defects (such as scratches and die marks) also can result from improper selection of the 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 of the part over time. Moreover, they cause the component to *warp* if a layer of material subsequently is 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 (see Fig. 13.7b).

15.10 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 is activated hydraulically. 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 of pulling force with a speed range of 6 to 60 m/min.

Very long rods and wire (many kilometers) 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*).

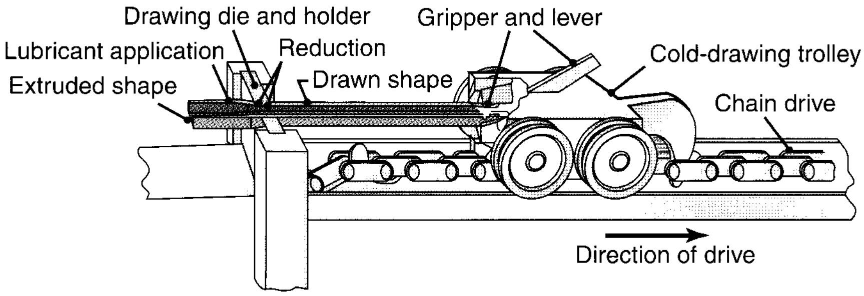


FIGURE 15.23 Cold drawing of an extruded channel on a draw bench to reduce its cross section. Individual lengths of straight rods or of cross sections are drawn by this method.

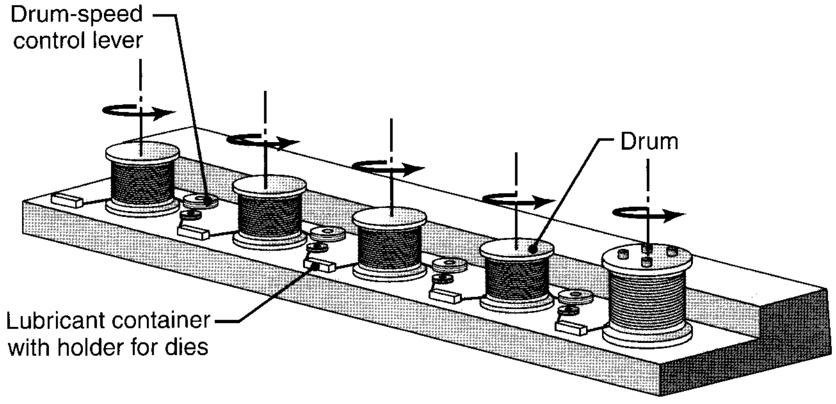


FIGURE 15.24 An illustration of multistage wire drawing typically used to produce copper wire for electrical wiring. *Source*: After H. Auerswald.

SUMMARY

- Extrusion is the process of pushing a billet through a die to reduce its cross section or to produce various solid or hollow cross sections. This process generally is 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 with good mechanical properties and dimensional tolerances.
- Rod, wire, and tube drawing basically involve the process of 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 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 with a good surface finish. External as well as internal defects (chevron cracking) can develop both in extrusion and in drawing. Their minimization or avoidance depends principally on the die angle, the reduction per pass, and the 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.

Blazynski, T.Z., **Plasticity and Modern Metal-forming Technology**, Elsevier, 1989.

Hosford, W.F., and Caddell, R.M., **Metal Forming: Mechanics and Metallurgy**, 3rd ed., Cambridge, 2007.

Inoue, N., and Nishihara, M. (eds.), **Hydrostatic Extrusion: Theory and Applications**, Elsevier, 1985.

Lange, K. (ed.), **Handbook of Metal Forming**, McGraw-Hill, 1985.

Michaeli, W., **Extrusion Dies**, 2nd ed., Hanser, 1992.

Nonferrous Wire Handbook, 2 vols. The Wire Association International, Inc., 1977 and 1981.

Saha, P., **Aluminum Extrusion Technology**, ASM International, 2000.

Sheppard, T., **Extrusion of Aluminum Alloys**, Chapman & Hall, 1997.

Tschaetch, H., **Metal Forming Practice: Processes, Machines, Tools**, Springer, 2007.

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?

QUALITATIVE PROBLEMS

15.14. Explain why extrusion is a batch, or semicontinuous, process. Do you think it can be made into a continuous process? Explain.

15.15. Explain why cold extrusion is an important manufacturing process.

15.16. What is the function of a stripper plate in impact extrusion?

15.17. Explain the different ways by which changing the die angle affects the extrusion process.

15.18. Glass is a good lubricant in hot extrusion. Would you use glass for impression-die forging also? Explain.

15.19. How would you go about avoiding center-cracking defects in extrusion? Explain why your methods would be effective.

15.20. 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.

15.21. Will the force in direct extrusion vary as the billet becomes shorter? If so, why?

15.22. Comment on the significance of grain-flow patterns, such as those shown in Fig. 15.6.

15.23. In which applications could you use the type of impact-extruded parts shown in Fig. 15.16?

15.24. Can spur gears be made by (a) drawing and (b) extrusion? Can helical gears? Explain.

15.25. 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?

15.26. What is the purpose of a dummy block in extrusion? Explain.

15.27. Describe your observations concerning Fig. 15.9.

15.28. 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?

15.29. Explain the advantages of bundle drawing.

15.30. Under what circumstances would backwards extrusion be preferable to direct extrusion?

15.31. Why is lubrication detrimental in extrusion with a porthole die?

15.32. Describe the purpose of a container liner in direct extrusion, as shown in Fig. 15.1.

QUANTITATIVE PROBLEMS

15.33. Estimate the force required in extruding 70–30 brass at 700°C if the billet diameter is 150 mm and the extrusion ratio is 30.

15.34. Assuming an ideal drawing process, what is the smallest final diameter to which an 80-mm diameter rod can be drawn?

15.35. If you include friction in Problem 15.34, would the final diameter be different? Explain.

15.36. 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 70 mm.

15.37. A planned extrusion operation involves steel at 1000°C with an initial diameter of 100 mm and a final diameter of 20 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?

15.38. A round wire made of a perfectly plastic material with a yield stress of 275 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 Eq. (15.3) and Eq. (15.4), estimate the drawing force required. Comment on the differences in your answer.

15.39. 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

15.40. Assume that the summary to this chapter is missing. Write a one-page summary of the highlights of the wire-drawing process.

15.41. Review the technical literature, and make a detailed list of the manufacturing steps involved in the manufacture of common metallic hypodermic needles.

15.42. 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.

15.43. 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.

15.44. 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.

Sheet-Metal Forming Processes and Equipment

CHAPTER 16

- This chapter describes the important characteristics of sheet metals and the forming processes employed to produce a wide variety of products.
- The chapter opens with a discussion of the shearing operation, a process that takes place to cut sheet metal into blanks of desired shapes or to remove portions of the material such as for holes or slots.
- A discussion 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, as well as operations such as stretch forming, rubber forming, spinning, peen forming, and superplastic forming.
- The important process of deep drawing is then described, along with deep drawability, as it relates to the production of containers with thin walls.
- The chapter ends with a discussion of part designs, equipment characteristics, and the economic considerations for all these operations.

Typical parts made by sheet-metal forming: Car bodies, aircraft fuselages, trailers, office furniture, appliances, fuel tanks, and cookware.

Alternative process: Die casting, thermoforming, pultrusion, injection molding, blow molding.

EXAMPLES:

CASE STUDIES:

16.1 Introduction

Products made of **sheet metals** are all around us. They include a very wide range of consumer and industrial products, such as beverage cans, cookware, file cabinets, metal desks, appliances, car bodies, trailers, and aircraft fuselages (Fig. 16.1). Sheet forming dates back to about 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 by forging, sheet-metal parts offer the advantages of light weight and versatile shape.

As described throughout this chapter, there are numerous processes employed for making sheet-metal parts. However, the term **pressworking** or **press forming** is used commonly in industry to describe general sheet-forming operations, because they typically are performed on *presses* (described in Sections 14.8 and 16.14) using a

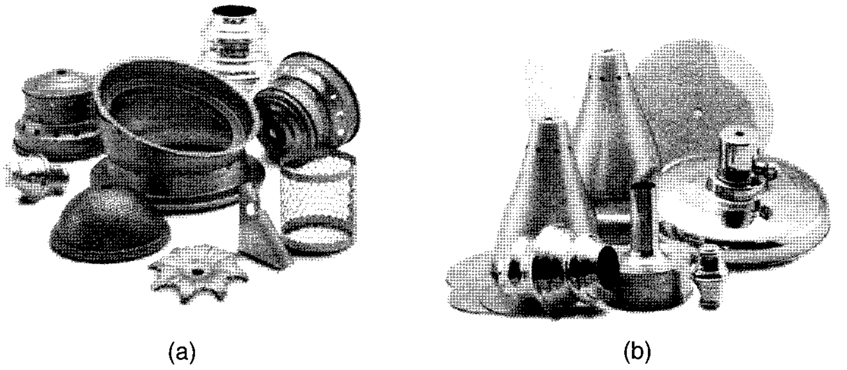


FIGURE 16.1 Examples of sheet-metal parts. (a) Stamped parts. (b) Parts produced by spinning. *Source*: (a) Courtesy of Williamsburg Metal Spinning & Stamping Corp.

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"). Note that this is a term similar to *forging* or *casting*, commonly used to describe parts made by those individual processes using dies or molds, respectively.

Low-carbon steel is the most commonly used sheet metal because of its low cost and generally good strength and formability characteristics. More recently developed alloys, such as TRIP and TWIP steels (see Section 5.5.7), have become popular for automotive applications because of their high strength; they are well suited for providing good crash protection in a lightweight design. Aluminum is the most common material for such sheet-metal applications as beverage cans, packaging, kitchen utensils, and applications where corrosion resistance is a concern. The common metallic materials for aircraft and aerospace applications are aluminum and titanium, although they are being replaced increasingly with composite materials, as described in Chapters 9 and 19.

Most manufacturing processes involving sheet metal are performed at room temperature. Hot stamping is occasionally performed in order to increase formability and decrease forming loads on machinery. Typical materials 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 rolled sheets then processed further into desired shapes by a wide variety of methods. The chapter also includes discussions on the characteristic features of sheet metals, the techniques employed to determine their formability, and the construction of forming-limit diagrams. All of the major processes of sheet forming and the equipment used to make sheet-metal products (as outlined in Table 16.1) are also described.

16.2 Shearing

Before a sheet-metal part is made, a *blank* of suitable dimensions first is removed from a large sheet (usually from a coil) by *shearing*. This sheet is cut by subjecting it to shear stresses, generally using a punch and a die (Fig. 16.2a). The typical features of the sheared edges of the sheet and of the slug are 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.

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

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
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

cracks eventually meet each other and complete separation occurs. The rough *fracture surfaces* are due to the cracks; the smooth and shiny *burnished surfaces* on the hole and the slug are from the contact and rubbing of the sheared edge against the walls of the punch and die, respectively.

The major processing parameters in shearing are

- The shape of the punch and die
- The speed of punching
- Lubrication
- The **clearance**, c, between the punch and the die.

The clearance is a major factor in determining the shape and the quality of the sheared edge. As the clearance increases, the zone of deformation (Fig. 16.3a) becomes larger and the sheared edge becomes rougher. The sheet tends to be pulled into the clearance region, and the perimeter or edges of the sheared zone become rougher. Unless such edges are acceptable as produced, secondary operations may be required to make them smoother (which will increase the production cost).

Edge quality can be improved with increasing punch speed; speeds may be 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 and thus adversely affect the formability of the sheet during subsequent operations, such as bending and stretching.

The ratio of the burnished area to the rough areas along the sheared edge (a) increases with increasing ductility of the sheet metal and (b) decreases with increasing sheet thickness and clearance. The extent of the deformation zone in

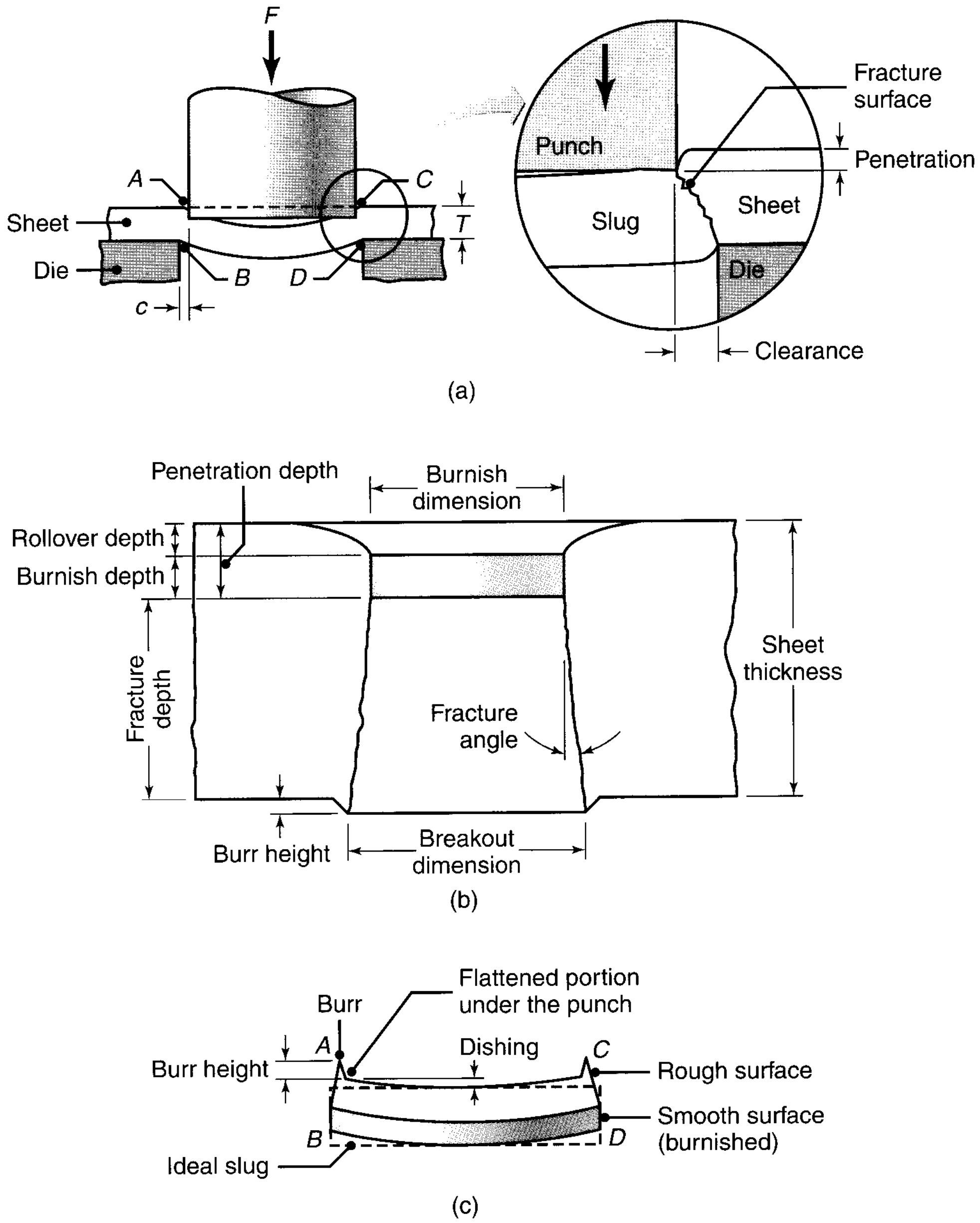


FIGURE 16.2 (a) Schematic illustration of shearing with a punch and die, indicating some of the process variables. Characteristic features of (b) a punched hole and (c) the slug. (Note that the scales of (b) and (c) are different.)

Fig. 16.3 depends on the punch speed. With increasing speed, the heat generated by plastic deformation is confined to a smaller and smaller zone. Consequently, the sheared zone is narrower, and the sheared surface is smoother and exhibits less burr formation. A **burr** is a thin edge or ridge, as shown in Figs. 16.2b and c. Burr 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.

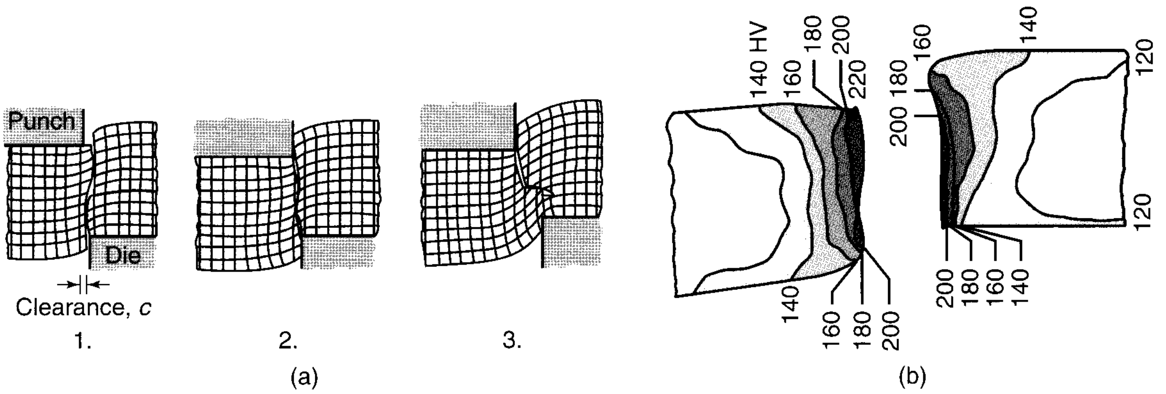


FIGURE 16.3 (a) Effect of the clearance, c, between punch and die on the deformation zone in shearing. As the clearance increases, the material tends to be pulled into the die rather than be sheared. In practice, clearances usually range between 2 and 10% of the thickness of the sheet. (b) Microhardness (HV) contours for a 6.4-mm thick AISI 1020 hot-rolled steel in the sheared region. *Source*: After H.P. Weaver and K.J. Weinmann.

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 along the periphery. The *maximum punch force*, F, can be estimated from the equation

$$F = 0.7TL(\text{UTS}), \tag{16.1}$$

where T is the sheet thickness, L is the total length sheared (such as the perimeter of a hole), and UTS 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 workpiece can, however, increase punch force significantly. Furthermore, in addition to the punch force, a force is required to strip the punch from the sheet during its return stroke. This second force, which is in opposite direction of the punch force, is difficult to estimate because of the many factors involved in the operation.

EXAMPLE 16.1 Calculation of Punch Force

Estimate the force required for punching a 25-mm diameter hole through a 3.2-mm thick annealed titanium-alloy Ti-6Al-4V sheet at room temperature.

Solution The force is estimated from Eq. (16.1), where the UTS for this alloy is found from Table 6.10 to be 1000 MPa. Thus,

$$F = 0.7(3.2)(\pi)(25)(1000) = 0.18 \text{ MN}.$$

16.2.1 Shearing Operations

The most common shearing operations are **punching**—where the sheared slug is scrap (Fig. 16.4a) or may be used for some other purpose—and **blanking**—where the slug is the part to be used and the rest is scrap. The operations described next, as

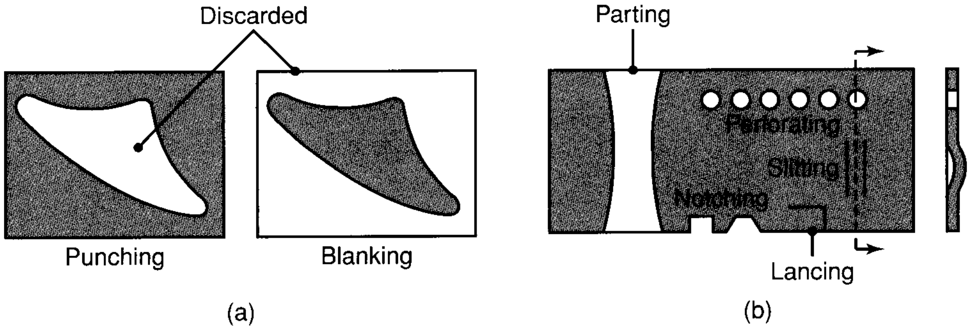


FIGURE 16.4 (a) Punching (piercing) and blanking. (b) Examples of various die-cutting operations on sheet metal. Lancing involves slitting the sheet to form a tab.

well as those described throughout the rest of this chapter, generally are carried out on computer-numerical-controlled machines with quick-change toolholders. Such machines are useful, particularly in making prototypes of sheet-metal parts requiring several operations to produce.

Die Cutting. This is a shearing operation that consists of the following basic processes (Fig. 16.4b):

- **Perforating**: punching a number of holes in a sheet
- **Parting**: shearing the sheet into two or more pieces
- **Notching**: removing pieces (or various shapes) from the edges
- **Lancing**: leaving a tab without removing any material.

Parts produced by these processes have various uses, particularly in assembly with other components. Perforated sheet metals with hole diameters ranging from around 1 mm to 75 mm have uses as filters, as screens, in ventilation, as guards for machinery, in noise abatement, and in weight reduction of fabricated parts and structures. They are punched in crank presses (see Fig. 14.17a) at rates as high as 300,000 holes per minute, using special dies and equipment.

Fine Blanking. Very smooth and square edges 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 tightly in place and prevents the type of distortion of the material shown in Figs. 16.2b and 16.3. The fine-blanking process, which was developed in the 1960s, involves clearances on the order of 1% of the sheet thickness and that may range from 0.5 to 13 mm in most cases. Dimensional tolerances are on the order of ±0.05 mm and less than ±0.025 mm in the case of edge perpendicularity.

Slitting. Shearing operations can be carried out by means of a pair of circular blades similar to those in a can opener (Fig. 16.6). In *slitting*, the blades follow either a straight line, a circular path, or a curved path. A slit edge normally has a burr, which may be folded over the sheet surface by rolling it (flattening) between two rolls. If not performed properly, slitting operations can cause various distortions of the sheared edges.

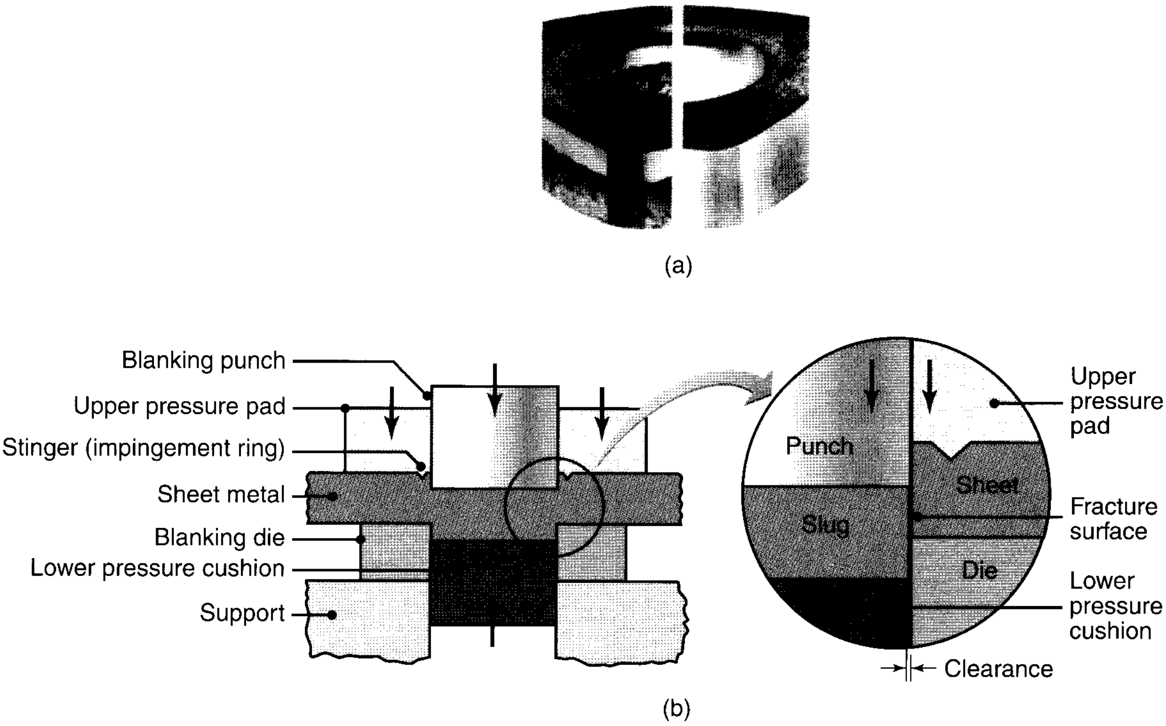


FIGURE 16.5 (a) Comparison of sheared edges produced by conventional (left) and by fine-blanking (right) techniques. (b) Schematic illustration of one setup for fine blanking. *Source*: Courtesy of Feintool U.S. Operations.

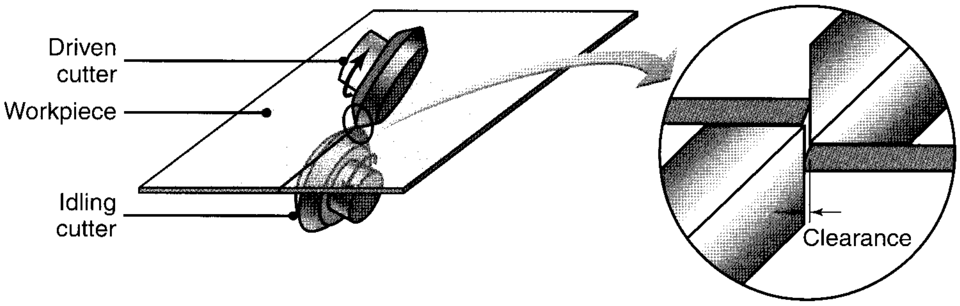


FIGURE 16.6 Slitting with rotary knives. This process is similar to opening cans.

Steel Rules. Soft metals (as well as 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 produced (a concept similar to that of a cookie cutter) and held on its edge on a flat wood or polymer base. The die is pressed against the sheet, which rests on the flat surface, and it shears the sheet along the shape of the steel rule.

Nibbling. In *nibbling*, 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 many overlapping holes are made. With manual or automatic control, sheets can be cut along any desired path. In addition to its flexibility, an advantage of nibbling is that intricate slots and notches, such as those shown in Fig. 16.4b, can be produced with standard punches. The process is economical for small production runs because no special dies are required.

Scrap in Shearing. The amount of scrap (*trim loss*) produced in shearing operations can be significant and can be as high as 30% on large stampings (see Table 40.3). Scrap can be a significant factor in manufacturing cost, and it can be reduced substantially by efficient arrangement of the shapes on the sheet to be cut (**nesting**, see Fig. 16.55). Computer-aided design techniques have been developed to minimize the scrap from shearing operations.

16.2.2 Tailor-welded Blanks

In the sheet-metal-forming processes to be described throughout this chapter, the blank is usually a one-piece sheet of one thickness cut from a large sheet. An important variation from these conditions involves *laser-beam butt welding* (see 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 and are then coiled.

Because of the small thicknesses involved, the proper alignment of the sheets prior to welding is important. The welded assembly subsequently is formed into a final shape (see Example 16.2). This technique is becoming increasingly important, particularly to the automotive industry. Because each subpiece now can have a different thickness, grade, coating, or other property, tailor-welded blanks possess the needed properties in the desired locations in the blank. The result is

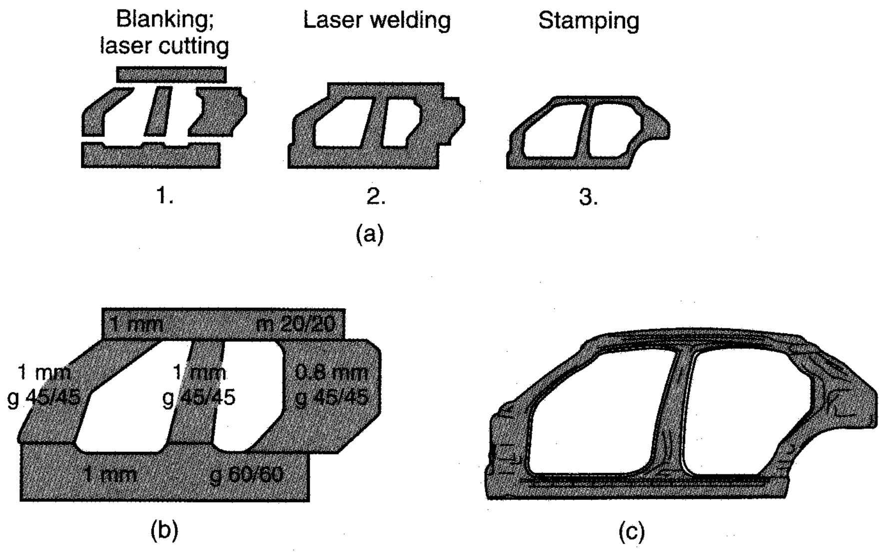
- Reduction in scrap
- Elimination of the need for subsequent spot welding (i.e., in the making of the car body)
- Better control of dimensions
- Improved productivity.

EXAMPLE 16.2 Tailor-welded Sheet Metal for Automotive Applications

An example of the use of tailor-welded sheet metals in automobile bodies is shown in Fig. 16.7. Note that five different pieces are blanked first, which includes cutting by laser beams. Four of these pieces are 1 mm thick, and one is 0.8 mm thick. The pieces are laser butt welded and 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 sheets—with or without coatings.

Laser-welding techniques are highly developed; as a consequence, weld joints are very strong and reliable. The growing trend toward welding and forming sheet-metal pieces makes possible significant flexibility in the product design, structural stiffness, formability, and crash behavior of an automobile. It also makes possible the use of different materials in one component, weight savings, and cost reduction in materials, scrap, equipment, assembly, and labor.

There are increasing applications for this type of production in automobiles. The various components shown in Fig. 16.8 utilize the advantages outlined above. For example, note in Fig. 16.8(b) 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 is varied (depending on its location and on its contribution to such characteristics as stiffness and strength) and thereby makes possible significant weight savings without loss of structural strength and stiffness.



Legend

g 60/60 (45/45)	Hot-galvanized alloy steel sheet. Zinc amount: 60/60 (45/45) g/m^2.
m 20/20	Double-layered iron–zinc alloy electroplated steel sheet. Zinc amount 20/20 g/m^2.

FIGURE 16.7 Production of an outer side panel of a car body by laser butt welding and stamping. *Source*: After M. Geiger and T. Nakagawa.

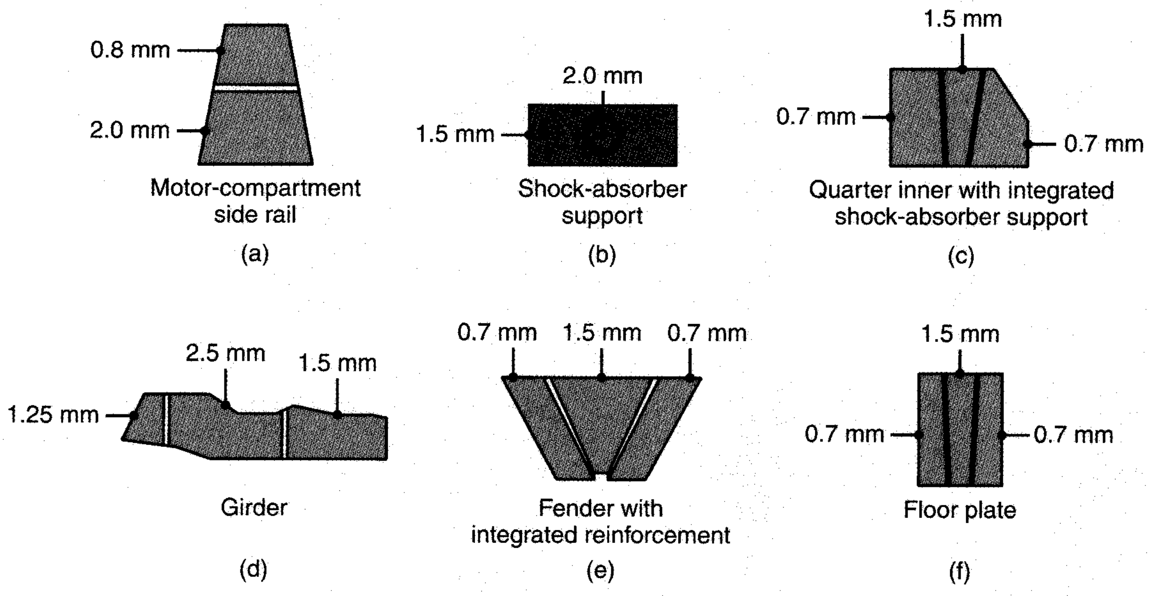


FIGURE 16.8 Examples of laser butt-welded and stamped automotive-body components. *Source*: After M. Geiger and T. Nakagawa.

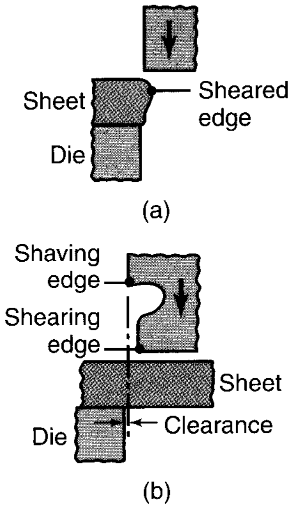


FIGURE 16.9 Schematic illustrations of the shaving process. (a) Shaving a sheared edge. (b) Shearing and shaving combined in one stroke.

16.2.3 Characteristics and Type of Shearing Dies

Clearance. Because the formability of the sheared part can be influenced by the quality of its sheared edges, clearance control is important. The appropriate clearance depends on

- The type of material and its temper
- The thickness and size of the blank
- Its proximity to the edges of other sheared edges or the edges of the original blank.

Clearances generally range between 2 and 8% of the sheet thickness, but they may be as small as 1% (as in fine blanking) or as large as 30%. The smaller the clearance, the better is the quality of the edge. If the sheared edge is rough and not acceptable, it can be subjected to a process called **shaving** (Fig. 16.9a), whereby the extra material from the edge is trimmed by cutting, as also depicted in Fig. 21.3.

As a general guideline, (a) clearances for soft materials are less than those for harder grades; (b) the thicker the sheet, the larger the clearance must be; and (c) as the ratio of hole diameter to sheet thickness decreases, clearances should be larger. 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 Shape. Note in Fig. 16.2a that the surfaces of the punch and of the die are both flat. Because the entire thickness is sheared at the same time, the punch force increases rapidly during shearing. The location of the regions being sheared at any particular instant can be controlled by *beveling* the punch and die surfaces (Fig. 16.10). This shape is similar to that of some paper punches, which you can observe by looking closely at the tip of the punch. Beveling is suitable particularly for shearing thick sheets because it reduces the force at the beginning of the stroke. It also reduces the operation's noise level, because the operation is smoother.

Note in Fig. 16.10c that the punch tip is symmetrical and in Fig. 16.10d that the die is symmetrical. Hence, there are no lateral forces acting on the punch to cause distortion. By contrast, the punch in Fig. 16.10b has a single taper and thus is subjected to a lateral force. Consequently, the punch and press setups in this latter case 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 (as it might at point *B* or *D* in Fig. 16.2a), causing damage.

Compound Dies. Several operations on the same sheet may be performed in one stroke at one station with a *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 rapidly become much more expensive to produce than those for individual shearing operations, especially for complex dies.

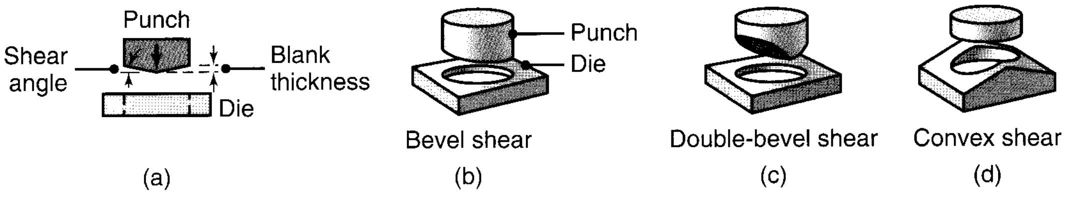


FIGURE 16.10 Examples of the use of shear angles on punches and dies.

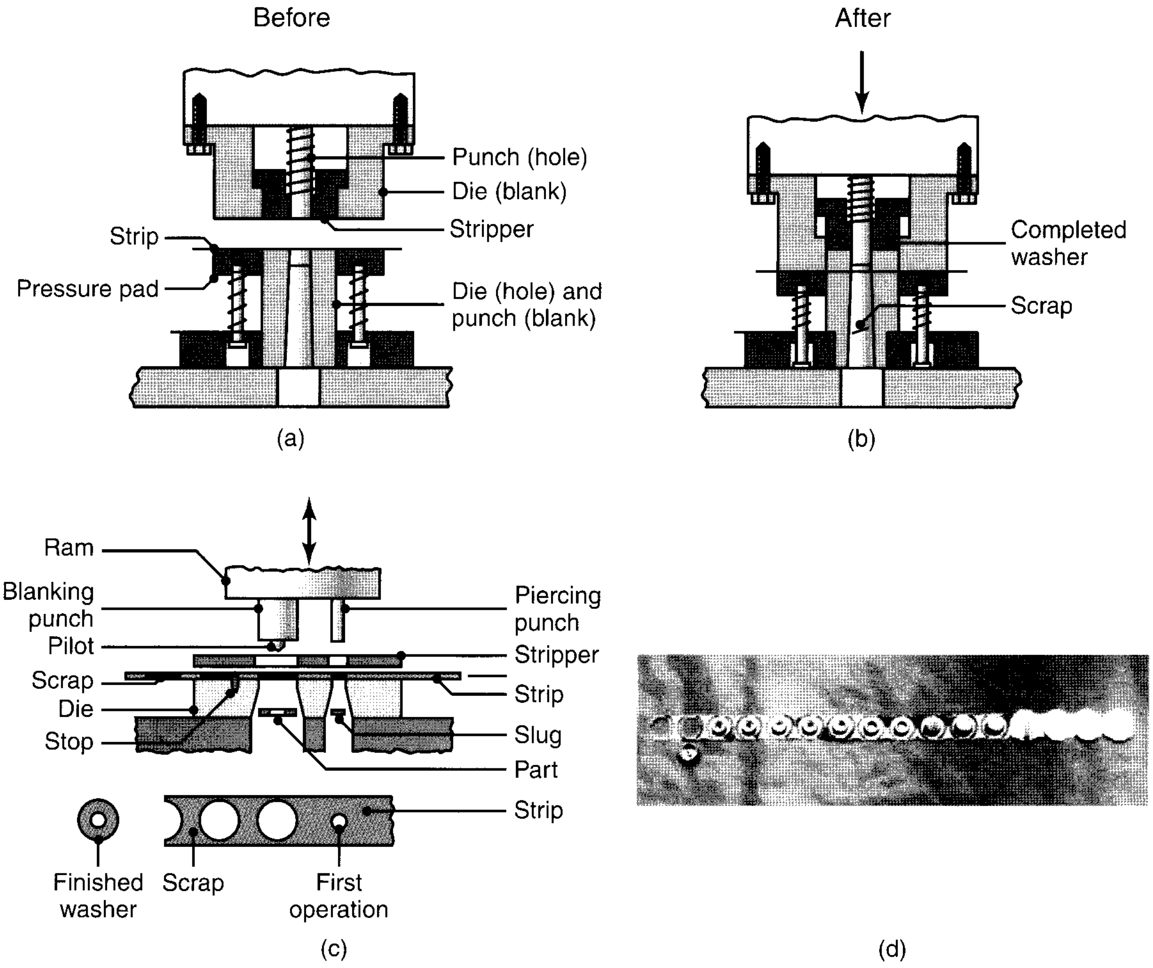


FIGURE 16.11 Schematic illustrations (a) before and (b) after blanking a common washer in a compound die. Note the separate movements of the die (for blanking) and the punch (for punching the hole in the washer). (c) Schematic illustration of making a washer in a progressive die. (d) Forming of the top 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.

Progressive Dies. Parts requiring multiple operations to produce can be made at high production rates in *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 of a series of punches (Fig. 16.11c). An example of a part made in progressive dies is shown in Fig. 16.11d; the part is the small round piece that supports the plastic tip in spray cans.

Transfer Dies. In a *transfer-die* setup, the sheet metal undergoes different operations at different stations of the machine that are arranged along a straight line or a circular path. After each step in a station, the part is transferred to the next station for further operations.

Tool and Die Materials. Tool and die materials for shearing generally are tool steels and (for high production rates) carbides (see Table 5.7). Lubrication is important for reducing tool and die wear, thus improving edge quality.

16.2.4 Miscellaneous Methods of Cutting Sheet Metal

There are several other methods of cutting sheets and, particularly, plates:

- **Laser-beam cutting** is an important process (Section 26.7) typically used with computer-controlled equipment to cut a variety of shapes consistently, in various thicknesses, and without the use of any dies. Laser-beam cutting also can be combined with punching and shearing. These processes cover different and complementary ranges. Parts with certain features can be produced best by one process; some with other features can be produced best by the other process. Combination machines incorporating both capabilities have been designed and built. (See also Example 27.1.)
- **Water-jet cutting** is effective on many metallic as well as nonmetallic materials (Section 27.8).
- Cutting with a **band saw**; this method is a chip-removal process.
- **Friction sawing** involves a disk or blade that rubs against the sheet or plate at high surface speeds (Section 24.5).
- **Flame cutting** is another common method, particularly for thick plates; it is used widely in shipbuilding and on heavy structural component (Section 30.8).

16.3 Sheet-metal Characteristics and Formability

After a blank is cut from a larger sheet or coil, it is formed into various shapes by several processes described in the rest of this chapter. We will now briefly review those characteristics of sheet metals that have important effects on these forming operations, as outlined in Table 16.2.

TABLE 16.2

Important Metal Characteristics for Sheet-forming Operations

Characteristic	Importance
Elongation	Determines the capability of the sheet metal to stretch 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 depressions on the sheet surface; can be eliminated by temper rolling, but sheet must be formed within a certain 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 (like an 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 techniques such 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 contain 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

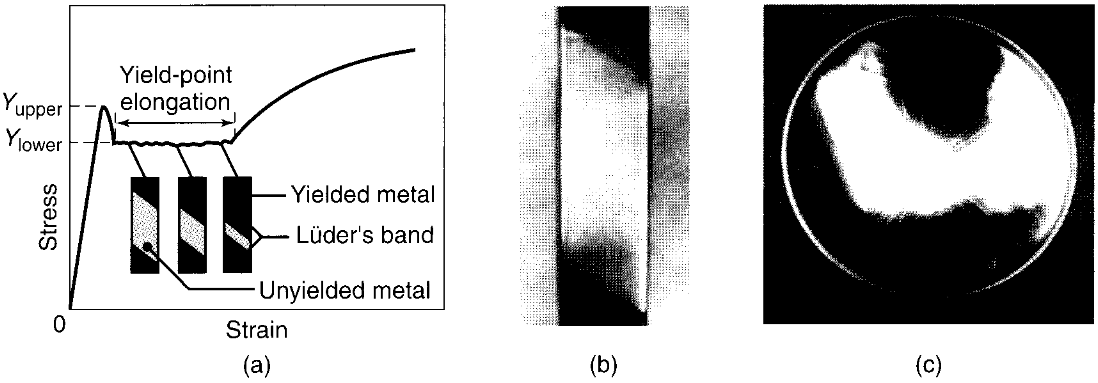
Elongation. Sheet-metal-forming processes rarely involve simple uniaxial stretching like that in a tension test. However, observations from tensile testing are useful and necessary for understanding the behavior of metals in these operations. Recall from Section 2.2 that a specimen subjected to tension first undergoes **uniform elongation** and that when the load exceeds the ultimate tensile strength of the material, the specimen begins to neck and thus elongation is no longer uniform.

Because the material usually is being stretched in sheet forming, high uniform elongation is desirable for good formability. The true strain at which necking begins is numerically equal to the *strain-hardening 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; this relationship is given in Eq. (2.9). The higher the value of m, the more diffuse the neck becomes. A diffuse neck is desirable in sheet-forming operations. 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 behavior results in **Lüder's bands** (also called *stretcher-strain marks* or *worms*) on the sheet (Fig. 16.12b)—elongated depressions on the surface of the sheet, such as can be found on the bottom of cans containing common household products (Fig. 16.12c). These marks may be objectionable in the final product, because coarseness on the surface degrades appearance and may 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 reducing the thickness of the sheet 0.5 to 1.5% by cold rolling (**temper** or **skin rolling**). Because of strain aging, however, the yield-point elongation reappears after a few days at room temperature or after a few hours at higher temperatures. To prevent this undesirable occurrence, the material should be formed within a certain time limit (which depends on the type of the steel).

Anisotropy. An important factor that influences sheet-metal forming is *anisotropy* (*directionality*) of the sheet. Recall that anisotropy is acquired during the thermomechanical processing of the sheet and that there are two types of anisotropy: *crystallographic anisotropy* (preferred orientation of the grains) and *mechanical fibering*



(a) (b) (c)

FIGURE 16.12 (a) Yield-point elongation in a sheet-metal specimen. (b) Lüder's bands in a low-carbon steel sheet. (c) Stretcher strains at the bottom of a steel can for household products. *Source*: (b) Courtesy of Caterpillar, Inc.

(alignment of impurities, inclusions, and voids throughout the thickness of the sheet). The relevance of this subject is discussed further in Section 16.4.

Grain Size. As described in Section 1.4, 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; the coarser the grain, the rougher is the surface appearance. An ASTM grain size of 7 or finer (Table 1.1) is preferred for general sheet-forming operations.

Dent Resistance of Sheet Metals. Dents commonly are found on cars, appliances, and office furniture. Dents usually are caused by dynamic forces from moving objects that hit the sheet metal. In typical automotive panels, for example, velocities at impact range up to 45 m/s. Thus, it is the *dynamic yield stress* (yield stress under high rates of deformation), rather than the static yield stress, that is the significant strength parameter.

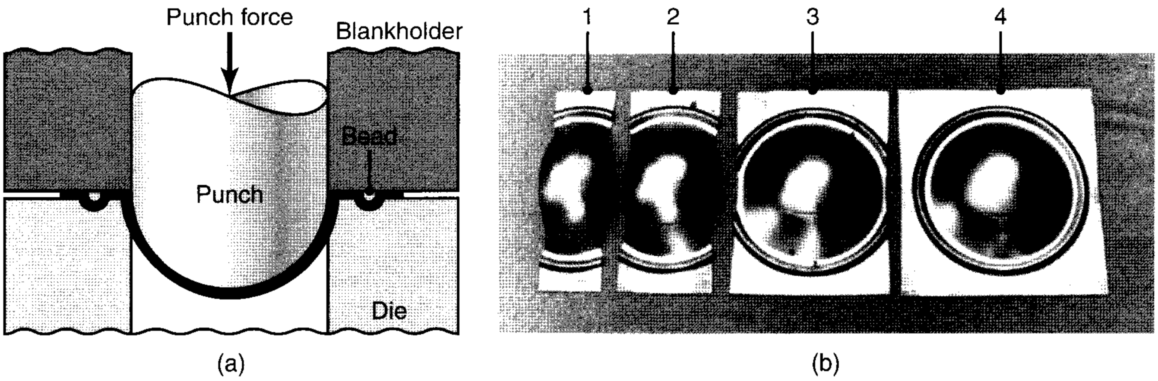
Dynamic forces tend to cause *localized* dents, whereas static forces tend to *diffuse* the dented area. This phenomenon may be demonstrated by trying to dent a piece of flat sheet metal, first by pushing a ball-peen hammer against it and then by striking it with the hammer. Note how localized the dent will be in the latter case. *Dent resistance* of sheet-metal parts has been found to (a) increase as the sheet thickness and its yield stress increase and (b) decrease as its elastic modulus and its overall panel stiffness increase. Consequently, panels rigidly held at their edges have lower dent resistance because of their higher stiffness.

16.4 Formability Tests for Sheet Metals

Sheet-metal formability is of great technological and economic interest, and it generally is defined as the ability of the sheet metal to undergo the desired shape change without failure, such as by necking, cracking, or tearing. As we shall see throughout the rest of this chapter, sheet metals (depending in part on geometry) 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 these different conditions. This section describes the methods that generally are used to predict formability.

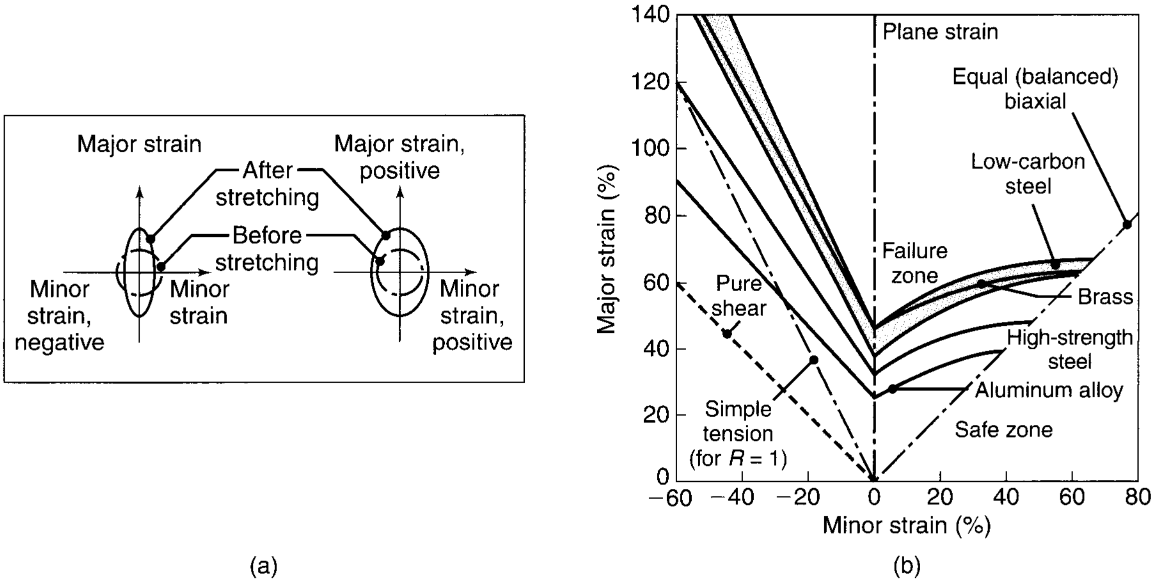
Cupping Tests. The earliest tests developed to predict sheet-metal formability were cupping tests (Fig. 16.13a). In the *Erichsen test*, the sheet specimen is clamped between two circular, flat dies and a steel ball or round punch is forced into the sheet until a crack begins to appear on the stretched specimen. The *punch depth, d*, at which failure occurs is a measure of the formability of the sheet. Although this and similar tests are easy to perform, they do not simulate the exact conditions of actual forming operations and hence are not particularly reliable, especially for complex parts.

Forming-limit Diagrams. An important advance in testing the formability of sheet metals is the development of *forming-limit diagrams*, as shown in Fig. 16.14. A forming-limit diagram (FLD) for a particular metal is constructed by first marking the flat sheet with a grid pattern of circles (see Fig. 16.15), using chemical or photoprinting techniques. The blank then is stretched over a punch (Fig. 16.13a), and the deformation of the circles is observed and measured in regions where failure (*necking* and *tearing*) has occurred. Although the circles typically are 2.5 to 5 mm in diameter, for improved accuracy of measurement, they should be made as small as is practical.



(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*: Courtesy of Inland Steel Company.



(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 16.4. *Source*: After S.S. Hecker and A.K. Ghosh.

In order to develop unequal stretching to simulate 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) approaches the state of *uniaxial stretching* (that is, simple tension). After a series of such tests is performed on a particular sheet metal and at different widths, a forming-limit diagram is constructed showing the boundaries between *failure* and *safe* regions (Fig. 16.14b).

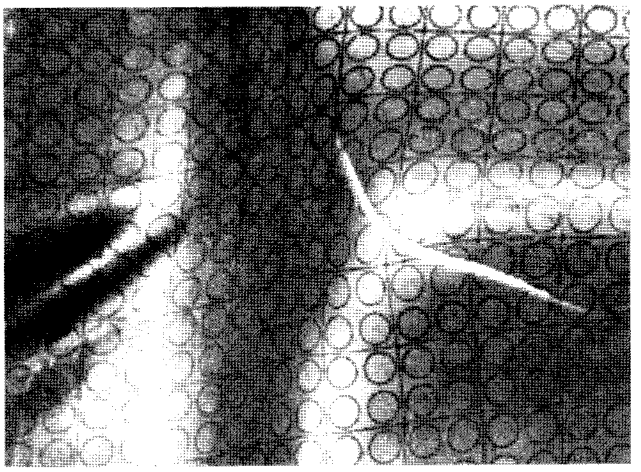


FIGURE 16.15 The deformation of the grid pattern and the tearing of sheet metal during forming. The major and minor axes of the circles are used to determine the coordinates on the forming-limit diagram in Fig. 16.14b. *Source*: S.P. Keeler.

In order to develop a forming-limit diagram, the major and minor engineering strains, as measured from the deformation of the original circles, are obtained. Note in Fig. 16.14a that the original circle has deformed into an ellipse. The *major axis* of the ellipse 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 magnitude of the stretching or shrinking in the *transverse* direction.

Note 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), and thus the minor strain is negative. (This behavior can be demonstrated easily by stretching a rubber band and observing the dimensional changes it undergoes.) On the other hand, if we place a circle on a spherical rubber balloon and inflate it, the minor and major strains are both positive and equal in magnitude.

By comparing the surface areas of the original circle and the deformed circle on the formed sheet, we also can determine whether the thickness of the sheet has changed during deformation. Because the volume remains constant in plastic deformation, we know that if the area of the deformed circle is larger than the original circle, the sheet has become thinner. This phenomenon can be demonstrated easily by blowing up a balloon and noting that it becomes more translucent as it is stretched (because it is getting thinner).

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 metal, and as can be noted, the higher the curve, the better is the formability of that particular metal. As expected, different materials and conditions (such as cold worked or heat treated) 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 sheet 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 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 on the formed sheet.

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 the forming of 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 forming-limit diagrams is to raise the curves in Fig. 16.14b. The thicker the sheet, the higher its formability curve, and thus the more formable the sheet is. On the other hand, in actual forming operations, a thick blank may not bend as easily around small radii without cracking (as described in Section 16.5 on bending). Friction and lubrication at the interface between the punch and the sheet metal are also important factors in the test results. With well-lubricated interfaces, the strains in the sheet are distributed more uniformly over the punch. Also, as expected, and depending on the material and its notch sensitivity, surface scratches, deep gouges, and blemishes can reduce formability significantly and thereby lead to premature tearing and failure of the part.

16.5 Bending Sheets, Plates, and Tubes

Bending is one of the most common industrial forming operations. We merely have to look at an automobile body, appliance, paper clip, or file cabinet to appreciate how many parts are shaped by bending. Furthermore, bending also imparts stiffness to the part by increasing its moment of inertia. Note, for example, how corrugations, flanges, beads, and seams improve the stiffness of structures without adding any weight. As a specific example, observe the diametral stiffness of a metal can with and without circumferential beads (see also *beading*).

The terminology used in the bending of a sheet or plate is shown 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) has become smaller in the outer region and larger in the inner region than the original width (as can be seen in Fig. 16.17c). This phenomenon may be observed easily 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,

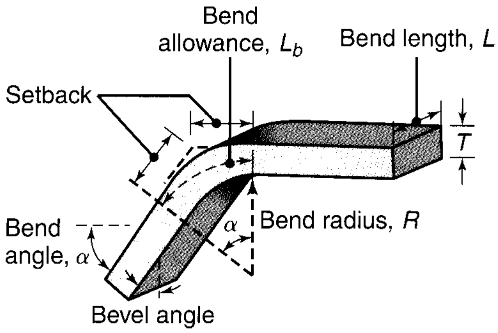


FIGURE 16.16 Bending terminology. Note that the bend radius is measured to the inner surface of the bent part.

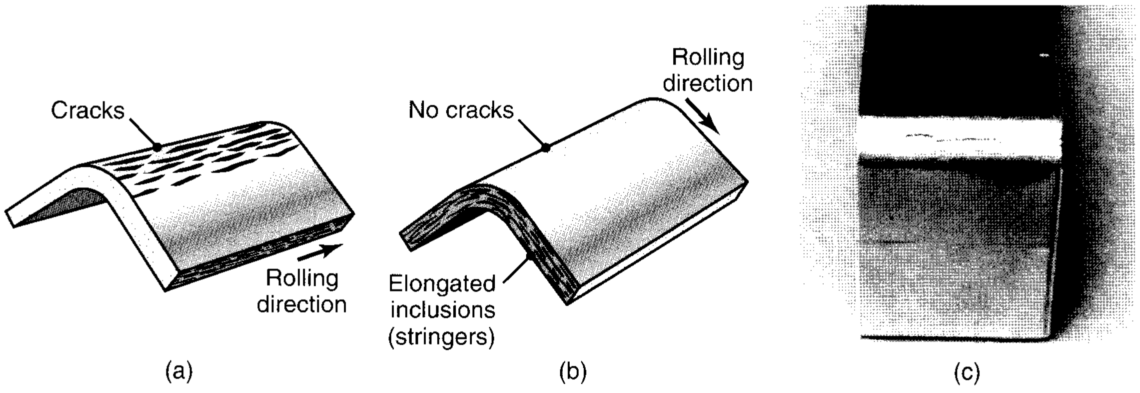


FIGURE 16.17 (a) and (b) The effect of elongated inclusions (stringers) on cracking as a function of the direction of bending with respect to the original rolling direction of the sheet. (c) Cracks on the outer surface of an aluminum strip bent to an angle of 90°. Note also the narrowing of the top surface in the bend area (due to the Poisson effect).

TABLE 16.3

Minimum Bend Radius for Various Metals at Room Temperature

Material	Condition	
	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	$2.6T$	$4T$

however, depends on the radius and the bend angle (as described in texts on mechanics of materials). 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. In practice, k values typically range 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,

$$L_b = \alpha\left[R + \left(\frac{T}{2}\right)\right]. \tag{16.3}$$

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 (that is, 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 (reciprocally) in terms of the thickness, such as $2T$, $3T$, $4T$, and so on (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.

There is an inverse relationship between *bendability* and the tensile reduction of the area of the material (Fig. 16.18). The *minimum bend radius*, R, is, approximately,

$$R = T\left(\frac{50}{r} - 1\right), \tag{16.5}$$

where r is the tensile reduction of area of the sheet metal. Thus, for $r = 50$, the minimum bend radius is zero; that is, the sheet can be folded over itself (see *hemming*, Fig. 16.23) in much the same way as a piece of paper is folded. To increase the bendability of metals, we may increase their tensile reduction of area either by heating or by bending in a high-pressure environment (which improves the ductility of the material;

see *hydrostatic stress*, Section 2.2.8). Bendability also depends on the *edge condition* of the sheet. Since rough edges are points of stress concentration, bendability decreases as edge roughness increases.

Another significant factor in edge cracking is the amount, shape, and hardness of *inclusions* present in the sheet metal and the amount of cold working that the edges undergo 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 improved greatly by removing the cold-worked regions by shaving or machining the edges of the part (see Fig. 16.9) or by annealing it to improve its ductility.

Anisotropy of the sheet is another important factor in bendability. Cold rolling results in anisotropy by *preferred orientation* or by *mechanical fibering* due to the alignment of any impurities, inclusions, and voids that may be present, as shown in Fig. 1.13. Prior to laying out or nesting (see Fig. 16.55) blanks for subsequent bending, caution should be exercised to cut in the proper direction from a rolled sheet; however, this choice may not always be possible in practice.

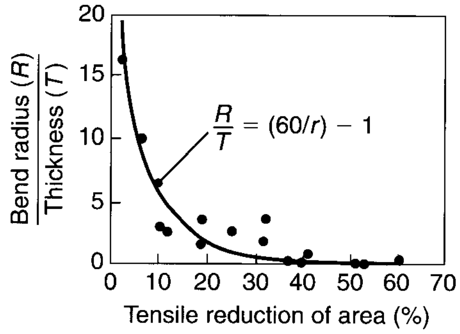


FIGURE 16.18 Relationship between R/T and tensile reduction of area for sheet metals. Note that sheet metal with a 50% tensile reduction of area can be bent over itself in a process like the folding of a piece of paper without cracking. *Source*: After J. Datsko and C.T. Yang.

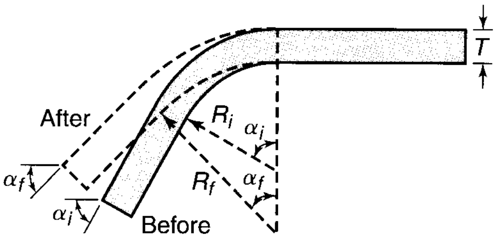


FIGURE 16.19 Springback in bending. The part tends to recover elastically after bending, and its bend radius becomes larger. Under certain conditions, it is possible for the final bend angle to be smaller than the original angle (negative springback).

Springback. Because all materials have a finite modulus of elasticity, plastic deformation always is followed by some elastic recovery when the load is removed (see Fig. 2.3). In bending, this recovery is called *springback*, which can be observed easily by bending and then releasing a piece of sheet metal or wire. Springback occurs not only in flat sheets and plates, but also in solid or hollow bars and tubes of any cross section. As noted in Fig. 16.19, the final bend angle after springback is smaller than the angle to which the part was bent, and the final bend radius is larger than before springback occurs.

Springback can be calculated approximately in terms of the radii R_i and R_f (Fig. 16.19) as

$$\frac{R_i}{R_f} = 4\left(\frac{R_iY}{ET}\right)^3 - 3\left(\frac{R_iY}{ET}\right) + 1. \qquad (16.6)$$

Note from this formula that springback increases (a) as the R/T ratio and the yield stress, Y, of the material increase and (b) 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 condition is caused by the nature of the deformation occurring just as the punch completes the bending operation at the end of the stroke. Negative springback does not occur in *air bending*, shown in Fig. 16.22a (also called *free bending*), 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 for by *overbending* the part (Figs. 16.20a and b). 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 (Figs. 16.20c and d)—a technique known as *bottoming* the punch. Another method is *stretch bending*, in which the part is subjected to tension while being bent (see also *stretch forming*, Section 16.6).

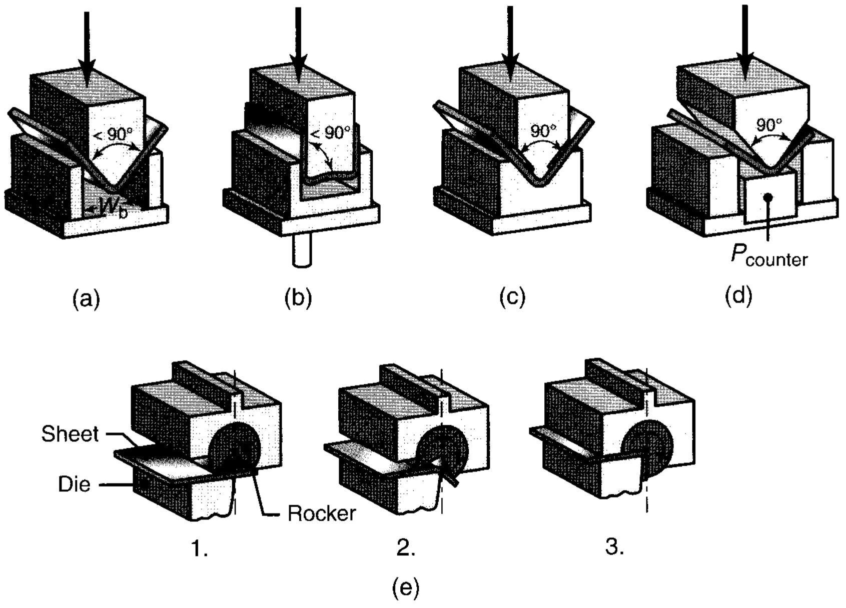
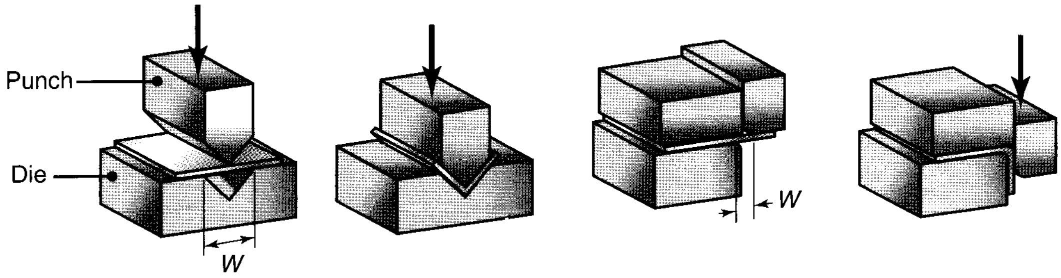


FIGURE 16.20 Methods of reducing or eliminating springback in bending operations.



(a) V-die (b) Wiping die

FIGURE 16.21 Common die-bending operations showing the die-opening dimension, *W*, used in calculating bending forces.

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 strength of the material, the length, L, of the bend, the thickness, T, of the sheet, and the die opening, W. (W is shown in Fig. 16.21.) Excluding friction, the *maximum bending force, P*, is

$$P = \frac{kYLT^2}{W}, \tag{16.7}$$

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 and Y is the yield stress of the material.

For a V-die, Eq. (16.7) is often modified as

$$P = \frac{(\text{UTS})LT^2}{W}, \tag{16.8}$$

where UTS is the ultimate tensile strength of the material. This equation applies well to situations in which the punch-tip radius and the sheet thickness are relatively small compared to the die opening, W.

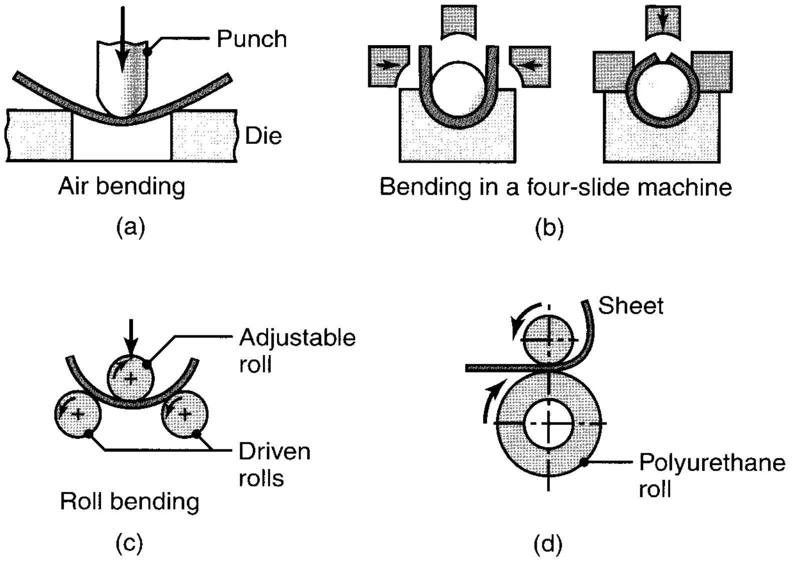


FIGURE 16.22 Examples of various bending operations.

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 completed. The force then increases sharply as the punch reaches the bottom of its stroke and the part touches the bottom of the die. In air bending (Fig. 16.22a), however, the force does not increase again after it begins to decrease, as it has no resistance to its free movement downward.

16.6 Miscellaneous Bending and Related Operations

Press-brake Forming. Sheet metal or plate can be bent easily with simple fixtures using a press. Sheets or narrow strips that are 7 m or even longer 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, their motions are only up and down, and they easily are adaptable to a wide variety of shapes. Also, the process can be automated easily 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 materials and also are chosen to improve die life. For most applications, however, carbon-steel or gray-iron dies generally are used.

Bending in a Four-slide Machine. Bending relatively short pieces can be done on a machine such as that shown in Fig. 16.22b. In these machines, the lateral movements of the dies are controlled and synchronized with the vertical die movement to form the part into desired shapes. This process is useful in 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. By adjusting the distance between the three rolls, various curvatures can be obtained. This process is flexible and is used widely for bending plates for applications such as boilers, cylindrical pressure vessels, and various curved structural members. Figure 16.22d

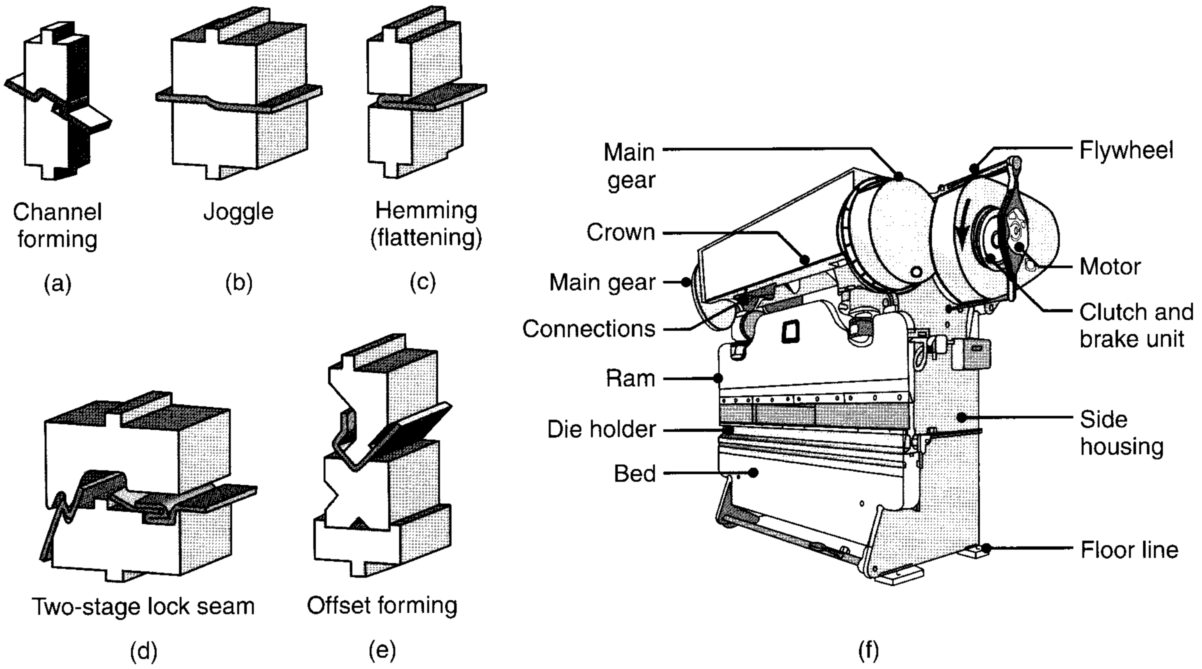


FIGURE 16.23 (a) through (e) Schematic illustrations of various bending operations in a press brake. (f) Schematic illustration of a press brake. *Source*: Courtesy of Verson Allsteel Company.

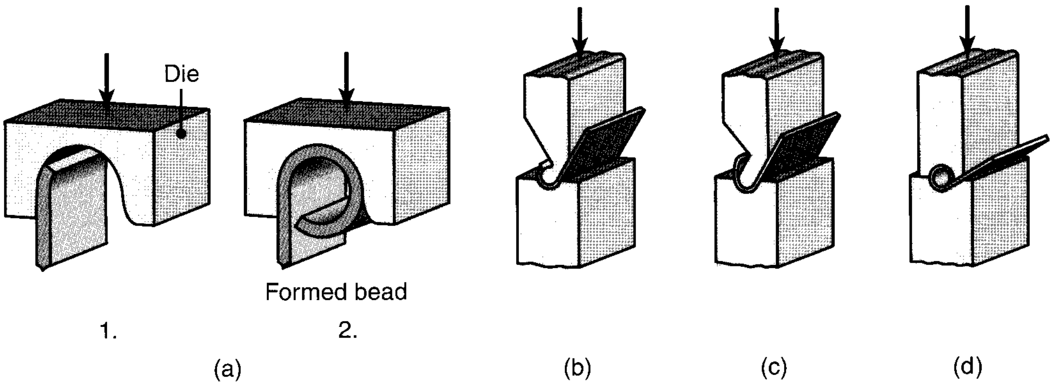


FIGURE 16.24 (a) Bead forming with a single die. (b) through (d) Bead forming with two dies in a press brake.

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.

Beading. In *beading*, the periphery of the sheet metal is bent into the cavity of a die (Fig. 16.24). The bead imparts stiffness to the part by increasing the moment of inertia of that section. Also, beads improve the appearance of the part and eliminate exposed sharp edges that can be hazardous.

Flanging. This is a process of bending the edges of sheet metals, usually to 90°. In **shrink flanging** (Fig. 16.25a), the flange is subjected to compressive hoop stresses that, if excessive, can cause the flange periphery to wrinkle. The wrinkling tendency

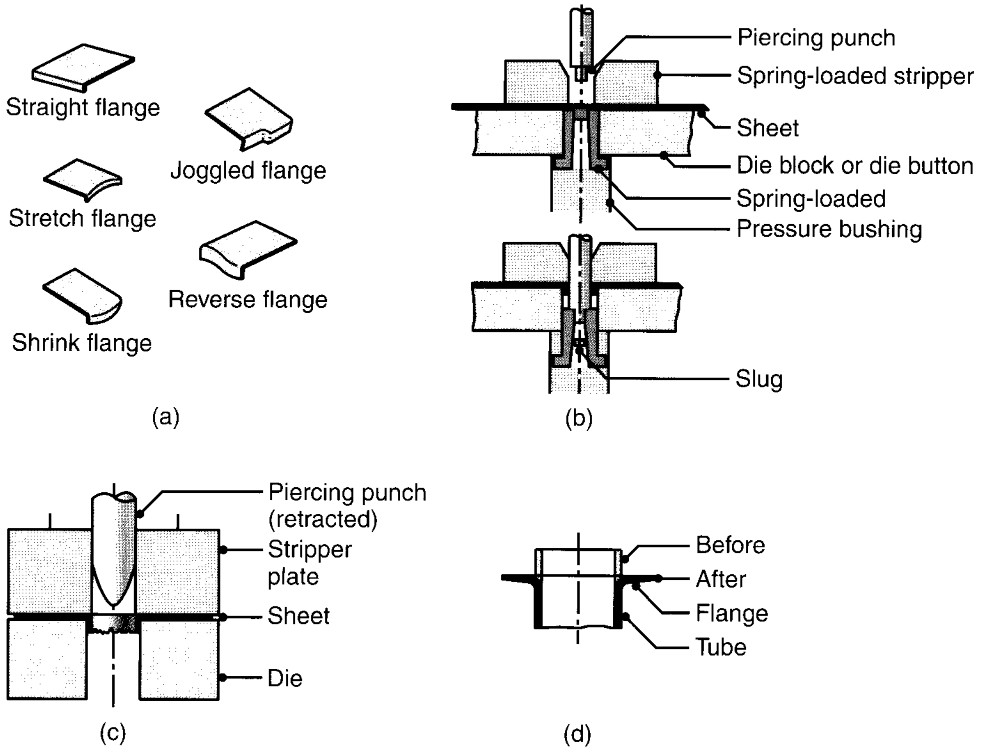


FIGURE 16.25 Various flanging operations. (a) Flanges on flat sheet. (b) Dimpling. (c) The piercing of sheet metal to form a flange. In this operation, a hole does not have to be prepunched before the punch descends. Note, however, the rough edges along the circumference of the flange. (d) The flanging of a tube. Note the thinning of the edges of the flange.

increases with decreasing radius of curvature of the flange. In **stretch flanging**, the flange periphery is subjected to tensile stresses that, if excessive, can lead to cracking along the periphery.

Roll Forming. This process, which is also called *contour-roll forming* or *cold-roll forming*, is used for forming continuous lengths of sheet metal and for large production runs. As it passes through a set of rolls, the metal strip is bent in consecutive stages (Fig. 16.26). The formed strip is then sheared into specific lengths and stacked continuously.

Typical roll-formed products are panels, door and picture frames, channels, gutters, siding, and pipes and tubing with lock seams (see Section 32.5). The length of the part is limited only by the amount of material supplied to the rolls from the coiled stock. Sheet thickness usually ranges from about 0.125 to 20 mm. Forming speeds are generally below 1.5 m/s, although they can be much higher for special applications.

The design and sequencing of the rolls (which usually are mechanically driven) require considerable experience. Dimensional tolerances and springback, as well as tearing and buckling of the strip, have to be considered. The rolls generally are made of carbon steel or of gray iron, and they may be chromium plated for a better surface finish of the formed product and for better wear resistance of the rolls. Lubricants may be used to reduce roll wear, to improve surface finish, and to cool the rolls and the sheet being formed.

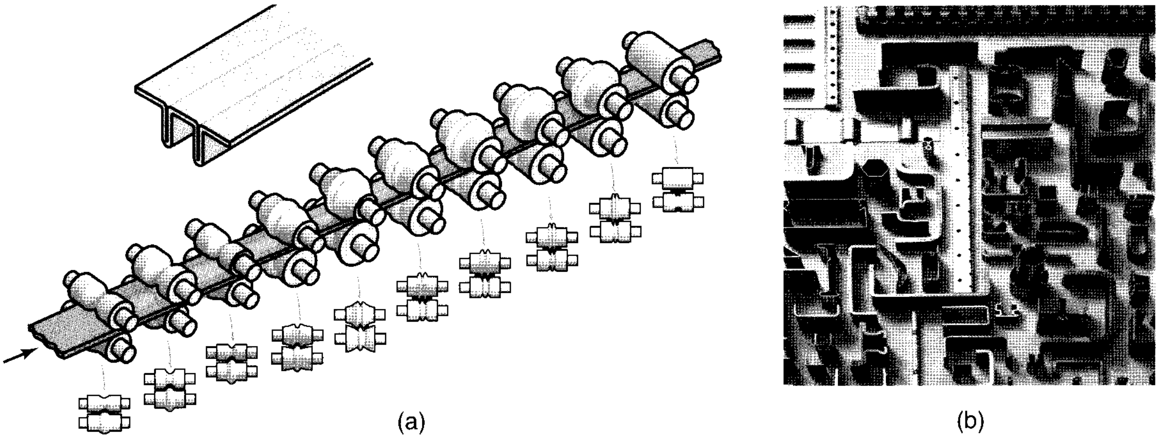


FIGURE 16.26 (a) Schematic illustration of the roll-forming process. (b) Examples of roll-formed cross sections. *Source*: (b) Courtesy of Sharon Custom Metal Forming, Inc.

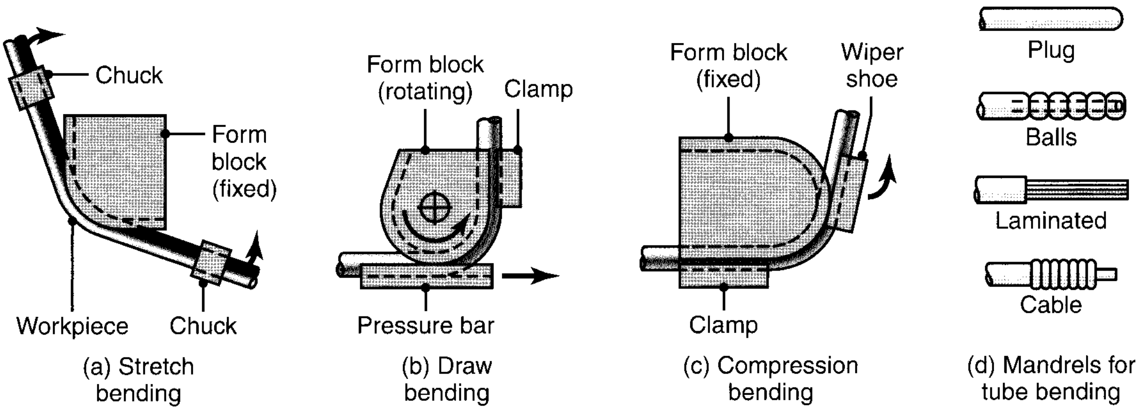


FIGURE 16.27 Methods of bending tubes. Internal mandrels or filling of tubes with particulate materials such as sand are often necessary to prevent collapse of the tubes during bending. Tubes also can be bent by a technique in which a stiff, helical tension spring is slipped over the tube. The clearance 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.

Tube Bending and Forming. Bending and forming tubes and other hollow sections requires special tooling because of the tendency for buckling and folding, as one notes when trying to bend a piece of copper tubing or even a plastic soda straw. The oldest method of bending a tube or pipe is to first pack its inside with loose particles (commonly sand) and then bend it into a suitable fixture. The function of the filler is to prevent the tube from buckling inward. After the tube has been bent, the sand is shaken out. Tubes also can be plugged with various flexible internal mandrels (Fig. 16.27) for the same purpose as the sand. Note that (because of its lower tendency for buckling) a relatively thick tube to be formed to a large bend radius can be bent safely without the use of fillers or plugs. (See also *tube hydroforming*, Section 16.8.)

Dimpling, Piercing, and Flaring. In *dimpling* (Fig. 16.25b), a hole first is punched and then expanded into a flange. Flanges also may be produced by *piercing* with 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), the process is called *flaring*. The condition of the edges (see Fig. 16.3) is important in these operations. 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 be a tendency for cracking along the outer periphery of the flange. To reduce this possibility, sheared or punched edges may be shaved off with a sharp tool (see 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 rollers for watertight and airtight joints, such as are needed 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. 16.28a). The punch is then retracted, the plug returns to its original shape (by total elastic recovery), and the formed part is removed by opening the split dies. Typical products made are coffee or water pitchers, beer barrels, and beads on oil drums. For parts with complex shapes, the plug (instead of being cylindrical) may be shaped in order 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; furthermore, they do not damage the surface finish of the part being formed.

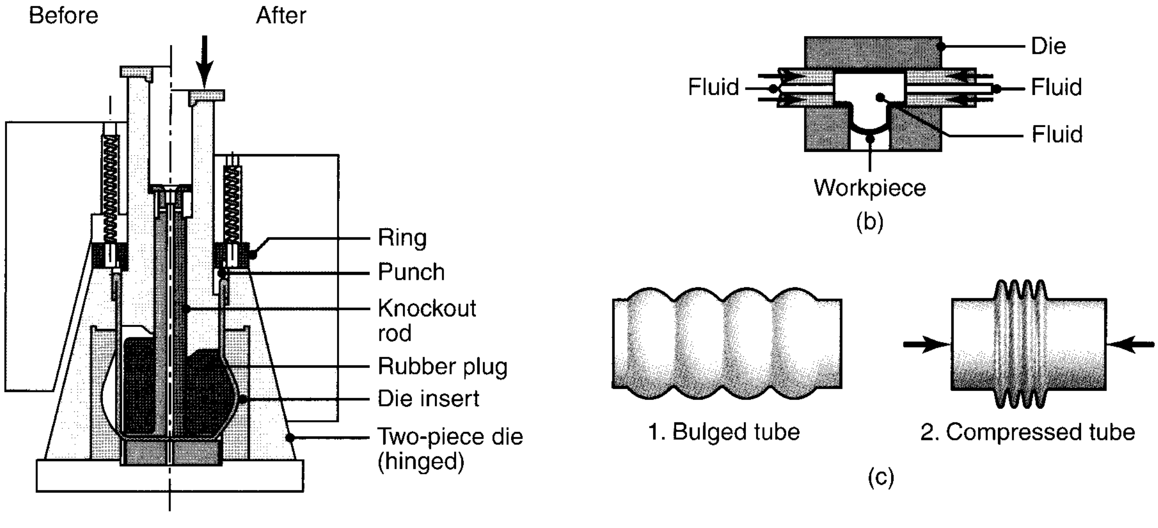


FIGURE 16.28 (a) The bulging of a tubular part with a flexible plug. Water pitchers can be made by this method. (b) Production of fittings for plumbing by expanding tubular blanks under internal pressure. The bottom of the piece is then punched out to produce a "T." *Source*: After J.A. Schey. (c) Steps in manufacturing bellows.

Segmented Dies. These dies consist of individual segments that are placed inside the part to be formed and expanded mechanically in a generally radial direction. They are then retracted to remove the formed part. Segmented dies are relatively inexpensive, and they can be used for large production runs.

Stretch Forming. In *stretch forming*, the sheet metal is clamped along its edges and then stretched over a male die (*form block* or *form punch*). The die moves upward, downward, or sideways, depending on the particular design of the machine (Fig. 16.29). 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 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 generally is used for low-volume production, it is versatile and economical, particularly for 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 width. Controlling the amount of stretching is important in order to prevent tearing. Stretch forming cannot produce parts with sharp contours or with reentrant corners (depressions on the surface of the die). Various 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 generally are made of zinc alloys, steel, plastics, or wood. Most applications require little or no lubrication.

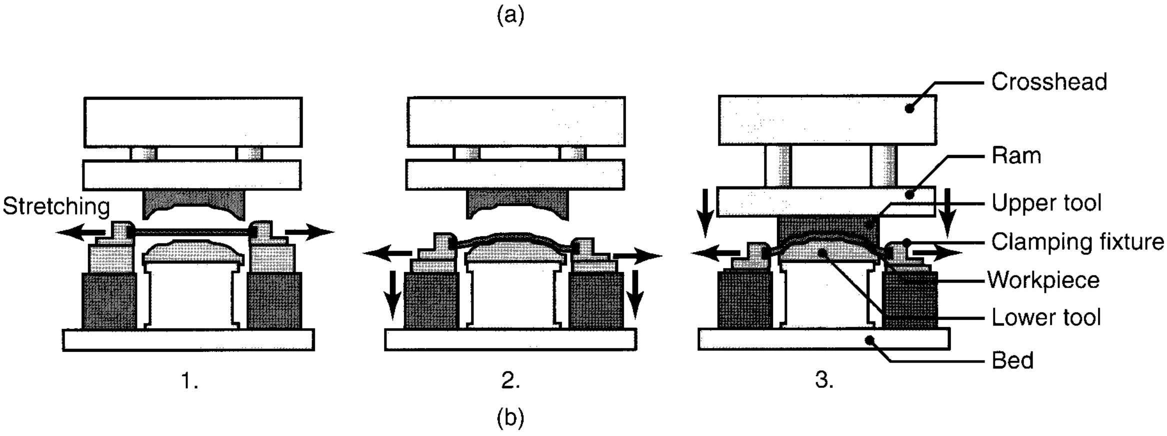


FIGURE 16.29 Schematic illustration of a stretch-forming process. Aluminum skins for aircraft can be made by this method. *Source*: (a) Courtesy of Cyril Bath Co.

16.7 Deep Drawing

Numerous parts made of sheet metal are cylindrical or box shaped, such as pots and pans, all types of containers for food and beverages (Fig. 16.30), stainless-steel kitchen sinks, canisters, and automotive fuel tanks. Such parts usually are made by a process in which a punch forces a flat sheet-metal blank into a die cavity (Fig. 16.31a). Although the process generally is called *deep drawing* (because of its capability for producing deep parts), it also is used to make parts that are shallow or have moderate depth. It is one of the most important metalworking processes because of its widespread use.

In the basic deep-drawing process, 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. 16.31b). The punch travels downward and forces the blank into the die cavity, forming a cup. The important variables in deep drawing are the properties of the sheet metal, the ratio of blank diameter, D_o; the punch diameter, D_p; the clearance, c,

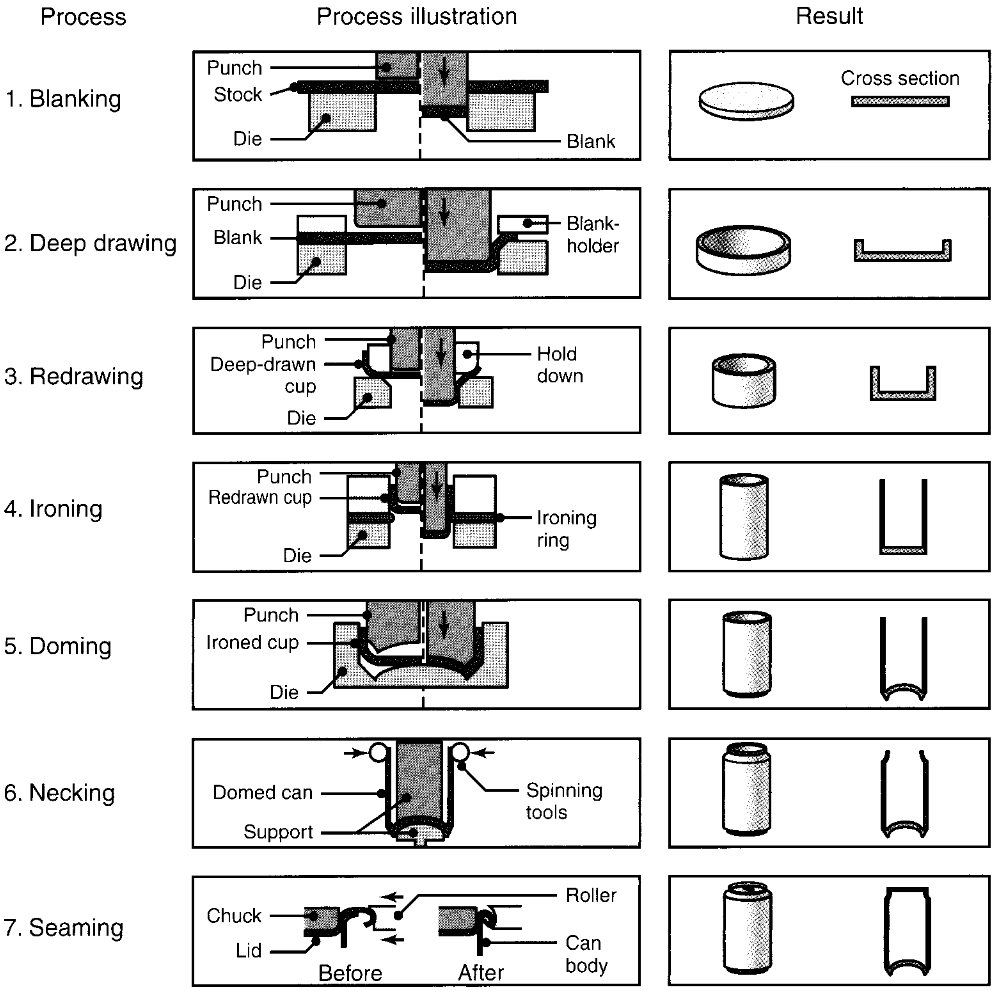


FIGURE 16.30 The metal-forming processes involved in manufacturing a two-piece aluminum beverage can.

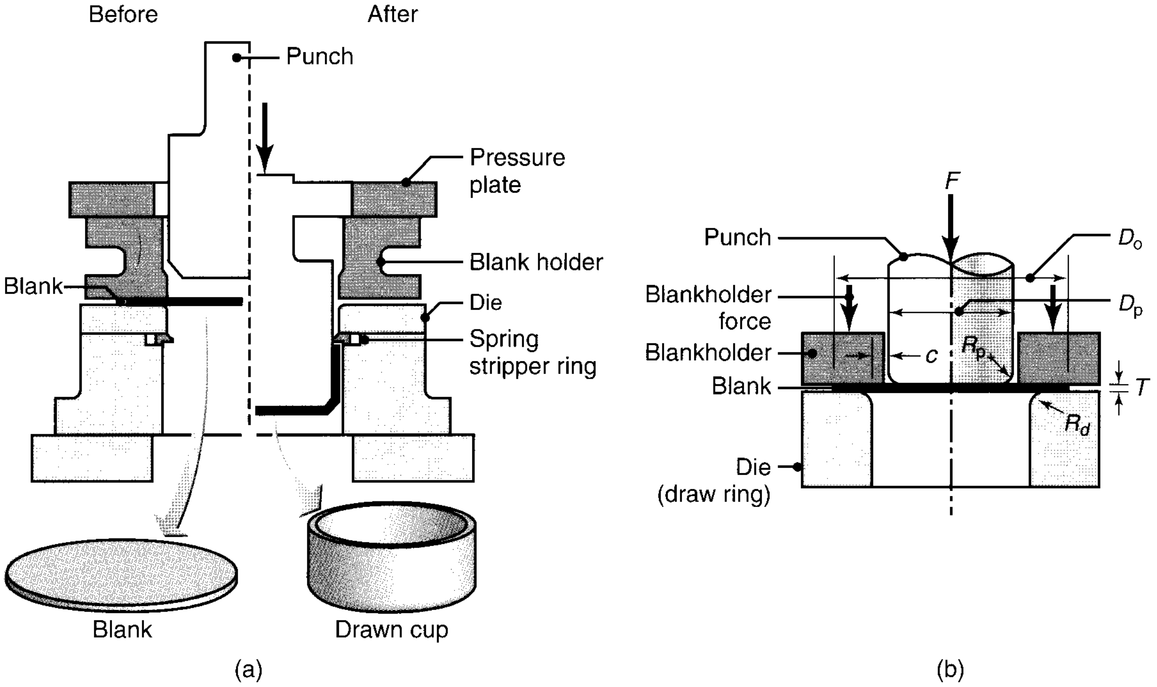


FIGURE 16.31 (a) Schematic illustration of the deep-drawing process on a circular sheet-metal blank. The stripper ring facilitates the removal of the formed cup from the punch. (b) Process variables in deep drawing. Except for the punch force, F, all the parameters indicated in the figure are independent variables.

between punch and die; the punch radius, R_p; the die-corner radius, R_d; the blankholder force; and friction and lubrication between all contacting surfaces.

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, such as a drinking glass. Wrinkling can be reduced or eliminated if a blankholder is loaded by a certain force. In order to improve performance, the magnitude of this force can be controlled as a function of punch travel.

Because of the many 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

$$F_{\max} = \pi D_p T(\mathrm{UTS})\left[\left(\frac{D_o}{D_p}\right) - 0.7\right], \tag{16.9}$$

where the nomenclature is the same as that in Fig. 16.31b. 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 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, can cause *tearing* of the cup.

16.7.1 Deep Drawability

In a deep-drawing operation, failure generally results from the *thinning* of the cup wall under high longitudinal tensile stresses. If we follow the movement of the material as it flows into the die cavity, it can be seen that the sheet metal (a) must be capable of undergoing a reduction in width due to a reduction in diameter and (b) must also resist thinning under the longitudinal tensile stresses in the cup wall. *Deep drawability* generally is expressed by the **limiting drawing ratio** (LDR) as

$$\text{LDR} = \frac{\text{Maximum blank diameter}}{\text{Punch diameter}} = \frac{D_o}{D_p}. \qquad (16.10)$$

Whether a sheet metal can be deep drawn successfully into a round cup-shaped part 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 the specimen undergoes in tension (Fig. 16.32):

$$R = \frac{\text{Width strain}}{\text{Thickness strain}} = \frac{\varepsilon_w}{\varepsilon_t}. \qquad (16.11)$$

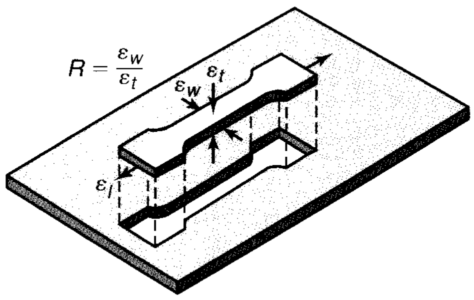


FIGURE 16.32 Strains on a tensile-test specimen removed from a piece of sheet metal. These strains are used in determining the normal and planar anisotropy of the sheet metal.

In order to determine the magnitude of R, a tensile-test specimen is first prepared and subjected to an elongation of 15 to 20%. The true strains that the sheet undergoes are calculated in the manner discussed in Section 2.2. Because cold-rolled sheets generally have anisotropy 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. For this condition, an average value, R_{avg}, is calculated from the equation

$$R_{\text{avg}} = \frac{R_0 + 2R_{45} + R_{90}}{4}, \qquad (16.12)$$

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.

The experimentally determined relationship between R_{avg} and the limiting drawing ratio is shown in Fig. 16.33. It has been established that no other mechanical property of sheet metal shows as consistent a relationship to 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 phenomenon called *earing* (Fig. 16.34). Ears are objectionable on deep-drawn cups because they have to be trimmed off, as they serve no useful purpose and interfere with further processing of the cup, resulting in scrap. Earing is caused by the **planar anisotropy**

TABLE 16.4

Typical Ranges of Average Normal Anisotropy, R_{avg}, for Various Sheet Metals	
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 (α)	3.0–5.0
Stainless steels	0.9–1.2
High-strength, low-alloy steels	0.9–1.2

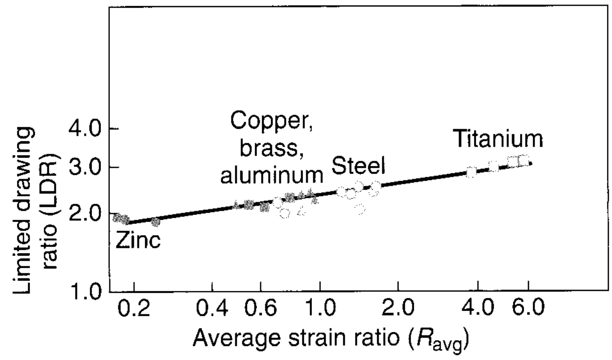


FIGURE 16.33 The relationship between average normal anisotropy and the limiting drawing ratio for various sheet metals. *Source*: After M. Atkinson.

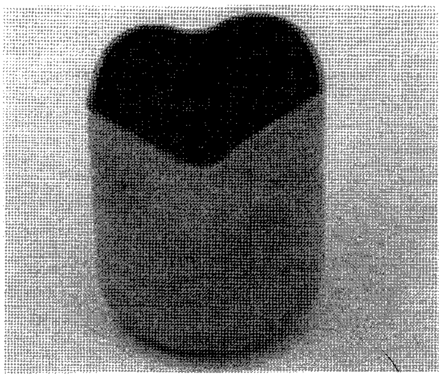


FIGURE 16.34 Earing in a drawn steel cup, caused by the planar anisotropy of the sheet metal.

of the sheet, and the number of ears produced may be two, four, or eight, depending on the processing history and microstructure of the sheet. If the sheet is stronger in the 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 is indicated by ΔR. It 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. 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. Sheet-metal textures are being developed continually to improve drawability by controlling the type of alloying elements in the material as well as various processing parameters during rolling of the sheet.

16.7.2 Deep-drawing Practice

Certain guidelines have been established 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. On the other hand, if the blankholder force is too low, wrinkling will occur.

Clearances are usually 7 to 14% greater than sheet thickness. If the clearance is 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 at the corners; if they are too large, the cup wall may wrinkle—a phenomenon called *puckering*.

Draw beads (Fig. 16.35) often are necessary to control the flow of the blank into the die cavity. Beads restrict the free flow of the sheet metal by bending and unbending it during the drawing cycle, thereby increasing the force required to pull the sheet into the die cavity. This phenomenon can be demonstrated simply by placing a strip of paper or aluminum foil through one's fingers in an arrangement similar to that shown in Fig. 13.35a. Note that a certain force is now required to pull the strip

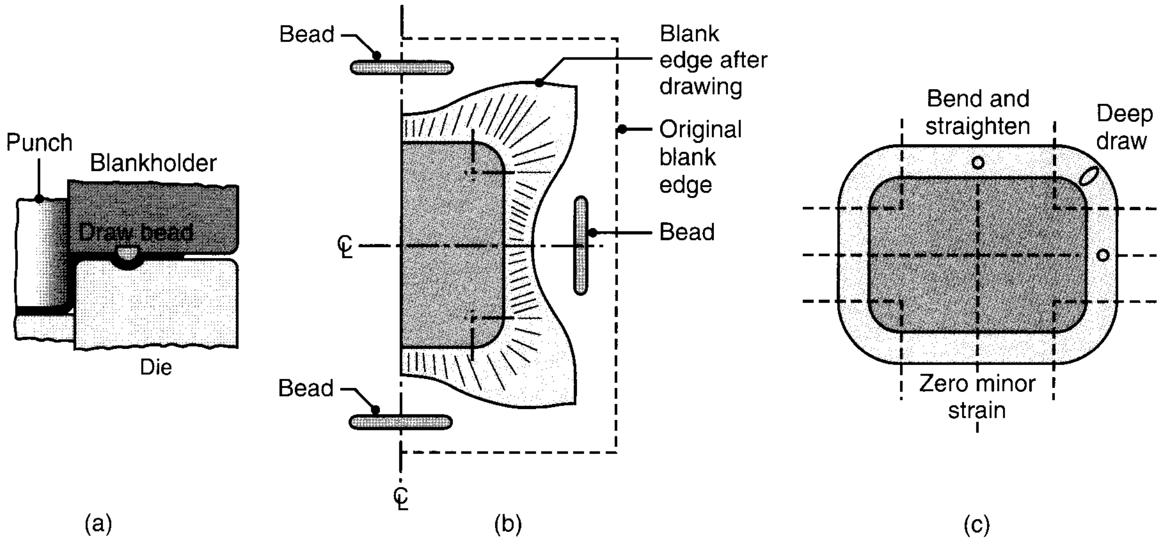


FIGURE 16.35 (a) Schematic illustration of a draw bead. (b) Metal flow during the drawing of a box-shaped part while using beads to control the movement of the material. (c) Deformation of circular grids in the flange in deep drawing.

through your fingers. Draw beads also help to reduce the required blankholder forces, because the beaded sheet has a higher stiffness (due to its higher moment of inertia) and, hence, a lower tendency to wrinkle. Draw-bead diameters may range from 13 to 20 mm—the latter applicable to large stampings, such as automotive panels.

Draw beads also are useful in drawing *box-shaped* and *nonsymmetric* parts, because they can present significant difficulties in practice (Figs. 16.35b and c). Note in Fig. 16.35c, for example, that various regions of the part undergo different types of deformation during drawing. (Recall also the fundamental principle that the material flows in the direction of least resistance.)

In order to avoid tearing of the sheet metal during forming, it often is necessary to incorporate the following factors:

- Proper design and location of draw beads
- Large die radii
- Effective lubrication
- Proper blank size and shape
- The cutting off of corners of square or rectangular blanks at 45° to reduce tensile stresses developed during drawing
- The use of blanks free of internal and external defects.

Ironing. Note in Fig. 16.31 that 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. The reason for this is that the cup rim consists of material from the outer diameter of the blank; hence, it has been reduced in diameter more (and thus becomes thicker) than the material constituting the rest of the cup wall. As a result, the cup will develop a nonuniform wall thickness. The thickness of the cup wall can be controlled by a process called *ironing*, in which a drawn cup is pushed through one or more ironing rings (see Fig. 16.30). The clearance between the ironing rings and the punch is less than the cup wall thickness, so the cup after ironing has essentially a constant wall thickness (that is equal to the clearance, except for some small elastic recovery). Aluminum beverage cans, for example, typically undergo two or three ironing operations in one stroke in which the drawn cup is pushed through a set of ironing rings.

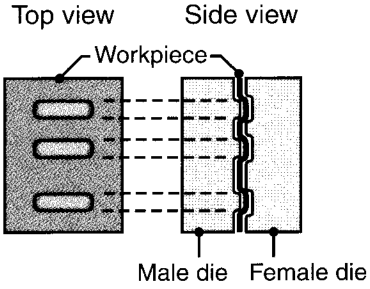


FIGURE 16.36 An embossing operation with two dies. Letters, numbers, and designs on sheet-metal parts can be produced by this process.

Redrawing. Containers that are too difficult to draw in one operation generally undergo *redrawing* (see Fig. 16.30). Because of the volume constancy of the metal, the cup becomes longer as it is redrawn to smaller diameters. In *reverse redrawing*, the cup is placed upside down in the die and thus is subjected to bending in the direction opposite to its original configuration.

Drawing without Blankholder. Deep drawing also may be carried out successfully without a blankholder, provided that the sheet metal is sufficiently thick to prevent wrinkling. A typical range of the diameter is

$$D_o - D_p < 5T, \tag{16.14}$$

where T is the sheet thickness. The dies are contoured specially for this operation.

Embossing. This is an operation consisting of shallow or moderate draws made with male and female matching shallow dies (Fig. 16.36). Embossing is used principally for the stiffening of flat sheet-metal panels and for purposes of decorating, numbering, and lettering, such as letters on the lids of aluminum beverage cans.

Tooling and Equipment for Drawing. The most common tool and die materials for deep drawing are tool steels and cast irons and include dies produced from ductile-iron castings made by the lost-foam process. Other materials, such as carbides and plastics, also may be used (see 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 manufactured on equipment such as high-speed machining on computer-controlled lathes.

The equipment for deep drawing is usually a *double-action hydraulic press* or a *mechanical press*, the latter generally being favored because of its higher operating speed. In the double-action hydraulic press, the punch and the blankholder are controlled independently. Punch speeds generally range between 0.1 and 0.3 m/s.

CASE STUDY 16.1 Manufacturing of Food and Beverage Cans

Can manufacturing is a major and competitive industry worldwide, 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 15 g), and they are under an internal pressure of 620 kPa—reliably and without leakage of their contents. There are stringent requirements for the surface finish of the can, 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. Thus, the cost of empty cans alone in a six-pack is 24 cents, which also indicates the importance of recycling cans.

Food and beverage cans may be produced in a number of ways, the most common ones being two-piece and three-piece cans. Two-piece cans consist of the can body and the lid (Fig. 16.37a). The body is made of one piece that has been drawn and ironed—hence the industry practice of referring 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.

Drawn and ironed can bodies are produced from a number of alloys, but the most common are

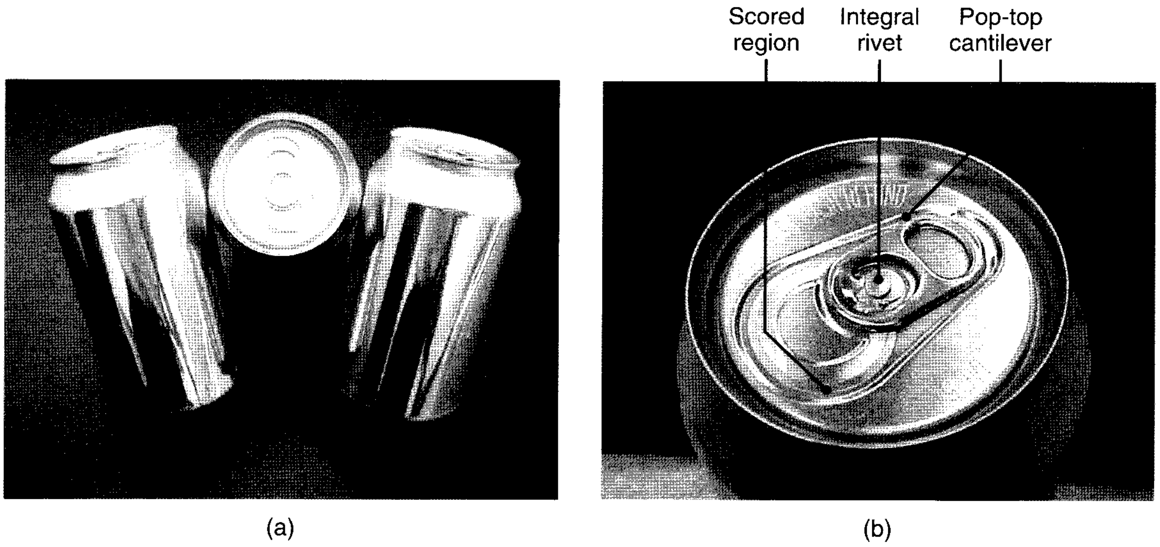


FIGURE 16.37 (a) Aluminum beverage cans. Note the excellent surface finish. (b) Detail of the can lid, showing the integral rivet and scored edges for the pop-top.

3004-H19 aluminum (see Section 6.2) and electrolytic tin-plated ASTM A623 steel. Aluminum lids are used for both steel and aluminum cans and are produced from 5182-H19 or 5182-H48. The lid presents a demanding set of design requirements, as can be appreciated by reviewing Fig. 16.37b. Not only must the can lid be scored easily (curved grooves around the tab), but an integral rivet is formed and headed 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 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 placed around the periphery of the lid. This polymer layer seals the can's contents after the lid is seamed to the can body, as described next.

The traditional method of manufacturing the can bodies is shown in Fig. 16.30. The process starts with 140-mm-diameter blanks produced from rolled sheet stock. These blanks are (a) deep drawn to a diameter of around 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 the can bottom. The deep-drawing and ironing operations are performed in a special type of press that typically produces 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 through spinning (Section 16.9) or by die necking (a forming operation similar to that shown in Fig. 15.19a, where a thin-walled tubular part is pushed into the die), and it is then spin-flanged. The reason for necking the can top is that the 5182 aluminum used for the lid is relatively expensive. Thus, by tapering the top of the can, a smaller volume of material is needed, thereby reducing the cost. Also, it should be noted that the cost of a can often is calculated to millionths of a dollar hence any design feature that reduces its cost will be exploited by this competitive industry.

Source: Courtesy of J.E. Wang, Texas A&M University.

16.8 Rubber Forming and Hydroforming

In the processes described in the preceding sections, it has been noted that the dies generally are made of solid materials, such as steels and carbides. However, in *rubber forming* (also known as the *Guerin process*), one of the dies in a set is made of a flexible material, typically a polyurethane membrane. Polyurethanes are used widely because of their abrasion resistance, fatigue life, and resistance to cutting or tearing.

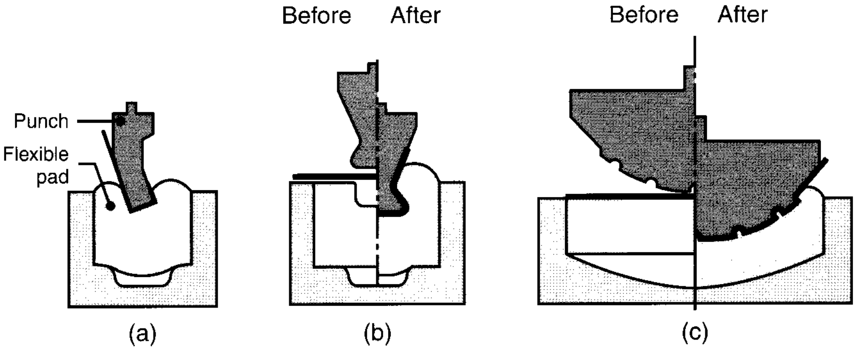


FIGURE 16.38 Examples of the bending and embossing of sheet metal with a metal punch and with a flexible pad serving as the female die. *Source*: Courtesy of Polyurethane Products Corporation.

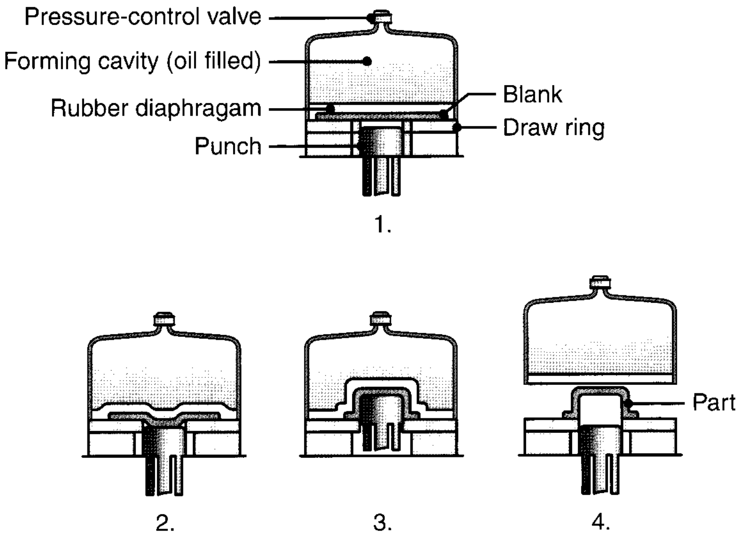


FIGURE 16.39 The hydroform (or fluid-forming) process. Note that, in contrast to the ordinary deep-drawing process, the pressure in the dome forces the cup walls against the punch. The cup travels with the punch; in this way, deep drawability is improved.

In the bending and embossing of sheet metal by this process, the female die is replaced with a rubber pad (Fig. 16.38). Note that the outer surface of the sheet is protected from damage or scratches, because it is not in contact with a hard metal surface during forming. Pressures in rubber forming are typically on the order of 10 MPa.

In the **hydroform**, or *fluid-forming process* (Fig. 16.39), the pressure over the rubber membrane is controlled throughout the forming cycle with a maximum pressure of up to 100 MPa. This procedure allows close control of the part during forming and prevents wrinkling or tearing. Deeper draws are obtained than in 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, which then reduces the longitudinal tensile stresses in the cup and thus delays fracture.

The control of frictional conditions in rubber forming as well as other sheet-forming operations can be a critical factor in making parts successfully. The use of proper lubricants and their application methods is also important.

In *tube hydroforming* (Fig. 16.40), metal tubing is formed in a die and pressurized internally by a fluid, usually water. This process, which now is being applied more widely, can form simple tubes as well as various intricate hollow shapes (Fig. 16.40b). Parts made by the process include automotive-exhaust and tubular structural components.

When selected properly, rubber-forming and hydroforming processes have the advantages of (a) the capability to form complex shapes, (b) forming parts with laminated sheets of various materials and coatings, (c) flexibility and ease of operation, (d) the avoidance of damage to the surfaces of the sheet, (e) low die wear, and (f) low tooling cost.

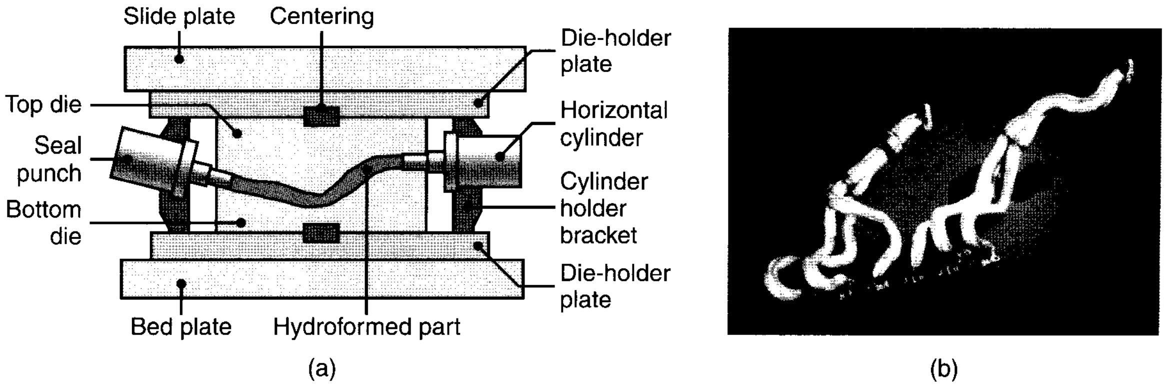


FIGURE 16.40 (a) Schematic illustration of the tube-hydroforming process. (b) Example of tube-hydroformed parts. Automotive-exhaust and structural components, bicycle frames, and hydraulic and pneumatic fittings are produced through tube hydroforming. *Source*: Courtesy of Schuler GmBH.

CASE STUDY 16.2 Tube Hydroforming of an Automotive Radiator Closure

The conventional assembly used to support an automotive radiator, or radiator closure, is constructed through the stamping of components that are subsequently welded together. To simplify the design and to achieve weight savings, a hydroformed assembly was designed, as shown in Fig. 16.41. Note that the design uses varying cross sections, an important design feature to reduce weight and to provide surfaces to facilitate assembly and mounting of the radiator.

Tube hydroforming is a versatile process that is capable of producing complex shapes. A typical tube hydroforming processing sequence consists of the following steps:

1. Bending of tube to the desired configuration
2. Tube hydroforming to achieve the desired shape
3. Finishing operations, such as end shearing and inspection
4. Assembly, including welding of components.

The sequence of operations on one of the tube components of the closure is shown in Fig. 16.42.

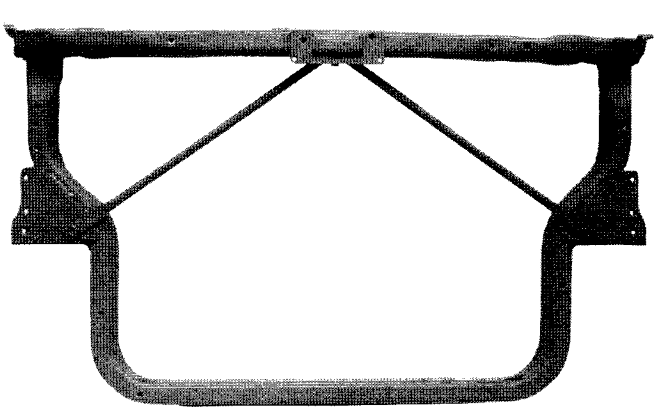


FIGURE 16.41 Hydroformed automotive radiator closure.

(continued)

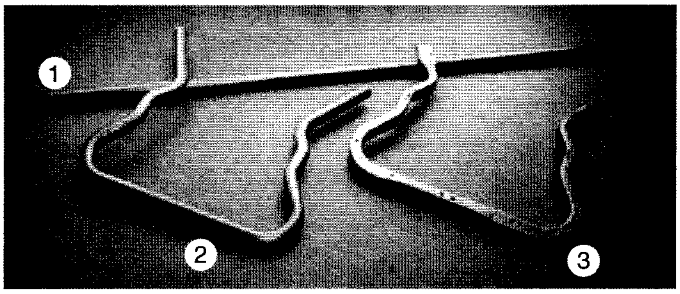
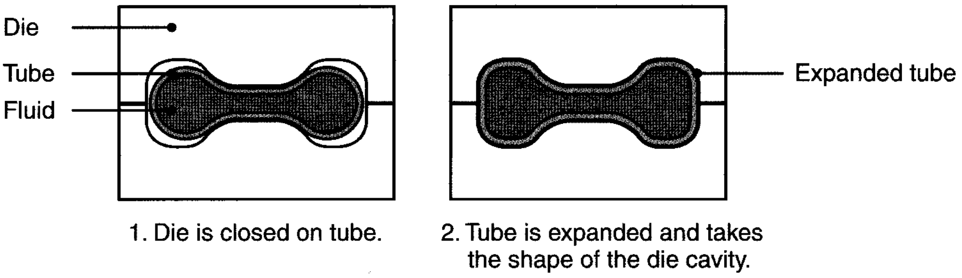
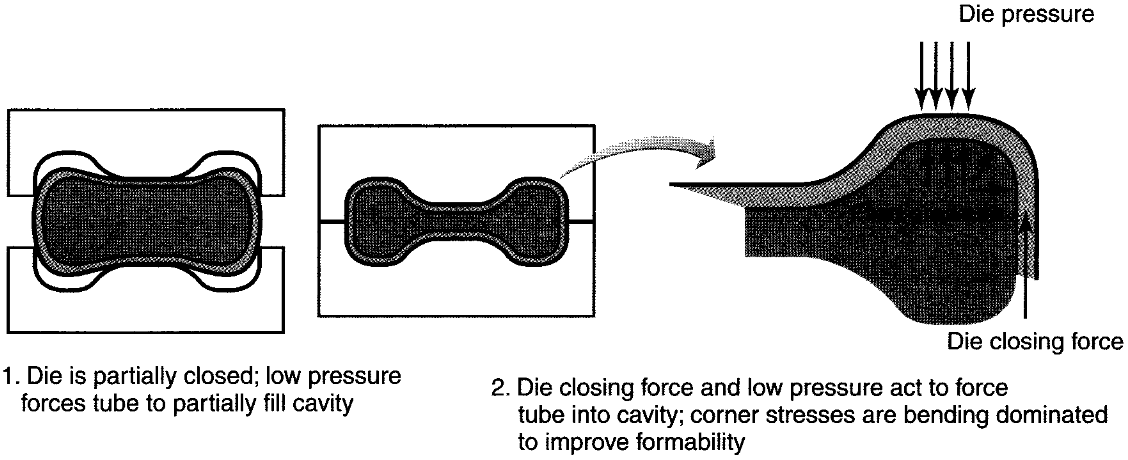


FIGURE 16.42 Sequence of operations in producing a tube-hydroformed component: (1) tube as cut to length; (2) after bending; (3) after hydroforming.



(a) Conventional hydroforming



(b) Pressure sequence hydroforming

FIGURE 16.43 Schematic illustration of expansion of a tube to a desired cross section through (a) conventional hydroforming and (b) pressure sequence hydroforming.

The tube, constructed of steel with a 300 MPa yield strength, is bent to shape (see Fig. 16.27). The bent tube is then placed in a hydroforming press and 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 16.43a shows a typical cross section. Note that as the tube is

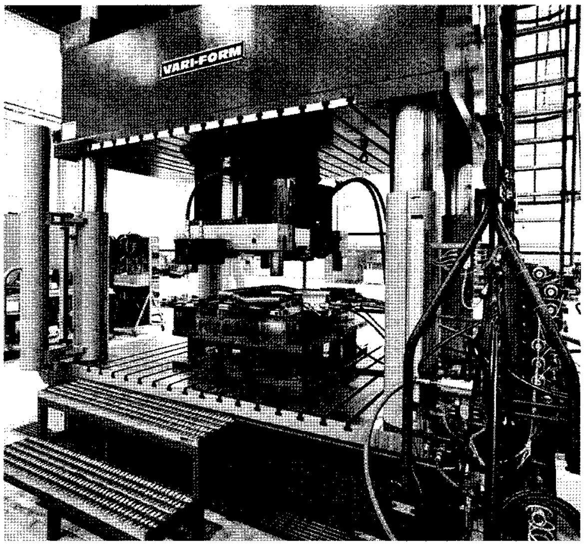


FIGURE 16.44 View of the tube-hydroforming press, with bent tube in place in the forming die.

expanded, there is significant wall thinning, especially at the corners, because of friction at the tube–die interface. The pressure sequence hydroforming process is therefore used on this part, as shown in Fig. 16.43b. 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 form the cross-sectional corners. After the die is completely closed, the internal pressure is increased to lock in the form and provide support for hole piercing. This sequence has the benefit of forming the sharp corners in the cross section by bending, as opposed to pure stretching in conventional hydroforming. The resulting wall thickness is much more uniform, producing a more structurally sound component. Figure 16.44 shows a part being hydroformed.

The assembly shown 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 (see Section 30.4.3) and with threaded fasteners to aid in 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 is increased by as much as 150% in some directions. In addition, the cooling system surface area is increased by 43% in the new design.

Source: Courtesy of B. Longhouse, Vari-Form, Inc.

16.9 Spinning

Spinning is a process that involves the forming of axisymmetric parts over a mandrel by the use of various tools and rollers—a process is similar to that of forming clay on a potter's wheel.

Conventional Spinning. In *conventional spinning*, a circular blank of flat or preformed sheet metal is placed and held against a mandrel and rotated while a rigid tool deforms and shapes the material over the mandrel (Fig. 16.45a). The tool may be activated either manually or (for higher production rates) by computer-controlled mechanisms. The process typically involves a sequence of passes, and it requires considerable skill. Conventional spinning is suitable particularly for conical and

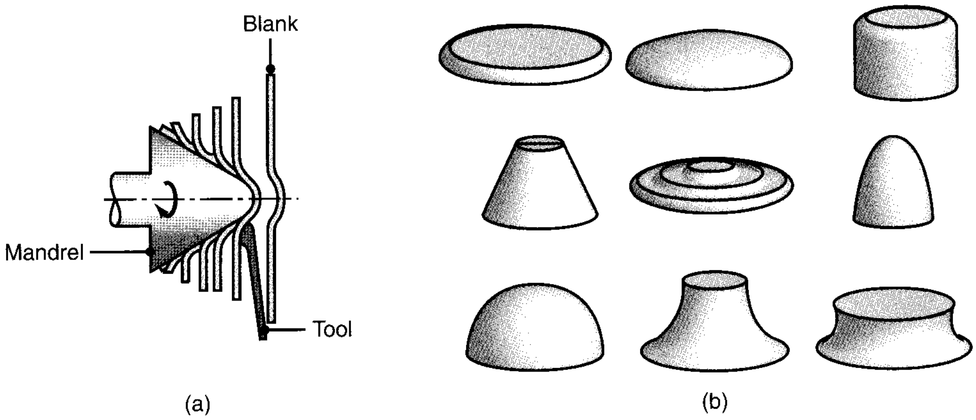


FIGURE 16.45 (a) Schematic illustration of the conventional spinning process. (b) Types of parts conventionally spun. All parts are axisymmetric.

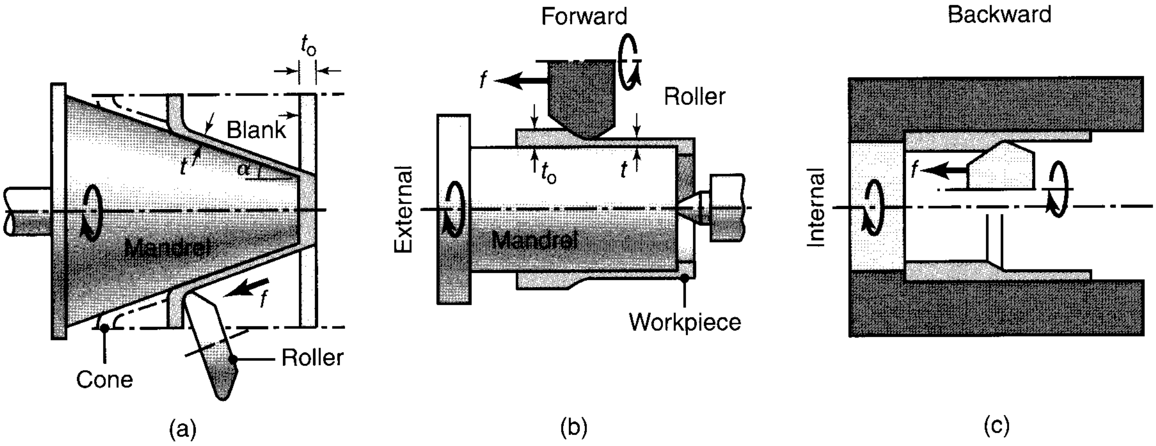


FIGURE 16.46 (a) Schematic illustration of the shear-spinning process for making conical parts. The mandrel can be shaped so that curvilinear parts can be spun. (b) and (c) Schematic illustrations of the tube-spinning process.

curvilinear shapes (Fig. 16.45b), which otherwise would be difficult or uneconomical to produce. Part diameters range up to 6 m. Although most spinning is performed at room temperature, thick parts and metals with high strength or low ductility require spinning at elevated temperatures.

Shear Spinning. Also known as *power spinning*, *flow turning*, *hydrospinning*, and *spin forging*, this operation produces an axisymmetric conical or curvilinear shape, reducing the sheet's thickness while maintaining its maximum (blank) diameter (Fig. 16.46a). A single forming roller can be used, but two rollers are preferable in order to balance the 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. This operation wastes little material, 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 in this process generally is defined as the maximum reduction in thickness to which a part can be subjected by spinning 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 just one spinning pass. For metals with low ductility, the operation is carried out at elevated temperatures by heating the blank in a furnace and transferring it rapidly to the mandrel.

Tube Spinning. In *tube spinning*, the thickness of hollow, cylindrical blanks is reduced or shaped by spinning them on a solid, round mandrel using rollers (Fig. 16.46). The reduction in wall thickness results in a longer tube. This operation may be carried out externally or internally; thus, various external and internal profiles can be produced from cylindrical blanks with constant wall thickness. The parts may be spun *forward* or *backward*; this nomenclature is similar to that of direct and indirect extrusion, as described in Section 15.2. The maximum thickness reduction per pass in tube spinning is related to the tensile reduction of area of the material, as it is in shear spinning. Tube spinning can be used to make rocket, missile, and jet-engine parts, pressure vessels, and automotive components, such as car and truck wheels.

Incremental Forming. *Incremental forming* is a term applied to a class of processes that are related to conventional metal spinning. The simplest version is *incremental stretch expanding*, shown in Fig. 16.47, wherein a rotating blank is deformed by a steel rod with a smooth hemispherical tip to produce axisymmetric parts. No special tooling or mandrel is used, and the motion of the rod determines the final part shape in one or more passes. The strain distribution within the workpiece depends on the tool path across the part profile, and proper lubrication is essential.

CNC incremental forming uses a CNC machine tool (see Section 37.3) that is programmed to follow contours at different depths across the sheet-metal surface. In this arrangement, the blank is clamped and is stationary, and the tool rotates to assist forming. 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. 20.2). Figure 16.47b depicts an example of a part that has been produced by CNC incremental forming. Note that the part does not have to be axisymmetric.

The main advantages of incremental forming are low tooling costs and high flexibility in the shapes that can be produced. CNC incremental forming has been

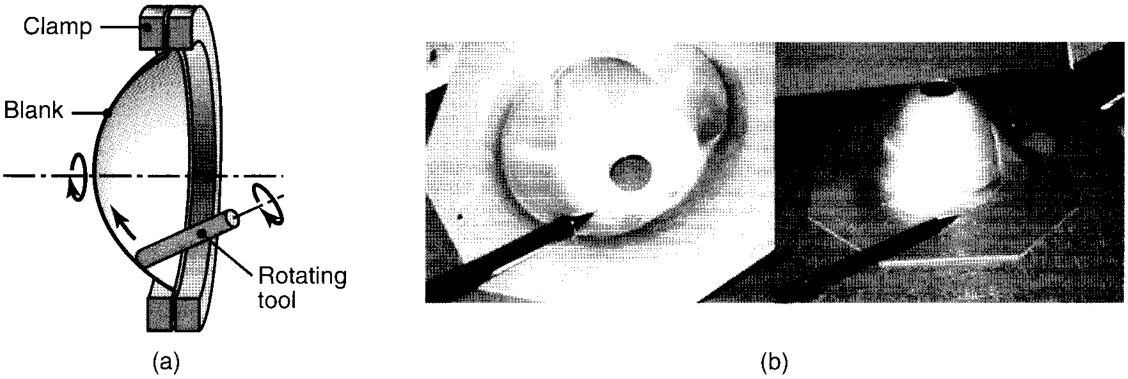


FIGURE 16.47 (a) Illustration of an incremental-forming operation. Note that no mandrel is used and that the final part shape depends on the path of the rotating tool. (b) An automotive headlight reflector produced through CNC incremental forming. Note that the part does not have to be axisymmetric. *Source*: After J. Jeswiet, Queen's University, Ontario.

used for rapid prototyping of sheet-metal parts because the lead times associated with hard tooling are not necessary. The main drawbacks to incremental forming include low production rates and limitations on materials that can be formed.

16.10 Superplastic Forming

The superplastic behavior of certain metals and alloys was described in Section 2.2.7, where tensile elongations on the order of 2000% were obtained within certain temperature ranges. Common examples of such materials are zinc–aluminum and titanium alloys, which have very fine grains—typically less than 10 to 15 μm (see Table 1.1). Superplastic alloys can be formed 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, to be described in Chapter 19). The behavior of the material in superplastic forming is similar to that of bubble gum or hot glass, which, when blown, expands many times its original diameter before it bursts.

Superplastic alloys, particularly Zn-22Al and Ti-6Al-4V, also can be formed by bulk-deformation processes, including closed-die forging, coining, hubbing, and extrusion. Commonly used die materials in superplastic forming are low-alloy steels, cast tool steels, ceramics, graphite, and plaster of paris. 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 out of one piece, with fine detail, close tolerances, and elimination of secondary operations.
- Weight and material savings can be realized because of the good formability of the materials.
- Little or no residual stresses develop in the formed parts.
- Because of the low strength of the material at forming temperatures, the tooling can be made of materials that have lower strength than those in other metalworking processes; hence, tooling costs are lower.

On the other hand, superplastic forming has the following limitations:

- The material must not be superplastic at service temperatures; otherwise the part will undergo shape changes.
- Because of the high strain-rate sensitivity of the superplastic material, 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; thus, cycle times are much longer than those of conventional forming processes. Consequently, superplastic forming is a batch-forming process.

Diffusion Bonding/Superplastic Forming. Fabricating complex sheet-metal structures by combining *diffusion bonding* with *superplastic forming* (SPF/DB) is an important process, particularly in the aerospace industry. Typical structures made are shown in Fig. 16.48, in which flat sheets are diffusion bonded (see Section 31.7) and formed. In this process, selected locations of the sheets are first diffusion bonded while the rest remains unbonded, using a layer of material (*stop-off*) to prevent bonding. The structure is then expanded in a mold, typically by using pressurized

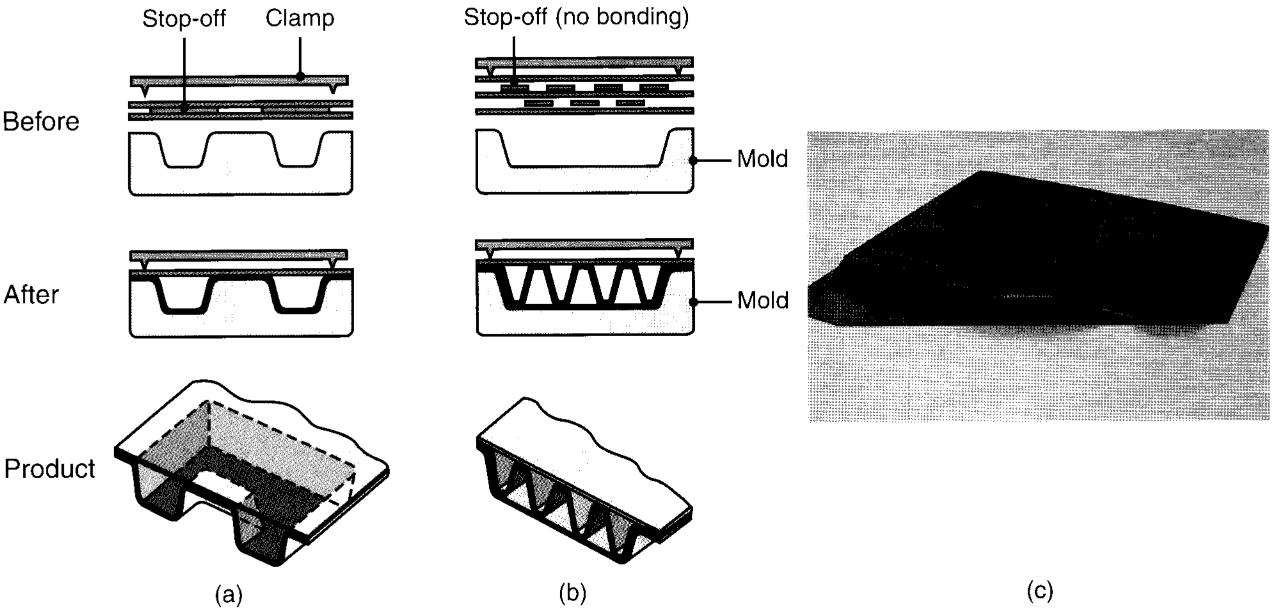


FIGURE 16.48 Types of structures made by superplastic forming and diffusion bonding of sheet metals. Such structures have a high stiffness-to-weight ratio. *Source*: (a) and (b) Courtesy of Rockwell International Corp., (c) Courtesy of Triumph Group, Inc.

neutral (argon) gas, thus taking the shape of the mold. These structures have high stiffness-to-weight ratios because they are thin, and by design, they have high section moduli. This important feature makes SPF/DB particularly attractive in aircraft and aerospace applications.

The SPF/DB process improves productivity by eliminating mechanical fasteners, and it 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 Specialized Forming Processes

Although not as commonly used as the other processes covered thus far, several other sheet-forming processes are used for specialized applications.

Explosive Forming. Explosives generally are used for demolition in construction, in road building, and for many destructive purposes. However, controlling their quantity and shape makes it possible to use explosives as a source of energy for sheet-metal forming. In *explosive forming*, first utilized to form metals in the early 1900s, the sheet-metal blank is clamped over a die and the entire assembly is lowered into a tank filled with water (Fig. 16.49a). The air in the die cavity is then evacuated, an explosive charge is placed at a certain height, and the charge is detonated.

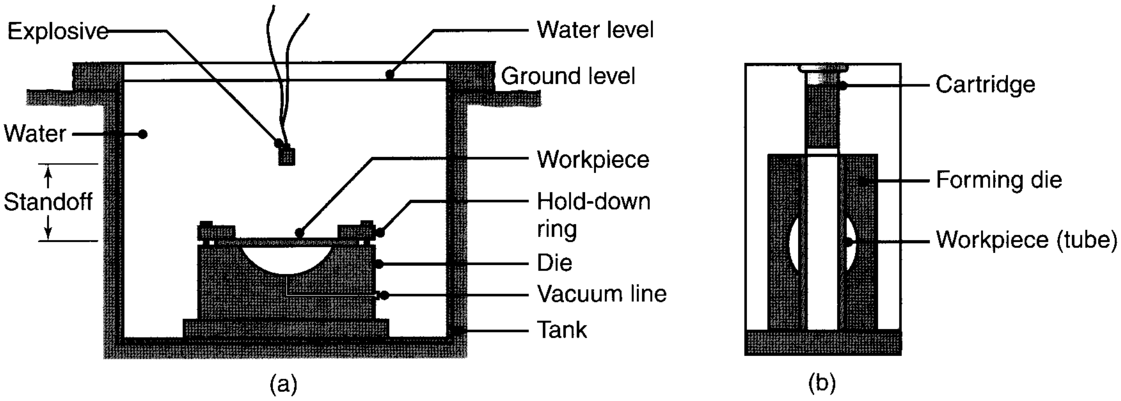


FIGURE 16.49 (a) Schematic illustration of the explosive-forming process. (b) Illustration of the confined method of the explosive bulging of tubes.

The explosive generates a shock wave with a pressure that is sufficient to form sheet metals. The *peak pressure*, p, generated in water is given by the expression

$$p = K\left(\frac{\sqrt[3]{W}}{R}\right)^{a}, \tag{16.15}$$

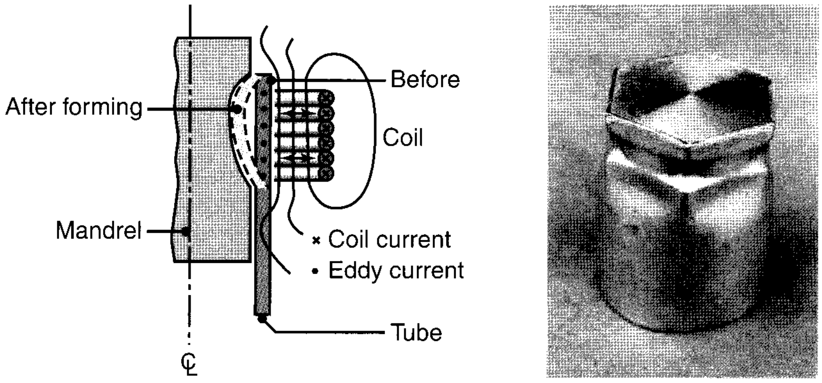
where p is in psi, K is a constant that depends on the type of explosive, such as 21.43×10^{6} for TNT (trinitrotoluene), W is the weight of the explosive in newtons, R is the distance of the explosive from the sheet-metal surface (called the *standoff*) in metres, and a is a constant, generally taken to be 1.15.

A variety of shapes can be formed through explosive forming, provided that the material is ductile at the high rates of deformation 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 plate. 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. 16.49b. In this case, a *cartridge* (canned explosive) is used as the source of energy. The process can be useful in the bulging and expanding of thin-walled tubes for specialized applications.

The mechanical properties of parts made by explosive forming are basically similar to those of others made by conventional forming methods. Depending on the number of parts to be produced, dies may 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 tube is then collapsed by magnetic forces over a solid piece, thus making the assembly an integral part (Fig. 16.50).

The mechanics of this process is based on the fact that a magnetic field produced by the coil (Fig. 16.50a) crosses the metal tube (which is an electrical conductor) and generates *eddy currents* in the tube. In turn, these currents produce their own magnetic field. The forces produced by the two magnetic fields oppose each other. The repelling force generated between the coil and the tube then collapses the



(a) (b)

FIGURE 16.50 (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.

tube over the inner piece. The higher the electrical conductivity of the workpiece, the higher the magnetic forces. It is not necessary for the workpiece material to have magnetic properties.

It has been shown that the basic advantages of this process is that the formability of the material is increased, dimensional accuracy is improved, and springback and wrinkling are reduced. Magnetic coil design is an important consideration for the success of the operation. Flat magnetic coils also can be made for use in operations such as embossing and shallow drawing of sheet metals. The process has been found to be particularly effective for aluminum alloys. First used in the 1960s, electromagnetically assisted forming is now 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. Peen forming is used to produce curvatures on thin sheet metals by *shot peening* (see Section 34.2) one surface of the sheet. As a result, the surface of the sheet is subjected to compressive stresses, which tend to expand the surface layer. Because the material below the peened surface remains rigid, the surface expansion causes the sheet to develop a curvature. The process also induces compressive surface residual stresses, which improve the fatigue strength of the sheet.

Peening is done with cast-iron or steel shot 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 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 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 Forming. This process involves the application of laser beams as a heat source in specific regions of the sheet metal. 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, for example, can be bent permanently without using dies. In **laser-assisted forming,**

the laser acts as a localized heat source, thus reducing the strength of the sheet metal at specific locations, improving formability, and increasing process flexibility. Applications include straightening, bending, embossing, and forming of complex tubular or flat components.

Microforming. This is a more recent development and describes a family of processes that are used to produce very small metallic parts and components. Examples of *minitaturized products* include a wristwatch with an integrated digital camera and one gigabyte of a computer storage component. Typical components made by microforming include small shafts for micromotors, springs, screws, and a variety of cold-headed, extruded, bent, embossed, coined, punched, or deep-drawn parts. Dimensions are typically in the submillimeter range, and weights are on the order of milligrams.

Electrohydraulic Forming. Also called *underwater spark* or *electric-discharge forming*, the source of energy in this forming process is a spark between electrodes that are connected with a short, thin wire. The rapid discharge of the energy from a capacitor bank through the wire generates a shock wave, similar to those created by explosives. The energy levels are lower than those in explosive forming, being typically a few kJ. The pressure developed in the water medium is sufficiently high to form the part. Electrohydraulic forming is a batch process and can be used in making various small parts.

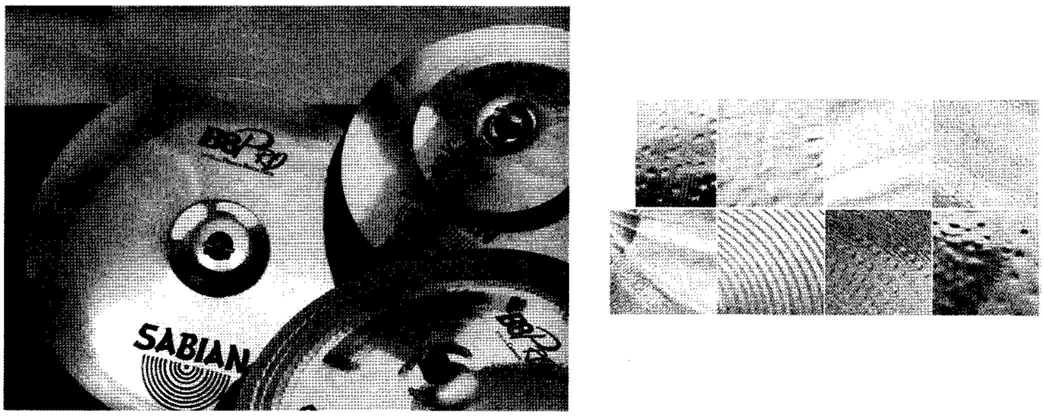
Gas Mixtures. As an energy source in this process, a gas mixture in a closed container is ignited. The pressure generated is sufficiently high to form sheet-metal parts. Although not often used in practice, the principle of this process is similar to that used for the generation of pressure in an internal combustion engine.

Liquefied Gases. Liquefied gases (such as liquid nitrogen) also have been used to develop pressures sufficiently high to form sheet metals. When allowed to reach room temperature in a closed container, liquefied nitrogen becomes gaseous and expands, developing the necessary pressure. Although not used in practice, the process is capable of forming relatively shallow parts.

CASE STUDY 16.3 Cymbal Manufacture

Cymbals (Fig. 16.51a) 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. 16.51b) is available to achieve the desired performance.

Cymbals are produced from metals—such as B20 bronze (80% Cu–20% Sn with a trace of silver), B8 bronze (92% Cu–8% Sn), nickel–silver alloy, and brass—by various methods of processing. The manufacturing sequence for producing B20 bronze cymbals is shown in Fig. 16.52. The metal is first cast into mushroom-shaped ingots and then cooled by ambient temperature. It then is 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, impart preferred grain orientation, and develop an even, round shape. The as-rolled blanks are then reheated and stretch formed (pressed) into the cup or bell shape that determines the cymbal's overtones. The cymbals then are center drilled or punched to create 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



(a) (b)

FIGURE 16.51 (a) A selection of common cymbals. (b) Detailed view of different surface textures and finishes of cymbals. *Source*: Courtesy of W. Blanchard, Sabian Ltd.

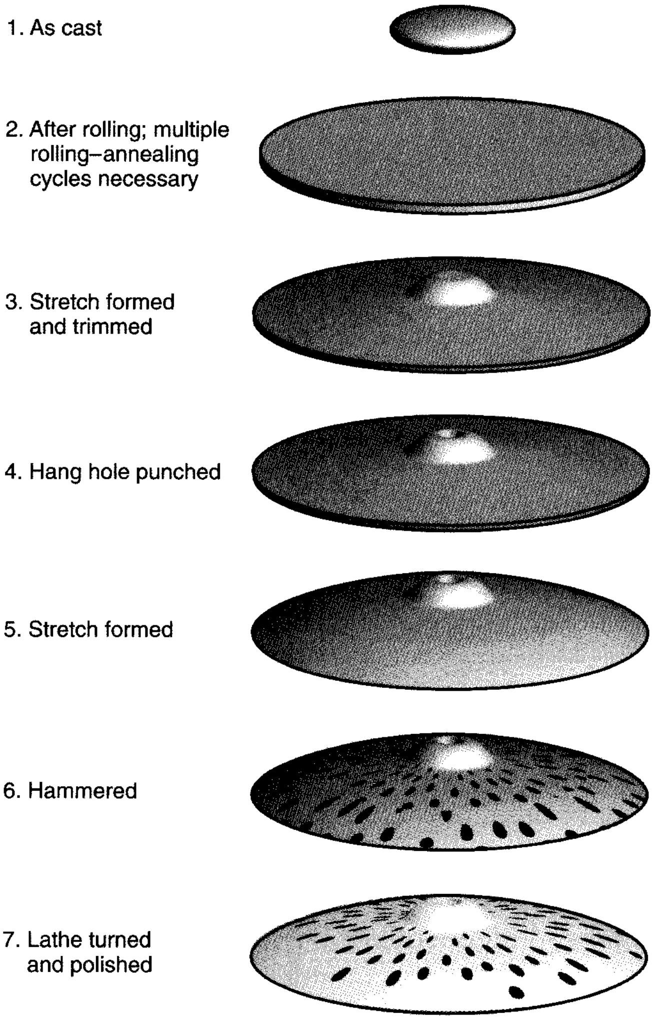
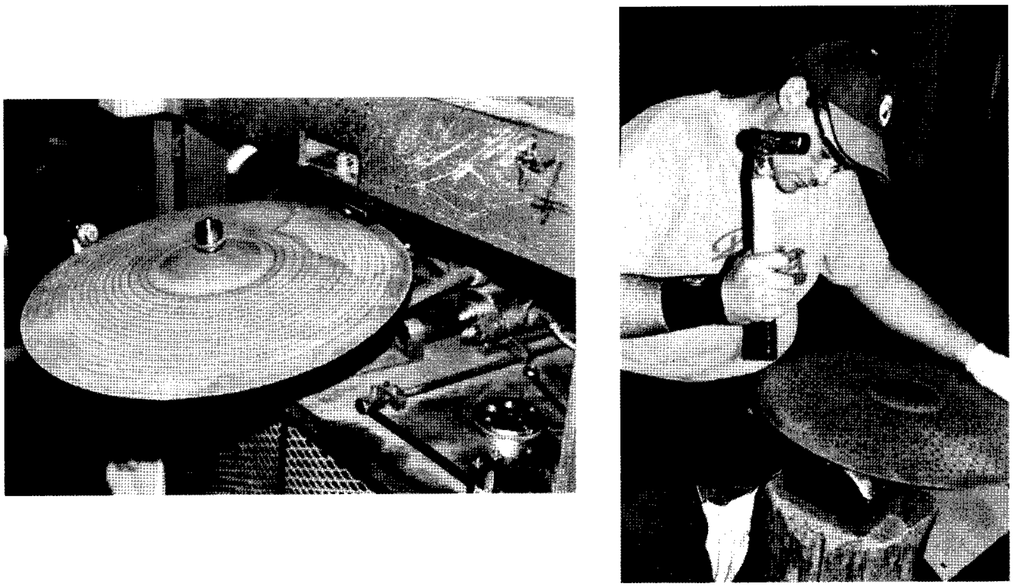


FIGURE 16.52 Manufacturing sequence for the production of cymbals. *Source*: Courtesy of W. Blanchard, Sabian Ltd.

(*continued*)



(a) (b)

FIGURE 16.53 Hammering of cymbals. (a) Automated hammering on a peening machine; (b) hand hammering of cymbals. *Source*: Courtesy of W. Blanchard, Sabian Ltd.

"Turkish dish" form that controls the cymbal's pitch.

Automatic peen-forming is done on machinery without templates, since the cymbals have already been pressed into shape, but the 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 in automatic peen-forming machines (Fig. 16.53a) or by hand (Fig. 16.53b). Hand hammering involves placing the bronze blank on a steel anvil, where the cymbals then are struck manually by hand hammers.

A number of finishing operations are performed on the cymbals. These can involve merely the cleaning and printing of identifying information, as some musicians prefer the natural surface appearance and sound of formed, hot-rolled bronze. More commonly, the cymbals are lathe turned (without any machining fluid) in order to remove the oxide surface and reduce the thickness of the cymbal to create the desired weight and sound. As a result of this process, the surface finish becomes lustrous and, in some cases, 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 the cymbal, and it is also an aesthetic feature that is appreciated by musicians. Various surface finishes associated with modern cymbals are shown in Fig. 16.51b.

Source: Courtesy of W. Blanchard, Sabian Ltd.

16.12 Manufacturing of Metal Honeycomb Structures

A *honeycomb structure* consists basically of a core of honeycomb or other corrugated shapes bonded to two thin outer skins (Fig. 16.54). 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.

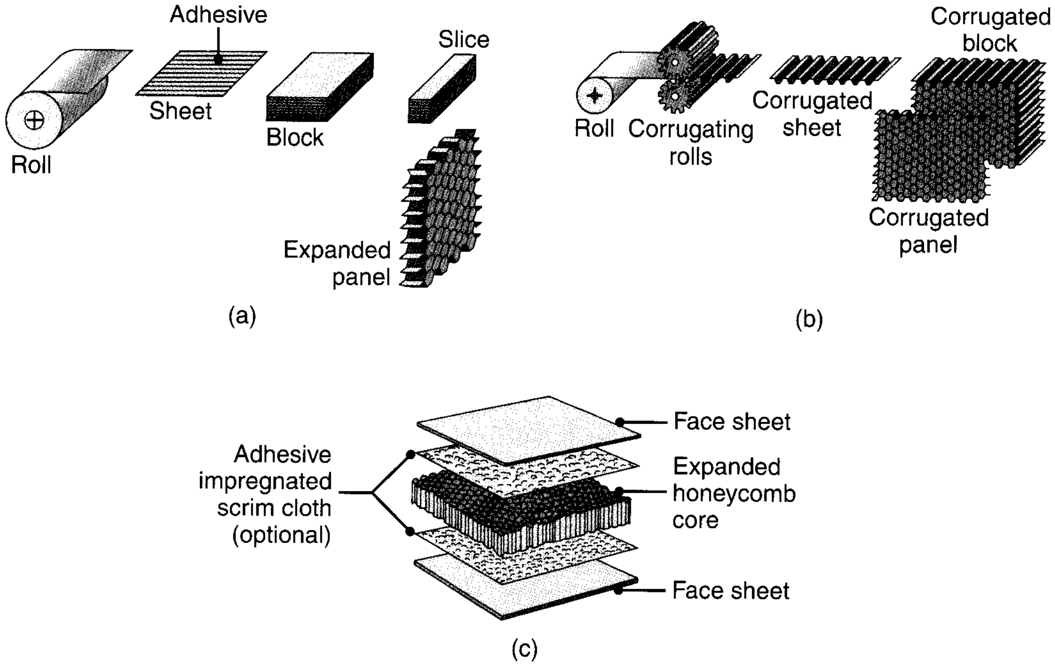


FIGURE 16.54 Methods of manufacturing honeycomb structures: (a) expansion process; (b) corrugation process; (c) assembling a honeycomb structure into a laminate.

Because of their light weight and high resistance to bending forces, metal honeycomb structures are used for aircraft and aerospace components, in buildings, and in transportation equipment. The chassis of the new Koenigsegg (Swedish) sports car, for example, is made partly of aluminum honeycomb with an integrated fuel tank.

Metal honeycomb manufacturing methods are described in this section because they involve operations that are best classified under sheet-metal-forming processes. Note, however, that honeycomb structures also may be made of nonmetallic materials, such as cardboard, polymers, and a variety of composite materials.

Honeycomb structures are made most commonly of 3000-series aluminum. However, they also are made of titanium, stainless steels, and nickel alloys for specialized applications and corrosion resistance. More recent developments include making honeycomb structures by using reinforced plastics, such as aramid-epoxy.

There are two basic methods of manufacturing honeycomb materials. In the **expansion process**, which is the more common method (Fig. 16.54a), sheets first are cut from a coil, and an *adhesive* (see 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 the adhesive joints. The block is finally cut into slices of the desired dimension and stretched to produce a honeycomb structure.

The **corrugation process** (Fig. 16.54b) is similar to the process used in making corrugated cardboard. The sheet metal first passes through a pair of specially designed rolls, becoming a corrugated sheet; it is then cut into desired lengths. Adhesive is applied to the node lines, the corrugated sheets are assembled into a block, and the block is cured. Because the sheets are preformed already, no expansion process is involved. The honeycomb material is finally made into a sandwich structure (Fig. 16.54c) by using face sheets that are joined by adhesives (or *brazed;* see Section 32.2) to the top and bottom surfaces.

16.13 Design Considerations in Sheet-metal Forming

As with most other processes described throughout this book, certain design guidelines and practices have evolved with time. Careful design using the best established design practices, computational tools, and manufacturing techniques is the best approach to achieving high-quality designs and realizing cost savings. The following guidelines apply to sheet-metal-forming operations, with the most significant design issues identified.

Blank Design. Material scrap is the primary concern in blanking operations. (See also Table 40.6.) Poorly designed parts will not *nest* properly, and there can be considerable scrap between successive blanking operations (Fig. 16.55). Some restrictions on blank shapes are made by the design application, but whenever possible, blanks should be designed to reduce scrap to a minimum.

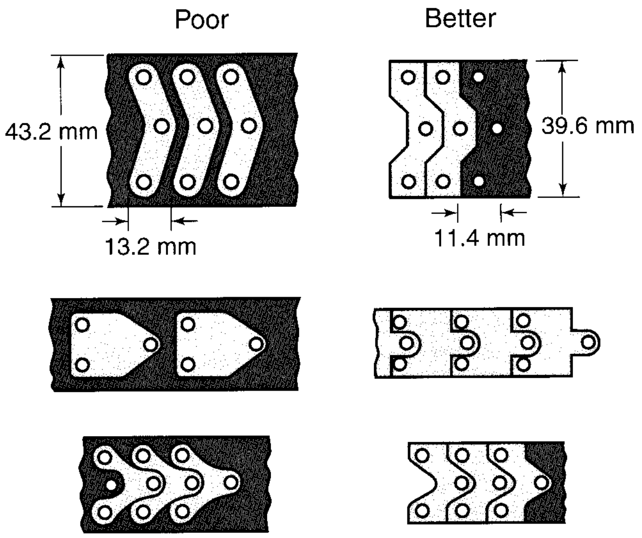


FIGURE 16.55 Efficient nesting of parts for optimum material utilization in blanking. *Source*: Courtesy of Society of Manufacturing Engineers.

Bending. In bending operations, the main concerns are material fracture, wrinkling, and the inability to form the bend. As shown in Fig. 16.56, a sheet-metal part with a flange that is to be bent will force the flange to undergo compression, which can cause buckling (see also *flanging*, Section 16.6). This problem can be controlled with a relief notch cut to limit the stresses from bending, or else a design modification as shown in the figure can be made to eliminate the problem. Right-angle bends have similar difficulties, and relief notches also can be used to avoid tearing (Fig. 16.57).

Because the bend radius is a highly stressed area, all stress concentrations should be removed from the bend-radius location. An example is parts with 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. 16.58a). Similarly, in bending flanges, tabs and notches should be avoided, since these stress concentrations will greatly reduce formability. When tabs are necessary, large radii should be used to reduce stress concentration (Fig. 16.58b).

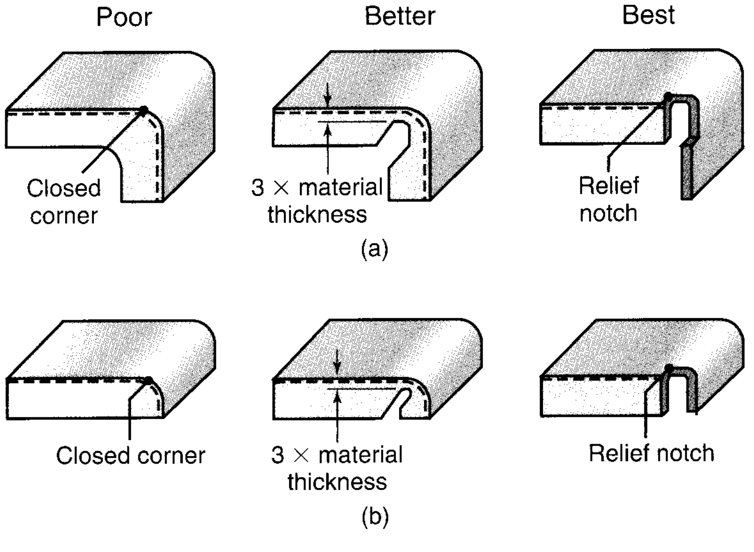


FIGURE 16.56 Control of tearing and buckling of a flange in a right-angle bend. *Source*: Courtesy of Society of Manufacturing Engineers.

When the bending and production of notches is to be used, it is important to orient the notches properly with respect to the grain direction. As shown in Fig. 16.17, bends ideally should be perpendicular to the rolling direction (or oblique if this is not possible) in order to avoid cracking. Bending sharp radii can be accomplished through scoring or embossing (Fig. 16.59), but it should be recognized that this can result in fracturing. Burrs should not be present in a bend allowance, since they are brittle

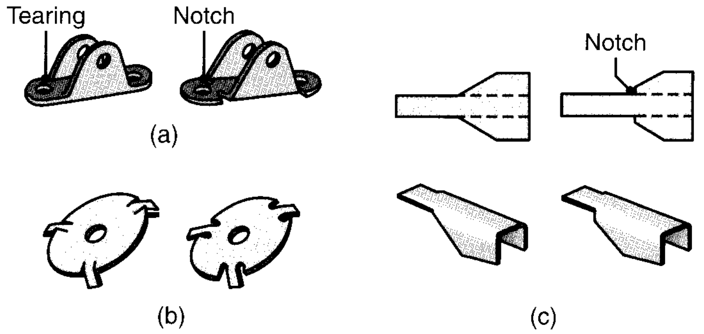


FIGURE 16.57 Application of notches to avoid tearing and wrinkling in right-angle bending operations. *Source*: Courtesy of Society of Manufacturing Engineers.

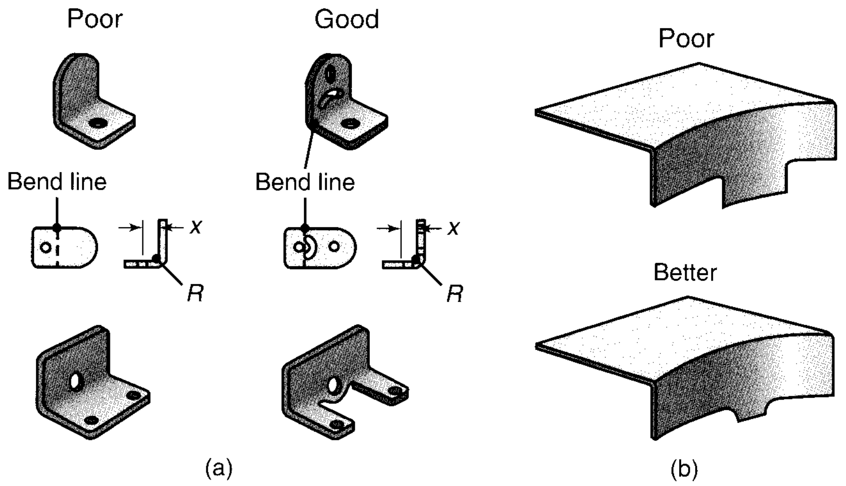


FIGURE 16.58 Stress concentrations near bends. (a) Use of a crescent or ear for a hole near a bend. (b) Reduction of severity of tab in flange. *Source*: Courtesy of Society of Manufacturing Engineers.

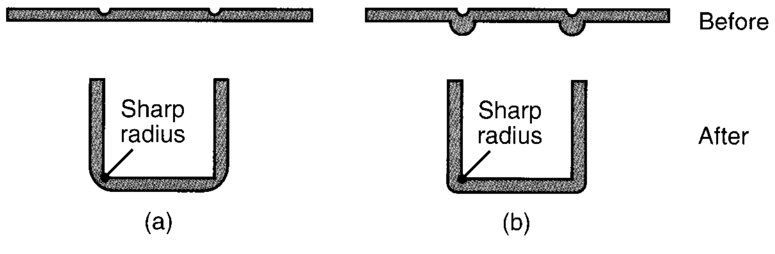


FIGURE 16.59 Application of (a) scoring or (b) embossing to obtain a sharp inner radius in bending. Unless properly designed, these features can lead to fracture. *Source*: Courtesy of Society of Manufacturing Engineers.

(because of strain hardening) and can fracture, leading to a stress concentration that propagates the crack into the rest of the sheet.

Roll Forming. In general, the process should 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 embossings can be located on the roll-formed shape.

Stamping and Progressive-die Operations. In progressive dies (see Section 16.2.3), the cost of the tooling and the number of stations are determined by the number and spacing of features on a part. Therefore, it is advantageous to hold the number of features to a minimum in order to minimize tooling costs. Closely spaced features may provide insufficient clearance for punches and may require two punches. Narrow cuts and protrusions also are problematic for forming with a single punch and die.

Deep Drawing. After a deep-drawing operation, a cup invariably will spring back towards its original shape. For this reason, designs that use a vertical wall in a deep-drawn cup may be difficult to form. Relief angles of at least 3° on each wall are easier to produce. Cups with sharp internal radii are difficult to produce, and deep cups often require one or more ironing operations.

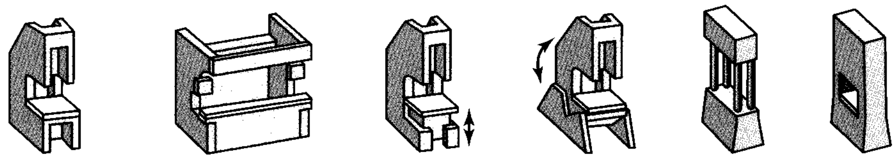
16.14 Equipment for Sheet-metal Forming

For most general pressworking operations, the basic equipment consists of mechanical, hydraulic, pneumatic, or pneumatic–hydraulic presses with a wide variety of designs, features, capacities, and computer controls. Typical designs for press frames are shown in Fig. 16.60 (see also Figs. 14.17 and 16.23f). The proper design, stiffness, and construction of such equipment is essential to the efficient operation of the system and to achieving a high production rate, good dimensional control, and high product quality.

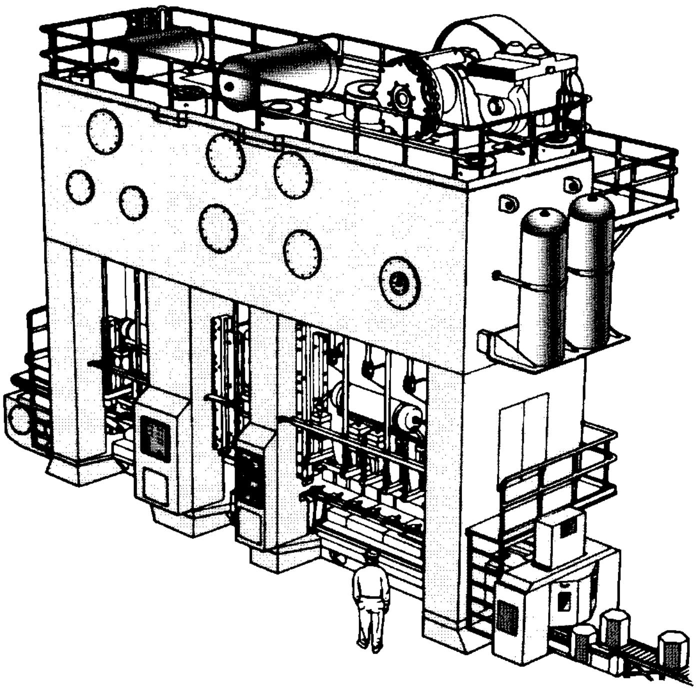
The traditional **C-frame** structure (Fig. 16.60a) has been used widely for ease of tool and workpiece accessibility, but it is not as stiff as the **box-type pillar** (Fig. 16.60e) or the **double-column frame** structure (Fig. 16.60f). Furthermore, accessibility has become less important due to advances 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, the size and shape of the dies, and the tooling required.
2. Size and shape of workpieces.
3. Length of stroke of the slide, the number of strokes per minute, the operating speed, and the *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 are typically used for reverse redrawing and for other complicated forming operations.
5. Maximum force required (press capacity and tonnage rating).
6. Type of mechanical, hydraulic, and computer controls.
7. Features for changing dies. Because the time required for changing dies in presses can be significant (as much as a few hours) and thus affect productivity, rapid die-changing systems have been developed. Following a system called *single-minute exchange of die* (SMED), die setups can be changed in less than 10 minutes by using computer-controlled hydraulic or pneumatic systems.
8. Safety features.



Basic C-frame design (a) Wide design (b) Adjustable bed (c) Open-back inclinable (d) Pillar (e) Double column (f)



(g)

FIGURE 16.60 (a) through (f) Schematic illustrations of types of press frames for sheet-forming operations. Each type has its own characteristics of stiffness, capacity, and accessibility. (g) A large stamping press. *Source*: (a) through (f) *Engineer's Handbook*, VEB Fachbuchverlag, 1965; (g) Verson Allsteel Company.

Because a press is a major capital investment, its present and future use for a broad variety of parts and applications should be investigated. Versatility and multiple use are important factors in press selection, particularly for product modifications and for making new products to respond to continually changing global markets.

16.15 Economics of Sheet-forming Operations

Sheet-metal forming involves economic considerations similar to those for the other processes that have been described. Sheet-forming operations are very versatile, and a number of different processes can be used to produce the same part. The costs involved (see also Chapter 40) depend on the particular operations (such as die and equipment costs and labor). For small and simple sheet-metal parts, die costs and

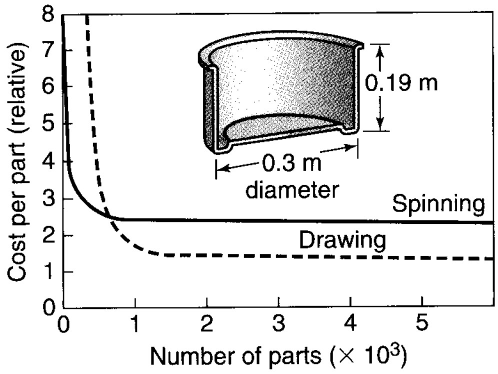


FIGURE 16.61 Cost comparison for manufacturing a round sheet-metal container either by conventional spinning or by deep drawing. Note that for small quantities, spinning is more economical.

lead times to make the dies are rather low. On the other hand, these costs for large-scale operations (such as stretch forming of aircraft panels and boat hulls) are very high. Furthermore, because the number of such parts needed is rather low, the cost per piece can be very high (see Fig. 14.18).

Similar considerations are involved in other sheet-forming operations. Deep drawing requires expensive dies and tooling, but a very high number of parts, such as containers, cans, and similar products, are produced with the same setup. These costs for other processes, such as punching, blanking, bending, and spinning, vary considerably, depending on part size and thickness.

Equipment costs vary widely and depend largely on the complexity of the forming operation, part loading and unloading features, part size and shape, and level of automation and computer control required. Automation, in turn, directly affects the amount of labor required and the skill level. The higher the extent of automation, the lower the skill level required. Furthermore, many sheet-metal parts generally require some finishing operations—one of the most common being deburring of the edges of the part, which generally is labor intensive, although some advances have been made in automated deburring (which itself requires computer-controlled equipment; hence, it can be costly).

As an example of the versatility of sheet-forming operations and the costs involved, note that a cup-shaped part can be formed by deep drawing, spinning, rubber forming, or explosive forming. Moreover, it also can be formed by impact extrusion, casting, or fabrication from different pieces. The part shown in Fig. 16.61 can be made either by deep drawing or by conventional spinning, but the die costs for the two processes are significantly different.

Deep-drawing dies have many components, and they cost much more than the relatively simple mandrels and tools employed in spinning. Consequently, the die cost per part in drawing will be high if only a few parts are needed. However, this part can be formed by deep drawing in a much shorter time than by spinning, even if the latter operation is automated and computer controlled. Furthermore, spinning generally requires more skilled labor. Considering these factors, the break-even point can be seen as around 700 parts, and deep drawing is more economical for quantities greater than that. Chapter 40 describes further details of the economics of manufacturing.

SUMMARY

- Sheet-metal-forming processes are among the most versatile of all 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 forming often prevents the thickness of the material from being reduced (to avoid necking and tearing).
- Important material parameters are the quality of the sheared edge of the sheet metal prior to forming, the capability of the sheet to stretch uniformly, the material's resistance to thinning, its normal and planar anisotropy, its grain size, and its yield-point elongation (for low-carbon steels).
- The forces and energy required in sheet-metal-forming processes are transmitted to the workpiece through solid tools and dies, by flexible rubber or polyurethane members, or by electrical, chemical, magnetic, and gaseous means.

- Because of the relatively thin materials used, springback, buckling, and wrinkling are significant problems in sheet forming. Springback is a function of the yield stress, the elastic modulus, and the ratio of bend radius to thickness. These problems can be reduced or eliminated by proper tool and die design, by minimizing the unsupported length of the sheet during processing, and by controlling the thickness of the incoming sheet and its mechanical properties.
- Among important developments is the superplastic forming of diffusion-bonded sheets. The process is capable of producing complex sheet-metal structures, particularly for aerospace applications (which require particularly high stiffness-to-weight ratios).
- Several test methods have been developed for predicting the formability of sheet metals. In bending operations, the tensile reduction of the area of the sheet gives an indication of its bendability (minimum bend radius); this also applies to, the spinnability parameter of metals (maximum reduction in thickness per pass).
- For general stamping operations, forming-limit diagrams are very useful, because they establish quantitative relationships among the major and minor principal strains that limit safe forming. For deep-drawing operations, the important parameter is the normal or plastic anisotropy of the sheet (the ratio of width strain to thickness strain in tensile testing).

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
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
Wrinkling

BIBLIOGRAPHY

ASM Handbook, Vol. 14B: **Metalworking: Sheet Forming**, ASM International, 2006.

Benson, S., **Press Brake Technology**, Society of Manufacturing Engineers, 1997.

Boljanovic, V., **Sheet Metal Forming Process and Die Design**, Industrial Press, 2004.

Davies, G., **Materials for Automobile Bodies**, Butterworth-Heinemann, 2003.

Fundamentals of Tool Design, 5th ed., Society of Manufacturing Engineers, 2005.

Gillanders, J., **Pipe and Tube Bending Manual**, FMA International, 1994.

Hosford, W.F., and Caddell, R.M., **Metal Forming: Mechanics and Metallurgy**, 3rd ed., Cambridge, 2007.

Hu, J., Marciniak, Z., and Duncan, J., **Mechanics of Sheet Metal Forming**. Butterworth-Heinemann, 2002.

Pearce, R., **Sheet Metal Forming**, Springer, 2006.

Progressive Dies, Society of Manufacturing Engineers, 1994.

Rapien, B.L., **Fundamentals of Press Brake Tooling**, Hanser Gardner, 2005.

Smith, D., **Fundamentals of Pressworking,** Society of Manufacturing Engineers, 1994.

Suchy, I., **Handbook of Die Design,** 2nd ed., McGraw-Hill, 2005.

Szumera, J., **Metal Stamping Process,** Industrial Press, 2003.

Tool and Manufacturing Engineers Handbook, 4th ed., Vol. 2: **Forming.** Society of Manufacturing Engineers, 1984.

Tschaetch, H., **Metal Forming Practice: Processes, Machines, Tools,** Springer, 2007.

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. What is microforming?

16.17. Explain the advantages of superplastic forming.

QUALITATIVE PROBLEMS

16.18. Explain the differences that you have observed between products made of sheet metals and those made by casting and forging.

16.19. Identify the material and process variables that influence the punch force in shearing, and explain how each of them affects this force.

16.20. Explain why springback in bending depends on yield stress, elastic modulus, sheet thickness, and bend radius.

16.21. Explain why cupping tests may not predict well the formability of sheet metals in actual forming processes.

16.22. Identify the factors that influence the deep-drawing force, F, in Fig. 16.31b, and explain why they do so.

16.23. Why are the beads in Fig. 16.35b placed in those particular locations?

16.24. 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.

16.25. 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.

16.26. Is it possible to have ironing take place in an ordinary deep-drawing operation? What is the most important factor?

16.27. 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.

16.28. What recommendations would you make in order to eliminate the cracking of the bent piece shown in Fig. 16.17c? Explain your reasons.

16.29. Give several specific examples from this chapter in which friction is desirable and several in which it is not desirable.

16.30. 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.

16.31. 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.

16.32. How would you produce the part shown in Fig. 16.40b other than by tube hydroforming?

QUANTITATIVE PROBLEMS

16.33. Calculate R_{avg} for a metal where the R values for the 0°, 45°, and 90° directions are 0.8, 1.7, and 1.8, respectively. What is the limiting drawing ratio (LDR) for this material?

16.34. Calculate the value of ΔR in Problem 16.33. Will any ears form when this material is deep drawn? Explain.

16.35. Estimate the limiting drawing ratio for the materials listed in Table 16.4.

16.36. 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.

16.37. 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.

16.38. For explosive forming, calculate the peak pressure in water for 1.2 N of TNT at a standoff distance of 1200 mm. Comment on whether or not the magnitude of this pressure is sufficiently high to form sheet metals.

16.39. 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.

16.40. Plot Eq. (16.6) in terms of the elastic modulus, E, and the yield stress, Y, of the material, and describe your observations.

16.41. What is the minimum bend radius for a 1.5-mm-thick sheet metal with a tensile reduction of area of 30%? Does the bend angle affect your answer? Explain.

16.42. 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.

16.43. 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.

16.44. What is the force required to punch a square hole 60 mm on each side in a 1-mm-thick 5052-O aluminum sheet by using flat dies? What would be your answer if beveled dies are used?

16.45. In Example 16.1, 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 15%?

16.46. 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

16.47. 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.

16.48. Consider several shapes to be blanked from a large sheet (such as oval, triangular, L-shaped, and so forth) by laser-beam cutting, and sketch a nesting layout to minimize scrap generation.

16.49. Give several specific product applications for (a) hemming and (b) seaming.

16.50. Many axisymmetric missile bodies are made by spinning. What other methods could you use if spinning processes were not available?

16.51. 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.

16.52. 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.

16.53. The design shown in Fig. P16.53 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.

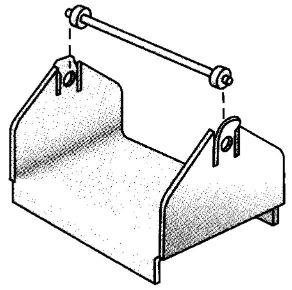


FIGURE P16.53

16.54. 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.

16.55. 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.

16.56. 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.

16.57. 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.

Powder-Metal Processing and Equipment

CHAPTER 17

- This chapter describes the powder metallurgy process for producing net-shape parts from metal powders.
- The chapter begins by examining methods of producing and blending of metal powders and investigates the shapes that powders will develop based on the process employed to make them.
- A number of secondary operations are then presented, such as compaction to consolidate the powder into a desired shape, and sintering to fuse the particles to achieve the required strength.
- Additional processes particular to powder metallurgy are then discussed, and design rules are presented.
- The chapter ends with a discussion of process capabilities and economics 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, and machining.

17.1 Introduction

In the manufacturing processes described thus far, the raw materials used have been metals and alloys either in a molten state (casting) or in solid form (metalworking). This chapter describes the **powder metallurgy** (PM) process, in which metal powders are compacted into desired and often complex shapes and sintered (heated without melting) to form a solid piece. This process first was used by the Egyptians 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 overall economics of the operation give this unique process its numerous attractive and expanding applications.

A wide range of parts and components are made by powder-metallurgy techniques (Fig. 17.1): (a) balls for ballpoint pens; (b) automotive components (which

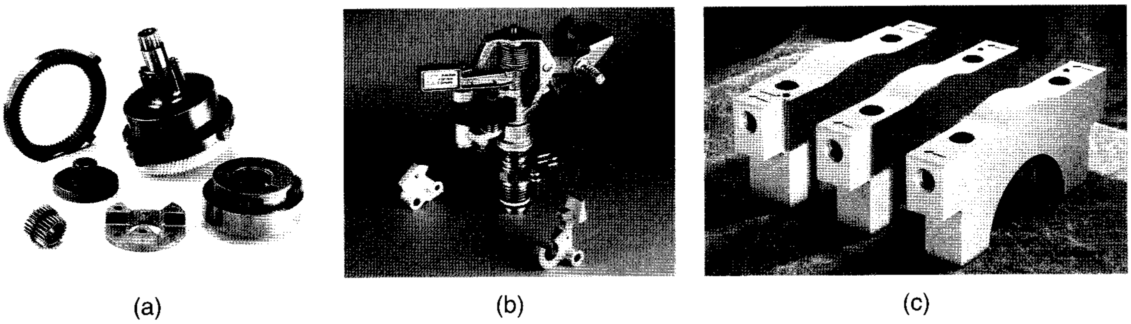

(a) (b) (c)

FIGURE 17.1 (a) Examples of typical parts made by powder-metallurgy 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 metal-powder 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*, 1998. Metal Powder Industries Federation, Princeton, New Jersey, 1998. (c) Courtesy of Zenith Sintered Products, Inc., Milwaukee, Wisconsin.

now constitute about 70% of the PM market) such as piston rings, connecting rods, brake pads, gears, cams, and bushings; (c) tool steels, tungsten carbides, and cermets as tool and die materials; (d) graphite brushes impregnated with copper for electric motors; (e) magnetic materials; (f) metal filters and oil-impregnated bearings with controlled porosity; (g) metal foams; (h) surgical implants, and (i) several others for aerospace, nuclear, and industrial applications. Advances in this technology now permit *structural* parts of aircraft, such as landing gear components, engine-mount supports, engine disks, impellers, and engine nacelle frames, to be made by PM.

Powder metallurgy has become competitive with processes such as casting, forging, and machining, particularly for relatively complex parts made of high-strength and hard alloys. Although most parts weigh less than 2.5 kg, they can weigh as much as 50 kg. It has been shown that PM parts can be mass-produced economically in quantities as small as 5000 per year and as much as 100 million per year for vibrator weights for cell phones.

The most commonly used metals in PM are iron, copper, aluminum, tin, nickel, titanium, and the refractory metals. For parts made of brass, bronze, steels, and stainless steels, *prealloyed powders* are used, where each powder particle itself is an alloy. Metal sources are generally bulk metals and alloys, ores, salts, and other compounds.

17.2 Production of Metal Powders

The powder-metallurgy process typically consists of the following operations, in sequence (Fig. 17.2):

1. Powder production;
2. Blending;
3. Compaction;
4. Sintering;
5. Finishing operations.

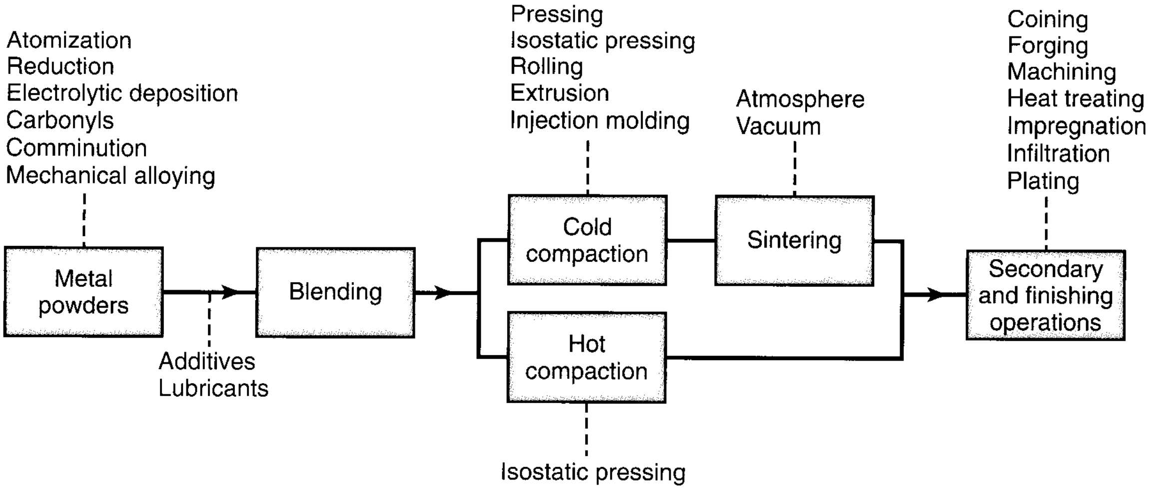


FIGURE 17.2 Outline of processes and operations involved in producing powder-metallurgy parts.

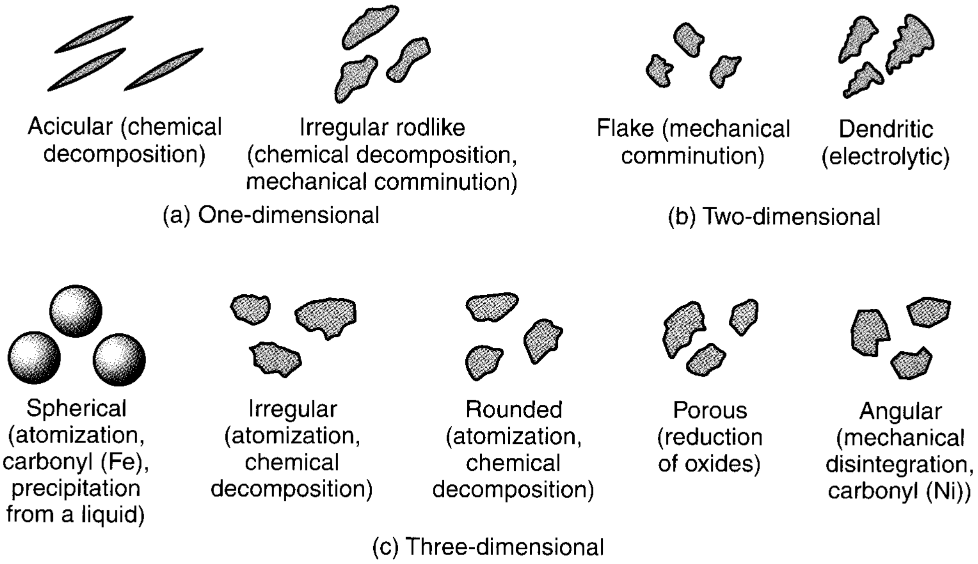


FIGURE 17.3 Particle shapes in metal powders, and the processes by which they are produced. Iron powders are produced by many of these processes (see also Fig. 17.4).

17.2.1 Methods of Powder Production

There are several methods of producing metal powders, and most of them can be produced by more than one method. The choice depends 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 the flow and permeability during compaction and in subsequent sintering operations. Particle sizes produced range from 0.1 to 1000 μm.

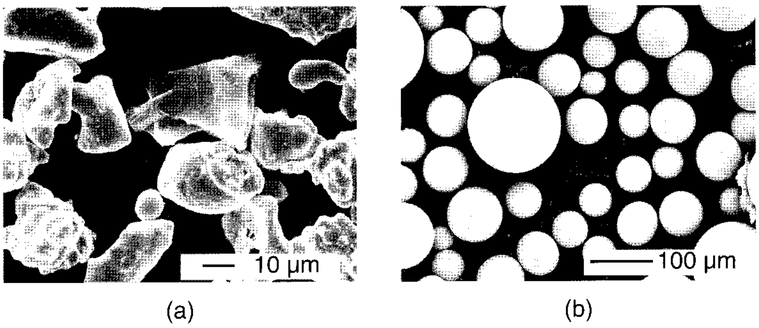


FIGURE 17.4 (a) Scanning-electron microscope image of iron-powder particles made by atomization. (b) Nickel-based superalloy (Udimet 700) powder particles made by the rotating electrode process; see Fig. 17.5d. *Source*: Courtesy of P.G. Nash, Illinois Institute of Technology, Chicago.

Atomization. Atomization involves a liquid-metal stream produced by injecting molten metal through a small orifice. 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 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 and higher production rates. Gas atomization usually results in more spherical particles (see Fig. 17.3c).

In *centrifugal atomization*, the molten-metal stream drops onto a rapidly rotating disk or cup, so that centrifugal forces break up the stream and generate particles (Fig. 17.5c). In another variation of this method, a consumable electrode is

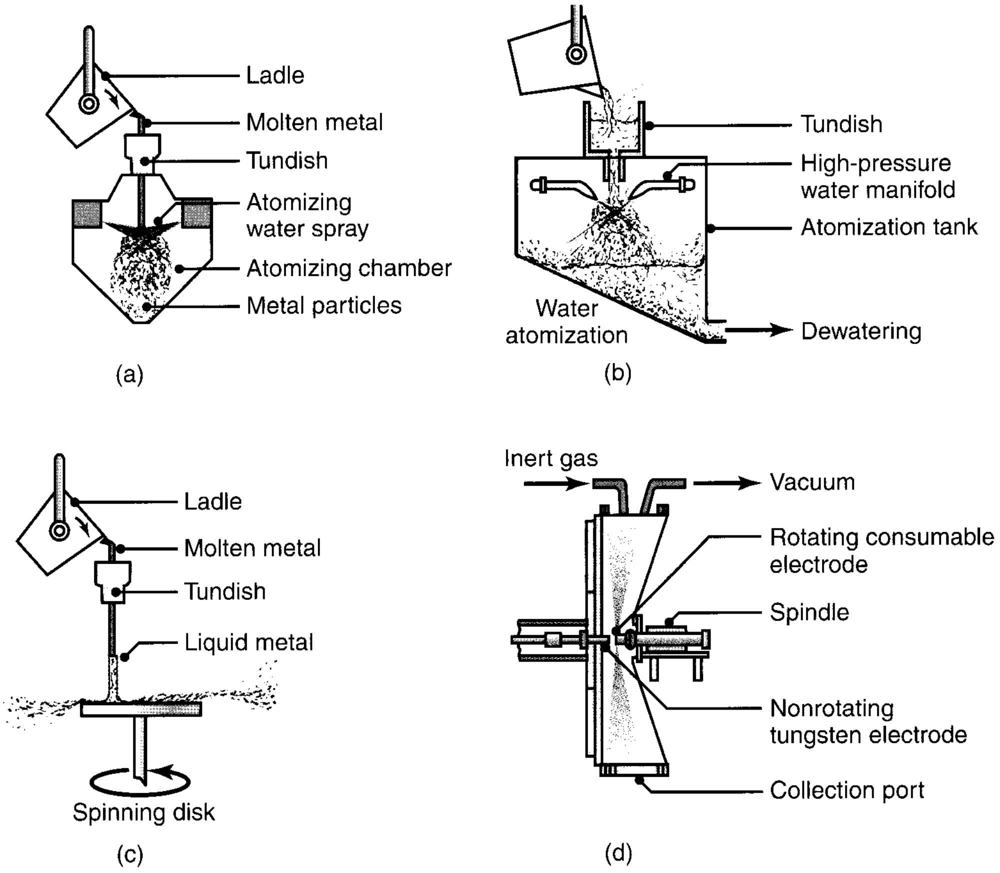


FIGURE 17.5 Methods of metal-powder production by atomization: (a) gas atomization; (b) water atomization; (c) centrifugal atomization with a spinning disk or cup; and (d) atomization with a rotating consumable electrode.

rotated rapidly (about 15,000 rev/min) in a helium-filled chamber (Fig. 17.5d). The centrifugal force breaks up the molten tip of the electrode into metal particles.

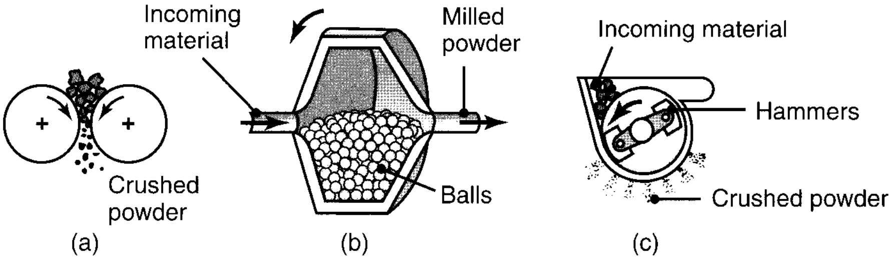


FIGURE 17.6 Methods of mechanical comminution to obtain fine particles: (a) roll crushing, (b) ball mill, and (c) hammer milling.

Reduction. The *reduction* of metal oxides (i.e., removal of 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 and have uniformly sized spherical or angular shapes.

Electrolytic Deposition. *Electrolytic deposition* utilizes either aqueous solutions or fused salts. The powders produced are among the purest available.

Carbonyls. *Metal carbonyls*, such as iron carbonyl [$Fe(CO)_5$] and nickel carbonyl [$Ni(CO)_4$] are formed by letting iron or nickel react with carbon monoxide. The reaction products are then decomposed to iron and nickel, and they turn into small, dense, uniformly spherical particles of high purity.

Comminution. *Mechanical comminution* (*pulverization*) involves crushing (Fig. 17.6), milling in a ball mill, or grinding of 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 into a ball mill are impacted by the balls as the cylinder is rotated or its contents are agitated. This action has two effects: (a) the particles periodically fracture, resulting in smaller particles, and (b) the morphology of the particles is affected. With brittle materials, the powder particles produced have angular shapes; with ductile metals, they are flaky and are not particularly suitable for powder-metallurgy applications.

Mechanical Alloying. In *mechanical alloying*, powders of two or more pure metals 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. The dispersed phase can result in

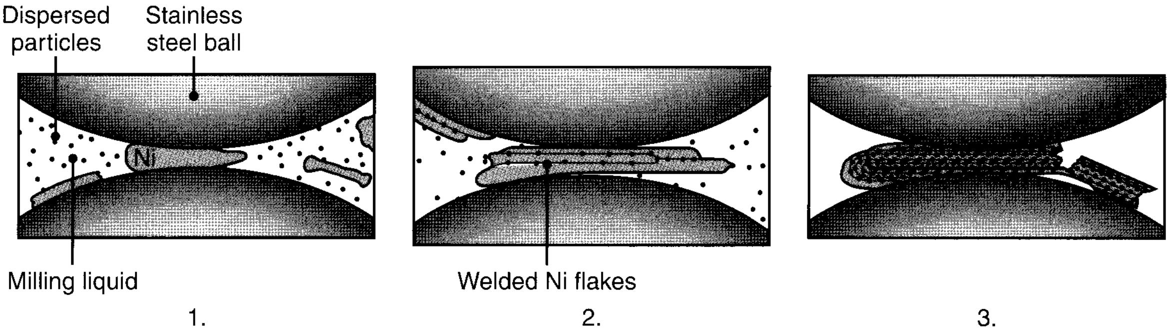


FIGURE 17.7 Mechanical alloying of nickel particles with dispersed smaller particles. As nickel particles are flattened between two balls, the second, smaller phase is impressed into the nickel surface and eventually is dispersed throughout the particle due to successive flattening, fracture, and welding events.

strengthening of the particles or can impart special electrical or magnetic properties to the powder.

Miscellaneous Methods. Other, less commonly used, methods for making powders are as follows:

- **Precipitation** from a chemical solution
- Production of fine metal chips by **machining**
- **Vapor condensation.**

More recent developments include techniques based on *high-temperature extractive metallurgical processess*—in turn based on the reaction of volatile halides (a compound of halogen and an electropositive element) with liquid metals and the controlled reduction and reduction/carburization of solid oxides.

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). Because these powders are pyrophoric (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). When the material is subjected to large plastic deformation by compression and shear at stress levels of 5500 MPa during processing of the powders, the particle size is reduced and the material becomes pore free and possesses enhanced properties.

Microencapsulated Powders. These metal powders are coated completely with a binder. For electrical applications (such as magnetic components of ignition coils and other pulsed AC and DC applications), the binder acts like an insulator, preventing electricity from flowing between particles and thus reducing eddy-current losses. The powders are compacted by warm pressing, and they 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 usually is controlled by *screening*—that is, by passing the metal powder through screens (*sieves*) of various mesh sizes. Screen analysis is achieved by using a vertical stack of screens, 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. For example, a mesh size of 30 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.)

In addition to screening, several other methods 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, and a detector then digitizes the signals and computes the particle-size distribution.

4. **Optical methods** (such as particles blocking a beam of light), in which the particle is sensed by a photocell.
5. **Suspending particles** in a liquid and then detecting particle size and distribution by electrical sensors.

Particle Shape. A major influence on processing characteristics, particle shape usually 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. This ratio ranges from unity for a spherical particle to about 10 for flake-like or needle-like particles.

Shape Factor. Also called the *shape index*, shape factor (SF) is a measure of the ratio of the surface area of the particle to its volume—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 for details). The maximum is called the *mode size*.

Other properites of metal powders that have an effect on their behavior in processing are (a) *flow properties* when the powders are filled into dies, (b) *compressibility* when they are being compacted, (c) *density*, as defined in various terms such as theoretical density, apparent density, and the density when the powder is shaken or tapped in the die cavity.

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 and other materials can be mixed in order to impart special physical and mechanical properties and characteristics to the PM product. Note that mixtures of metals can be produced by alloying the metal before producing a powder, or else blends can be produced. Proper mixing is essential to ensure the uniformity of mechanical properties throughout the part.
- Even when a single metal is used, the powders may vary significantly in size and shape; hence, they must be blended to obtain uniformity from part to part. The ideal mix is one in which all of the particles of each material (and of each size and morphology) are distributed uniformly.
- *Lubricants* can be mixed with the powders to improve their flow characteristics. They reduce friction between the metal particles, improve flow of the powder metals into the dies, and improve die life. Lubricants typically 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), are used to develop sufficient *green strength* (see Section 17.3), and additives also can be used to facilitate sintering.

Powder mixing must be carried out under controlled conditions in order to avoid contamination or deterioration. Deterioration is caused by excessive mixing, which may alter the shape of the particles and cause work hardening, making subsequent compaction more difficult. Powders can be mixed in air, in inert atmospheres (to avoid oxidation), or in liquids (which act as lubricants and

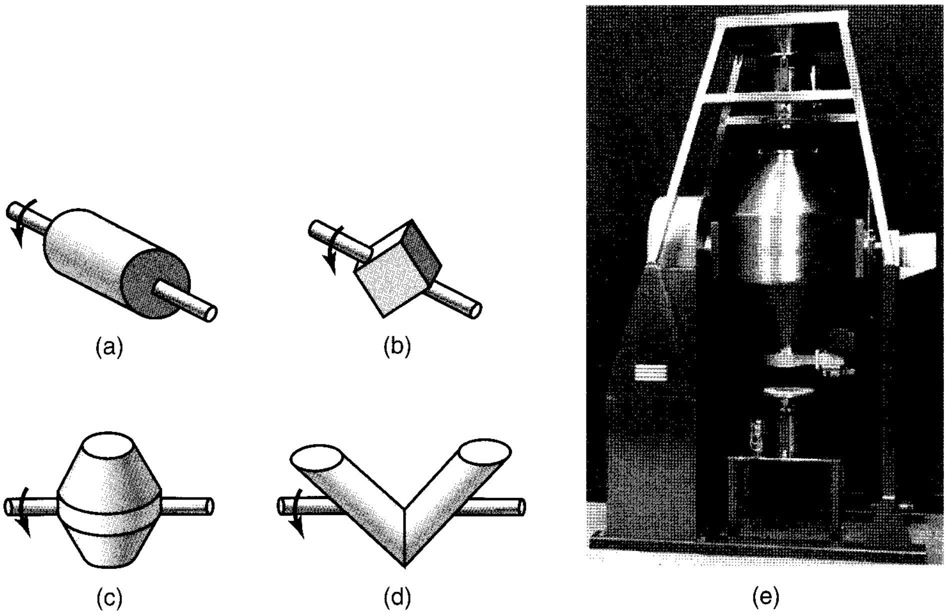


FIGURE 17.8 (a) through (d) Some common bowl geometries for mixing or blending powders. (e) A mixer suitable for blending metal powders. Since metal powders are abrasive, mixers rely on the rotation or tumbling of enclosed geometries, as opposed to using aggressive agitators. *Source*: Courtesy of Kemutec Group, Inc.

make the mix more uniform). Several types of blending equipment are available (Fig. 17.8).

Hazards. Because of their high surface area–to-volume ratio, metal powders can be explosive, particularly aluminum, magnesium, titanium, zirconium, and thorium. Great care must be exercised both during blending and in storage and handling. Precautions include (a) grounding equipment, (b) preventing sparks (by using non-sparking tools) and avoiding friction as a source of heat, and (c) avoiding dust clouds, open flames, and chemical reactions.

17.3 Compaction of Metal Powders

Compaction is the step in which the blended powders are pressed into various shapes in dies, as shown in sequence in Fig. 17.9. The purposes of compaction are to obtain the required shape, density, and particle-to-particle contact and to make the part sufficiently strong for further processing. The powder (*feedstock*) is fed into the die by a feed shoe, and the upper punch descends into the die. The presses used are actuated either hydraulically or mechanically, and the process generally is carried out at room temperature, although it can be done at elevated temperatures.

The pressed powder is known as **green compact**, since it has a low strength, just as is seen in green parts in slip casting (Section 18.2.1). The green parts are very fragile (similar to chalk) and can crumble or become damaged very easily; this situation is exacerbated by poor pressing practices. To obtain higher green strengths, the powder must be fed properly into the die cavity and proper pressures must be developed throughout the part.

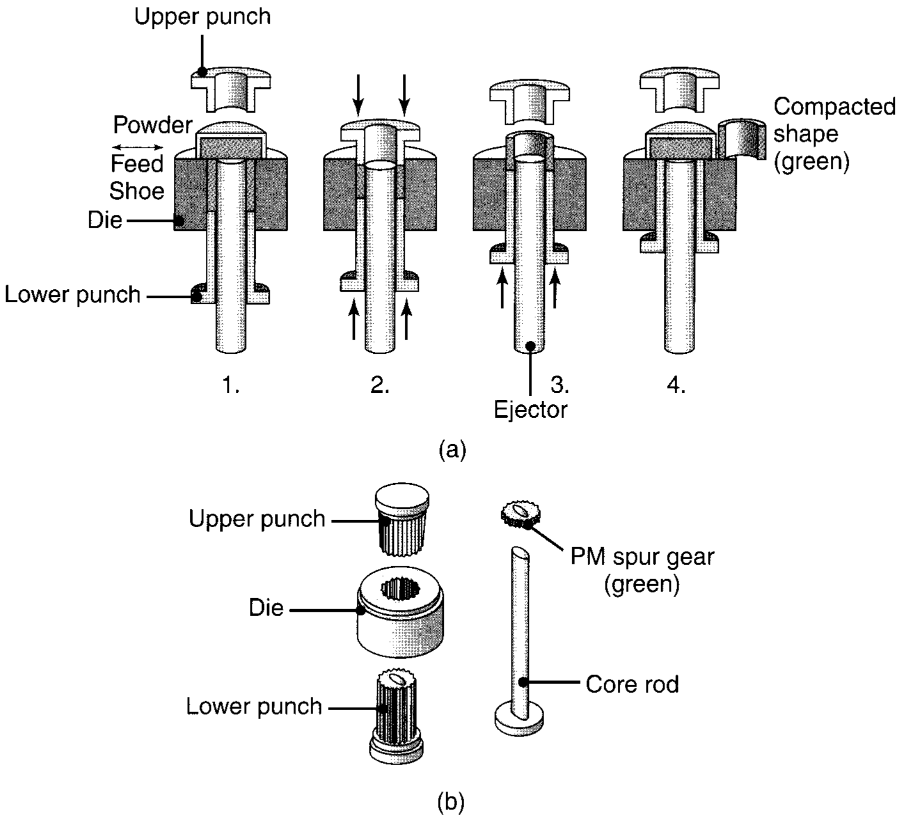


FIGURE 17.9 (a) Compaction of metal powder to form a bushing. The pressed-powder part is called green compact. (b) Typical tool and die set for compacting a spur gear. *Source*: Courtesy of the Metal Powder Industries Federation.

The *density* of the green compact depends on the pressure applied (Fig. 17.10a). As the compacting pressure is increased, the compact density approaches that of the metal in its bulk form. An important factor in density is the size distribution of the particles. If all of the particles are of the same size, there always will be some porosity when they are packed together, theoretically a porosity of at least 24% by volume. (Observe, for example, a box filled with rice, and note that there are always open spaces between the grains.) Introducing smaller particles into the powder mix will fill the spaces between the larger powder particles and thus result in a higher density of the compact (see also *porous aluminum*, Section 6.2).

The higher the density of the compacted part, the higher are its strength and elastic modulus (Fig. 17.10b). The reason is that the higher the density, the higher the amount of solid metal in the same volume, and hence the greater its strength (resistance to external forces). 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 with separate movements in order to ensure that the density is more uniform throughout the part (Fig. 17.11). Recall a similar discussion regarding the compaction of sand in mold making (see Fig. 11.07). However, density variation in components such as gears, cams, bushings, and structural parts may be desirable. For example, densities can be increased in critical locations where high strength and wear resistance are important and reduced where they are not.

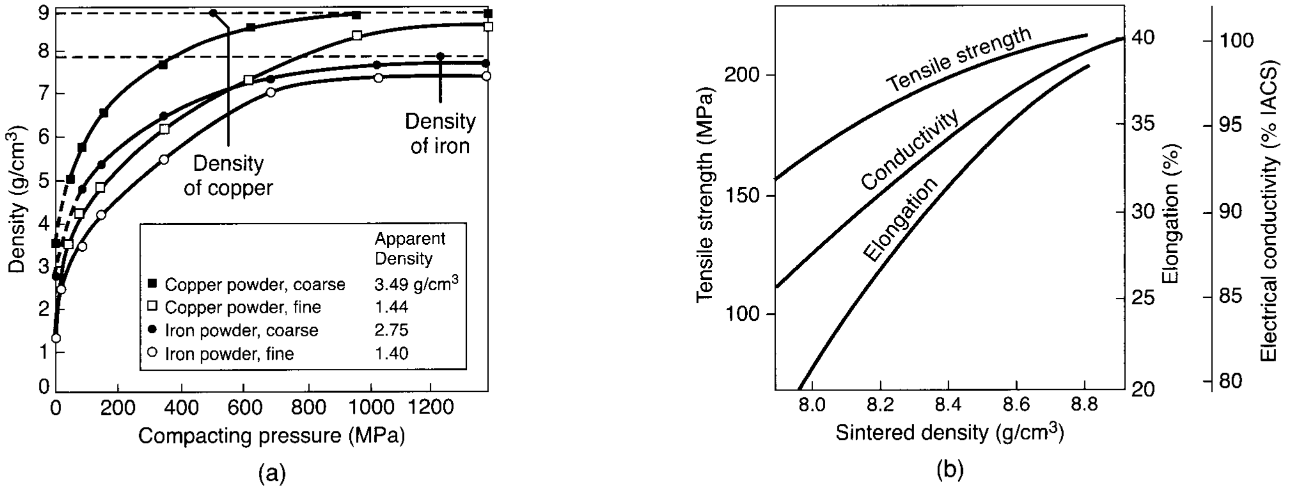


FIGURE 17.10 (a) Density of copper- and iron-powder compacts as a function of compacting pressure. Density greatly influences the mechanical and physical properties of PM parts. (b) Effect of density on tensile strength, elongation, and electrical conductivity of copper powder. *Source*: (a) After F.V. Lenel, (b) After the International Annealed Copper Standard (IACS) for electrical conductivity.

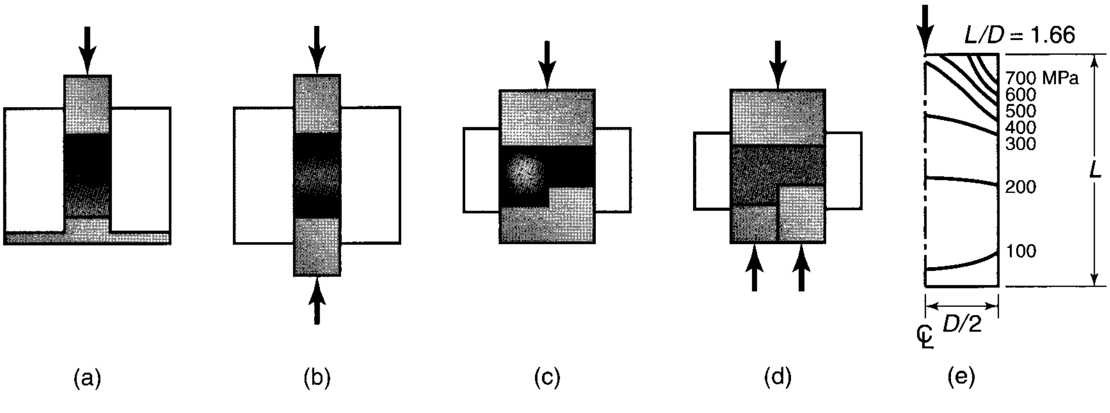


FIGURE 17.11 Density variation in compacting metal powders in various dies: (a) and (c) single-action press; (b) and (d) double-action press. Note in (d) the greater uniformity of density from pressing with two punches with separate movements compared with (c). (e) Pressure contours in compacted copper powder in a single-action press. *Source*: After P. Duwez and L. Zwell.

17.3.1 Equipment

The pressure required for pressing metal powders ranges from 70 MPa for aluminum to 800 MPa for high-density iron parts (see Table 17.1). The compacting pressure required depends on the characteristics and shape of the particles, on the method of blending, and on the lubricant.

Press capacities are usually around 1.8 to 2.7 MN, although presses with much higher capacities are used for special applications. Most applications actually

TABLE 17.1

Compacting Pressures for Various Powders

Metal	Pressure (MPa)
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

require less than 0.9 MN. For small tonnage, crank- or eccentric-type mechanical presses are used; for higher capacities, toggle or knuckle-joint presses are employed (see Fig. 14.17b). Hydraulic presses (Fig. 17.12) with capacities as high as 45 MN can be used for large parts. Press selection depends on part size and the 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.

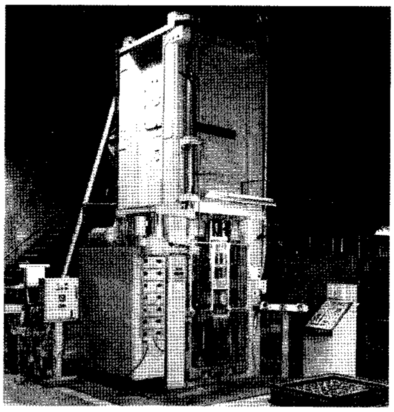


FIGURE 17.12 A 7.3-MN mechanical press for compacting metal powder. *Source*: Courtesy of Cincinnati Incorporated.

17.3.2 Isostatic Pressing

Green compacts may be subjected to *hydrostatic pressure* in order to achieve more uniform compaction and, hence, density.

In **cold isostatic pressing** (CIP), the metal powder is placed in a flexible rubber mold typically made of neoprene rubber, urethane, polyvinyl chloride, or another elastomer (Fig. 17.13). The assembly then is 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

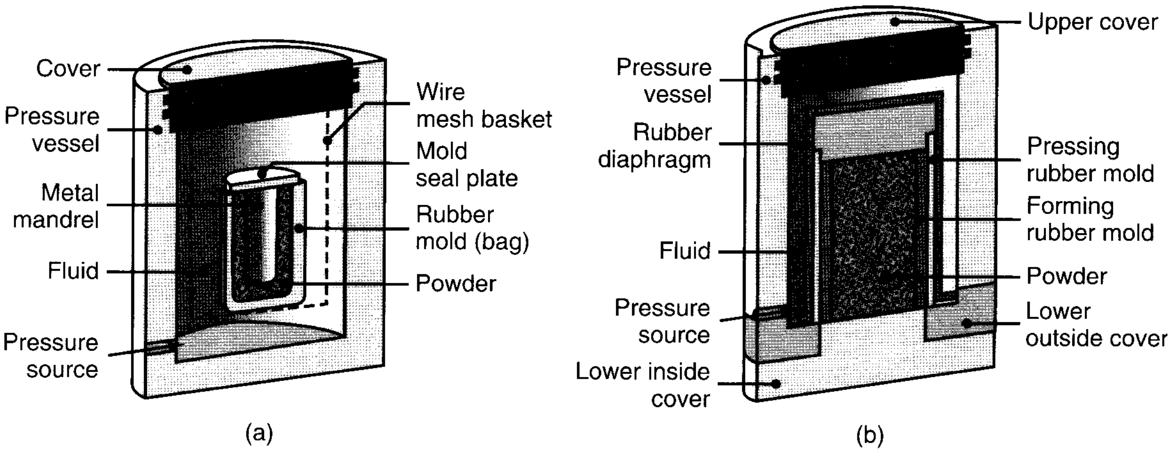


FIGURE 17.13 Schematic diagram of cold isostatic pressing. Pressure is applied isostatically inside a high-pressure chamber. (a) The wet bag process to form a cup-shaped part. The powder is enclosed in a flexible container around a solid-core rod. (b) The dry bag process used to form a PM cylinder.

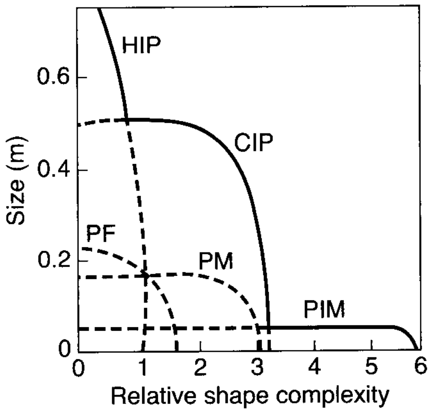


FIGURE 17.14 Capabilities, with respect to part size and shape complexity, available from various PM operations. PF = powder forging. *Source*: Courtesy of the Metal Powder Industries Federation.

ranges for CIP and other compacting methods in terms of the size and complexity of a part are shown in Fig. 17.14. A typical application is automotive cylinder liners.

In **hot isostatic pressing** (HIP), the container generally is made of a high-melting-point sheet metal and the pressurizing medium is high-temperature inert gas or a vitreous (glass-like) fluid (Fig. 17.15). Common conditions for HIP are pressures as high as 100 MPa—although they can be three times as high—and temperatures of 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. Consequently, it has gained wide acceptance in making high-quality parts.

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, (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.

The main advantages of hot isostatic pressing are the following:

- Because of the uniformity of pressure from all directions and the absence of die-wall friction, it produces fully dense compacts of practically uniform grain structure and density (hence, isotropic properties), irrespective of part shape. Parts with high length-to-diameter ratios have been produced with very uniform density, strength, toughness, and good surface detail.
- HIP is capable of handling much larger parts than those in other compacting processes.

The limitations of HIP are as follows:

- Wider dimensional tolerances are needed than those obtained in other compacting processes
- HIP requires higher equipment cost and production time than are required by other processes
- HIP is applicable only to relatively small production quantities, typically less than 10,000 parts per year.

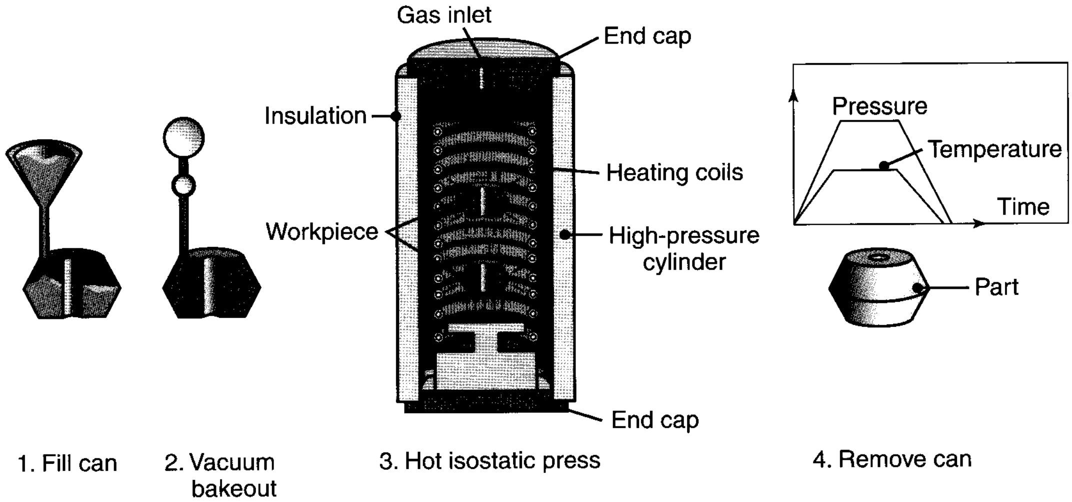


FIGURE 17.15 Schematic illustration of hot isostatic pressing. The pressure and temperature variation versus time are shown in the diagram.

EXAMPLE 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. 17.16. The 0.2-kg valve lifter rides on the camshaft and opens and closes the engine valves. Consequently, it is desired to have a tungsten-carbide face for wear resistance and a steel shaft for fatigue resistance. Before the HIP valve lifter was developed, parts were produced through furnace brazing, but resulted in occasional field failures and relatively high scrap rates. The required annual production of these parts is over 400,000, so high scrap rates are particularly objectionable.

The HIP-clad product consists of a 9% Co bonded tungsten-carbide (WC) face made from powder (pressed and sintered), a steel sheet-metal cap fitted over the WC disk, a copper-alloy foil interlayer, and a steel shaft. The steel cap is electron-beam welded to the steel shaft, and then the assembly is hot isostatically pressed to provide a very strong bond. The HIPing takes place at 1010°C 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 drastically reduced scrap rates to under 0.2%. No field failures have been experienced in over four years of full production. Production costs also were substantially reduced because of the hot isostatic pressing step.

Source: Courtesy of the Metal Powder Industries Federation.

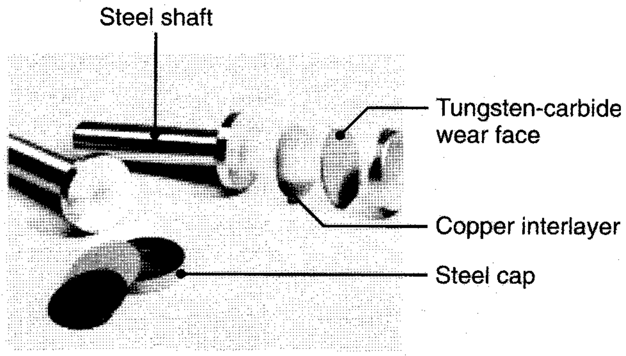


FIGURE 17.16 A valve lifter for heavy-duty diesel engines produced from a hot-isostatic-pressed carbide cap on a steel shaft. *Source*: Courtesy of the Metal Powder Industries Federation.

17.3.3 Miscellaneous Compacting and Shaping Processes

Powder-injection Molding. In this process, also called metal-injection molding (MIM), very fine metal powders (<10 μm) are blended with a 25 to 45% polymer or a wax-based binder. The mixture then undergoes a process similar to die casting (Section 11.3.5; see also *injection molding of plastics* in Section 19.3); it is injected into the mold at a temperature of 135° to 200°C. 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. The parts then are sintered in a furnace at temperatures as high as 1375°C. Subsequent operations (such as hole tapping, metal infiltration, and heat treating) also may be performed.

Generally, metals that are suitable for powder-injection molding (PIM) are those which melt at temperatures above 1000°C; examples are carbon and stainless steels, tool steels, copper, bronze, and titanium. Typical parts made are components for watches, 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 as follows:

- Complex shapes having wall thicknesses as small as 5 mm can be molded and then removed easily from the dies.
- Mechanical properties are nearly equal to those of wrought products.
- Dimensional tolerances are good.
- High production rates can be achieved by using multicavity dies.
- Parts produced by the PIM process compete well against small investment-cast parts, small forgings, and complex machined parts. However, the PIM process does not compete well with zinc and aluminum die casting or with screw machining.

The major limitations of PIM are the high cost and limited availability of fine metal powders.

EXAMPLE 17.2 Mobile Phone Components Produced through Metal Injection Molding

Figure 17.17 shows PM components used on mobile phones to allow them to flip open. The components were produced from 17-4 PH stainless steel, achieving a final density of 7.6 g/cm^3, a hardness of 30 HRC, and a tensile strength of 1,200 MPa. The parts shown are very complex, making them ideal candidates for MIM. The complexity arises because multiple components have been combined to form these parts, resulting in significant cost savings. The parts are produced within a maximum of $+0.23\%$ of the linear tolerance. The parts, as designed, have successfully completed accelerated life testing.

These components are integral to the unique opening mechanism deployed by the Motorola PEBL mobile phone. The dual hinge consists of the center barrel and a collar that extends to house intricate slots and ends with the two straight legs. This feature enables the single-movement opening of the clamshell mobile phone. To accomplish this, the components must exhibit both the capability of design flexibility and material strength. The only other process option was machining, which would have required considerably higher material utilization and added cost, estimated to be five times the MIM component cost.

Source: Courtesy of the Metal Powder Industries Federation.

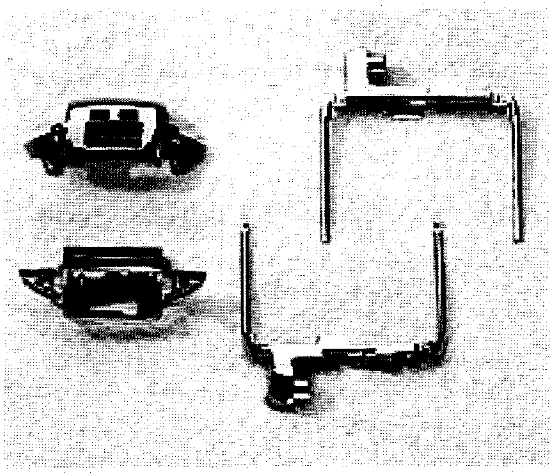


FIGURE 17.17 Powder-metal components for mobile phones to achieve a flip-open feature.

Forging. In *powder forging* (PF), the part produced from compaction and sintering serves as the preform in a hot-forging operation. These products are almost fully dense and have a good surface finish, good dimensional tolerances, and a uniform and fine grain size. The superior properties obtained make forging particularly suitable for such applications as highly stressed automotive (such as connecting rods) and jet-engine components.

Rolling. In *powder rolling* (also called **roll compaction**), the metal powder is fed into the roll gap in a two-high rolling mill (Fig. 17.18) and is compacted into a continuous strip at speeds of up to 0.5 m/s. The rolling process 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.

Extrusion. Powders can be compacted by *extrusion*, whereby the powder is encased in a metal container and hot extruded. After sintering, preformed PM parts may be reheated and forged in a closed die to their final shape. Superalloy powders, for example, are hot extruded for enhanced properties.

Pressureless Compaction. In *pressureless compaction*, the die is gravity filled with metal powder and the powder is 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. Spray deposition is a *shape-generation process* (Fig. 17.19). The basic components of the spray-deposition process for metal powders are (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.

Although there are several variations, the best known is the *Osprey process* shown in Fig. 17.19. After the metal is atomized, it is deposited onto a cooled preform

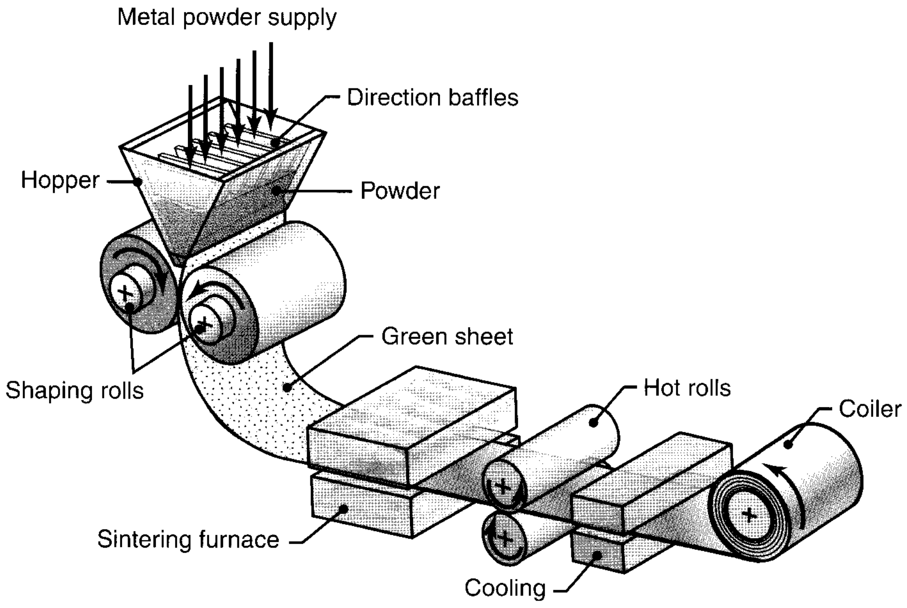


FIGURE 17.18 An illustration of powder rolling.

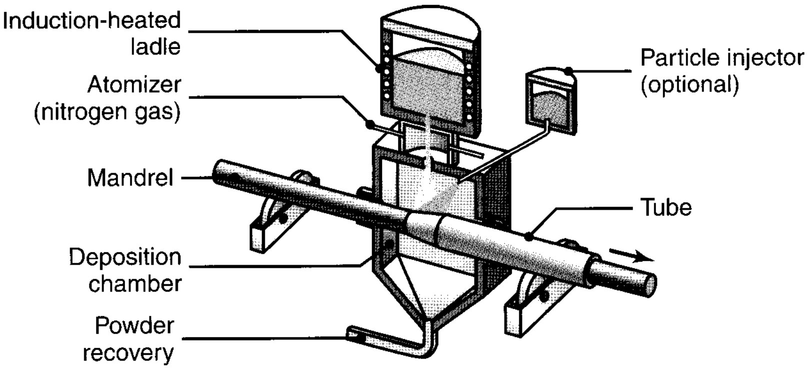


FIGURE 17.19 Spray deposition (*Osprey process*) in which molten metal is sprayed over a rotating mandrel to produce seamless tubing and pipe.

mold, usually made of copper or ceramic, where it solidifies. The metal particles bond together, developing a density that normally is above 99% of the solid-metal density. Spray-deposited forms may be subjected to additional shaping and consolidation processes, such as forging, rolling, and extrusion. The grain size is fine, and the mechanical properties are comparable to those of wrought products made of the same alloy.

Ceramic Molds. Ceramic molds for shaping metal powders are made by the technique used in investment casting. After the mold is made, it is filled with metal powder and placed in a steel container. The space between the mold and the container is filled with particulate material. The container is then evacuated, sealed, and subjected to hot isostatic pressing. Titanium-alloy compressor rotors for missile engines have been made by this process.

Selective Laser Sintering. Some powder-metal parts have been produced by *selective laser sintering*, a rapid-prototyping operation described in detail in Section 20.3.4.

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 5.7). 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 proper compaction and die life. Too large a clearance between the punch and the die will allow the metal powder to enter the gap, where it will severely interfere with the operation and cause eccentric parts. Diametral clearances are generally less than 25 μm. Die and punch surfaces must be lapped or polished (in the direction of tool movements in the die) for improved die life and overall performance.

17.4 Sintering

Sintering is the process whereby green compacts are heated in a controlled-atmosphere furnace to a temperature below the melting point, but sufficiently high to allow bonding (fusion) of the individual particles. As we have seen, the green

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

compact is brittle, and its *green strength* is low. The nature and strength of the bond between the particles and, hence, that of the sintered compact, depend on the complex mechanisms of diffusion, plastic flow, evaporation of volatile materials in the compact, recrystallization, grain growth, and pore shrinkage.

The principal variables in sintering are temperature, time, and the furnace atmosphere. Sintering 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 8 hours for tungsten and tantalum.

Continuous-sintering furnaces, which are used for most production, have three chambers:

1. *Burn-off chamber*, for volatilizing the lubricants in the green compact in order to improve bond strength and prevent cracking.
2. *High-temperature chamber*, for sintering.
3. *Cooling chamber*.

To obtain optimum properties, proper control of the furnace atmosphere is important for successful sintering. An oxygen-free atmosphere is essential to control the carburization and decarburization of iron and iron-based compacts and to prevent oxidation of the powders. A vacuum generally is 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.

Sintering mechanisms are complex and depend on the composition of the metal particles as well as on the processing parameters. The sintering 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. 17.20a); as a result, the strength, density, ductility, and thermal and electrical conductivities of the compact increase. At the same time, however, the compact shrinks. Hence, allowances should be made for shrinkage, as are done in casting.

A second sintering mechanism is **vapor-phase transport** (Fig. 17.20b). Because the material is heated to very close to its melting temperature, metal atoms will be 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

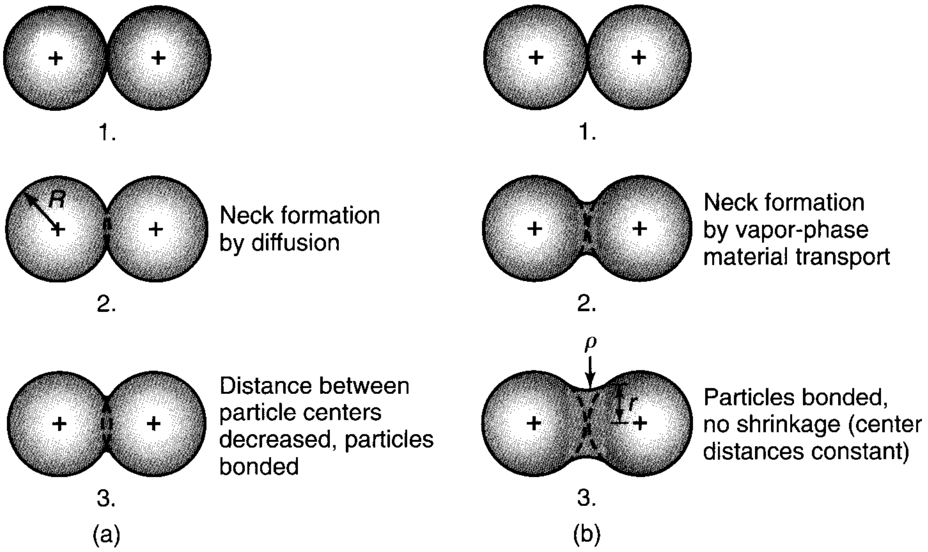


FIGURE 17.20 Schematic illustration of two mechanisms for sintering metal powders: (a) solid-state material transport; and (b) vapor-phase material transport. R = particle radius, r = neck radius, and ρ = neck-profile radius.

the other, the particle will melt and (because of surface tension) surround the particle that has not melted. An example of this mechanism, known as **liquid-phase sintering**, is cobalt in tungsten-carbide tools and dies (see Section 22.4). Stronger and denser parts can be obtained in this way. In **spark sintering** (an experimental process), loose metal powders are placed in a graphite mold, heated by electric current, subjected to a high-energy discharge, and compacted—all in one step. Another technique under development is **microwave sintering**, which reduces sintering times and thereby prevents grain growth, which can adversely affect strength.

Mechanical Properties. Depending on temperature, time, and the processing history, different structures and porosities can be obtained in a sintered compact and thus affect its properties. Porosity cannot be eliminated completely because (a) 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 material is less than 80% of its theoretical 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 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 remembered, however, that unless they are precision forged, forgings generally require some additional machining or finishing processes that a PM component may not require.

TABLE 17.3

Mechanical Properties of Selected PM Materials

Designation	MPIF type	Condition	Ultimate tensile strength (MPa)	Yield strength (MPa)	Elastic modulus (GPa)	Hardness	Elongation in 25 mm (%)
Ferrous					70		
FC-0208	N	AS	225	205	70	45 HRB	<0.5
		HT	295	—	110	95 HRB	<0.5
	R	AS	415	330	110	70 HRB	1
		HT	550	—	130	35 HRC	<0.5
	S	AS	550	395	130	80 HRB	1.5
		HT	690	655	145	40 HRC	<0.5
FN-0405	S	AS	425	240	145	72 HRB	4.5
		HT	1060	880	160	39 HRC	1
	T	AS	510	295	160	80 HRB	6
		HT	1240	1060		44 HRC	1.5
Aluminum							
601 AB, pressed bar		AS	110	48	—	60 HRH	6
		HT	252	241	—	75 HRH	2
Brass							
CZP-0220	T	—	165	76	—	55 HRH	13
	U	—	193	89	—	68 HRH	19
	W	—	221	103	—	75 HRH	23
Titanium							
Ti-6Al-4V		HIP	917	827	—	—	13
Superalloys							
Stellite 19		—	1035	—	—	49 HRC	<1

MPIF = Metal Powder Industries Federation, AS = as sintered, HT = heat treated, HIP = hot isostatically pressed.

TABLE 17.4

Comparison of Mechanical Properties of Some Wrought and Equivalent PM Metals (as Sintered)

Metal	Condition	Relative density (%)	UTS (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	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

Note: The density and strength of PM materials greatly increase with further processing, such as forging, isostatic pressing, and heat treatments.

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.

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 after sintering:

1. **Coining** and **sizing** are compacting operations performed under high pressure in presses. The purposes of these operations are to impart dimensional accuracy to the sintered part and to improve its strength and surface finish by further densification.
2. *Preformed and sintered* alloy-powder compacts subsequently may be cold or hot **forged** to the desired final shapes and sometimes by *impact forging*. These products have a good surface finish, good dimensional tolerances, and a 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 may be subjected to other finishing operations, such as
   - **Machining:** for producing various geometric features by milling, drilling, and tapping (to produce threaded holes).
   - **Grinding:** for improved dimensional accuracy and surface finish.
   - **Plating:** for improved appearance and resistance to wear and corrosion.
   - **Heat treating:** for improved hardness and strength.
4. The inherent porosity of PM components can be utilized 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. These bearings have a continuous supply of lubricant (due to capillary action) during their service lives (also referred to as *permanently lubricated*). Universal joints also are made by means of grease-impregnated PM techniques, thus no longer requiring traditional grease fittings.
5. **Infiltration** is a process whereby 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. The molten metal infiltrates the pores by capillary action and produces a relatively pore-free part having good density and strength. The most common application is the infiltration of iron-based compacts by copper or bronze. The advantages of infiltration are that the hardness

and tensile strength of the part are improved and the pores are filled, thus preventing moisture penetration (which could cause corrosion). Furthermore, since some porosity is desirable when an infiltrant is used, the part may be sintered only partially, resulting in lower thermal warpage.

6. **Electroplating** (Section 34.9) can be applied on PM parts, but special care is required to remove the electrolytic fluid, since it presents health hazards. Under some conditions, electroplating can seal a part and eliminate its permeability.

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. 17.21 through 17.23):

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 the compact. Thus, holes or recesses should be parallel to the axis of punch travel. Chamfers also should be provided to avoid damage to the edges during ejection.
3. PM parts should be made with the widest acceptable dimensional tolerances (consistent with their intended applications) in order to increase tool and die life and reduce production costs.
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 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.

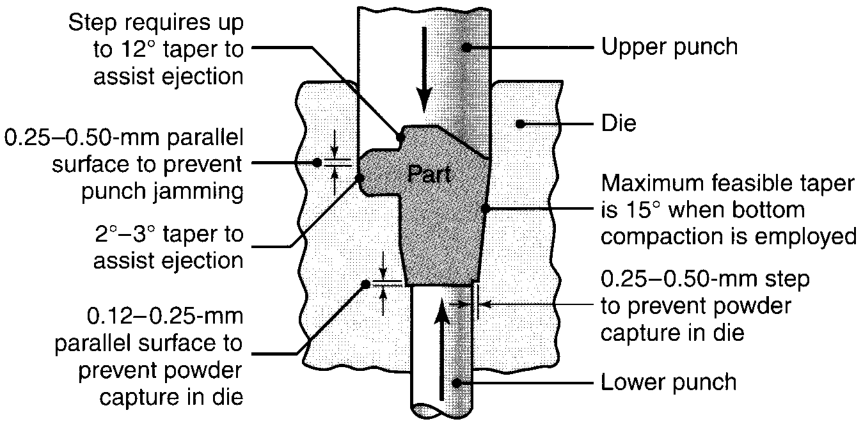


FIGURE 17.21 Die geometry and design features for powder-metal compaction.
Source: Courtesy of the Metal Powder Industries Federation.

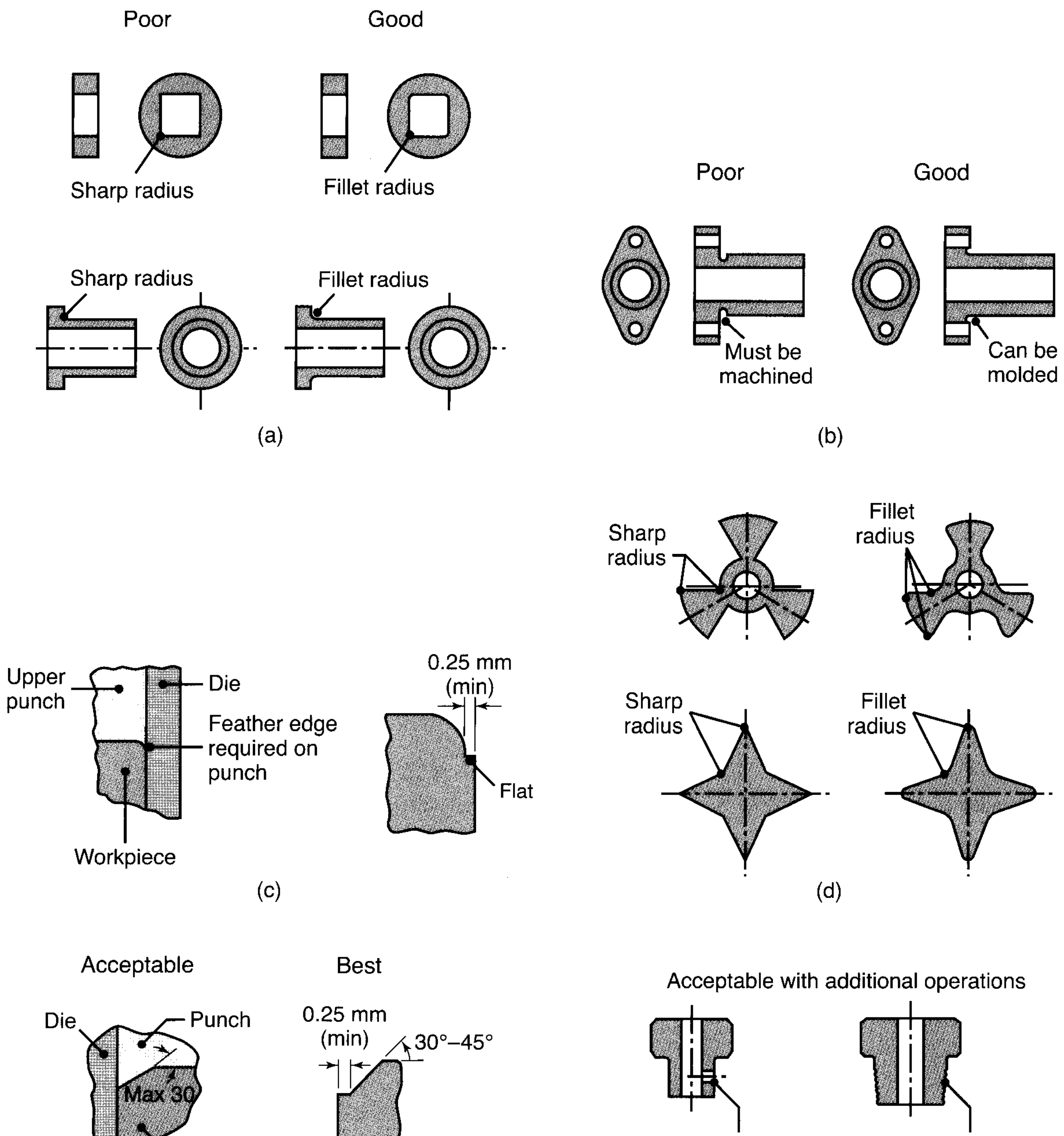


FIGURE 17.22 Examples of PM parts showing poor and good designs. Note that sharp radii and reentry corners should be avoided and that threads and transverse holes have to be produced separately by additional machining operations. *Source*: Courtesy of the Metal Powder Industries Federation.

6. Letters can be pressed if they are oriented perpendicular to the direction of pressing and can be raised or recessed. Raised letters are more susceptible to damage in the green stage and also prevent stacking during sintering.
7. Flanges or overhangs can be produced by a step in the die. However, long flanges can be broken upon ejection and may require more elaborate tooling. A long flange should incorporate a draft around the flange, a radius at the

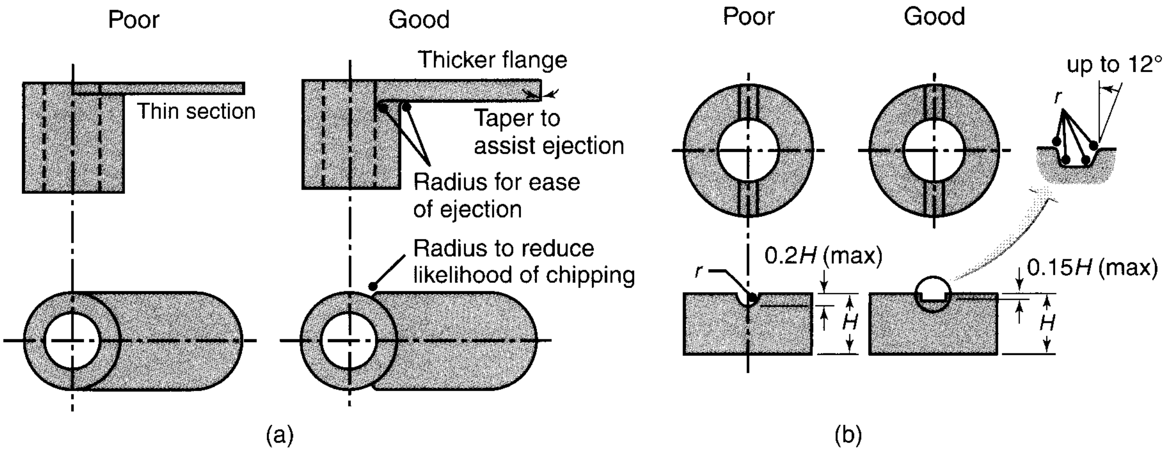


FIGURE 17.23 (a) Design features to use with unsupported flanges. (b) Design features for use with grooves. *Source*: Courtesy of Metal Powder Industries Federation.

bottom edge, and a radius at the juncture of the flange and/or component body to reduce stress concentrations and the likelihood of fracture.

8. A true radius cannot be pressed into the edge of a part because it would require the punch to be feathered (gently tapered) to a zero thickness, as shown in Fig. 17.22d. Chamfers or flats are preferred for pressing, and a 45° angle in a 0.25-mm flat is a common design approach.
9. Keys, keyways, and holes used for transmitting torques on gears and pulleys can be formed during powder compaction. Bosses can be produced, provided that proper drafts are used and their length is small compared to the overall component dimension.
10. Notches and grooves can be made if they are oriented perpendicular to the pressing direction. It is recommended that circular grooves not exceed a depth of 20% of the overall component, and rectangular grooves should not exceed 15%.
11. Parts produced through powder-injection molding have design constraints similar to those of parts produced by injection molding of polymers. With PIM, wall thicknesses should be uniform, to minimize distortion during sintering. Also, molds should be designed with smooth 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 improve significantly with additional operations, such as sizing, machining, and grinding.

17.7 Process Capabilities

The process capabilities of powder metallurgy may be summarized as follows:

- It is a technique for making parts from high-melting-point refractory metals and parts that may be difficult or uneconomical to produce by other methods.
- High production rates are possible on relatively complex parts using automated equipment and requiring little labor.
- Powder-metal processing offers good dimensional control and (in many instances) the elimination of machining and finishing operations; in this way, it reduces scrap and waste and saves energy.

- The availability of a wide range of compositions makes it possible to obtain special mechanical and physical properties, such as stiffness, vibration damping, hardness, density, toughness, and specific electrical and magnetic properties. Some of the newer highly alloyed superalloys can be manufactured into parts only by PM processing.
- It offers the capability of impregnation and infiltration for specific applications.

The limitations of PM are:

- The high cost of metal powder, particularly those for powder-injection molding, compared with that of raw materials to be cast or wrought.
- The high cost of tooling and equipment for small production runs.
- Limitations on part size and shape complexity.
- Mechanical properties, such as strength and ductility, that generally are lower than those obtained by forging. However, the properties of full-density PM parts made by HIP or by additional forging operations can be as good as those made by other processes.

17.8 Economics of Powder Metallurgy

Because powder metallurgy can produce parts at net or near-net shape and thus eliminate many secondary manufacturing and assembly operations, it increasingly has become 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.

As in other metalworking operations, the cost of dies and tooling in powder metallurgy depends on the part shape complexity and the method of processing the metal powders. Thus, tooling costs for processes such as hot isostatic pressing and powder-injection molding are higher than the more conventional powder processing. Because it is a near-net shape-manufacturing method, the cost of finishing operations in PM is low compared to other processes, such as casting and forging. However, if there are certain features to the part (such as threaded holes, undercuts, and transverse cavities and holes), then 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.

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 the part.

Labor costs are not as high in other processes, primarily because the individual operations (such as powder blending, compaction, and sintering) are performed on highly automated equipment. Thus, the skills required are not as high.

The near-net-shape capability of PM significantly reduces or eliminates scrap. For example, weight comparisons of 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.

TABLE 17.6

Forged and PM Titanium Parts and Cost Savings

Part	Weight (kg) 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

CASE STUDY 17.1 Powder Metallurgy Parts in a Snowblower

A collection of parts in the freewheeling steering system of a commercial snowblower is shown in Fig. 17.24. 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 to enhance 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 that form unique features and 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 (prealloyed iron with 1.9% nickel, 0.56% molybdenum, and 0.8% carbon with 2% admixed copper) with a tensile strength of 480 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, 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 long ID with a 0.05-mm tolerance. Both the pawl latch gear and the sprocket achieve a density of 6.7 g/cm^3 and a tensile strength of 690 MPa.

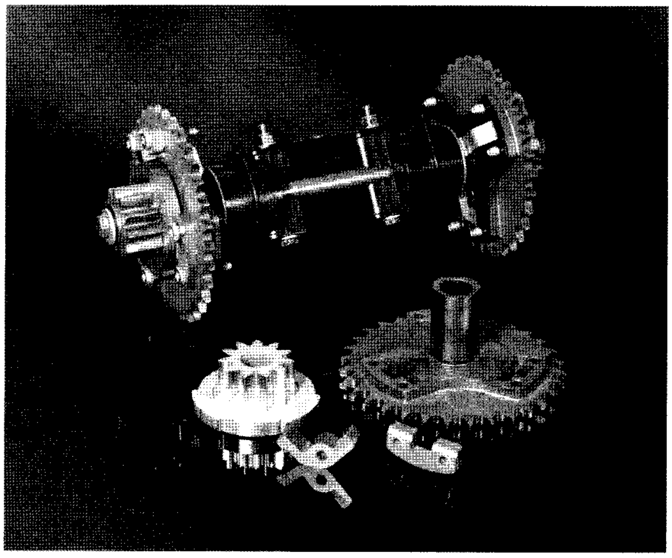


FIGURE 17.24 Collection of PM parts in a commercial snowblower.

All components passed normal life-cycle testing and product-life testing, including shock loading by engaging the drive in reverse when 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, as these are net-shape components. The only additional operations prior to final assembly are vibratory deburring and honing of the 32-tooth sprocket 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.

SUMMARY

- Powder metallurgy is a net-shape forming process consisting of metal powder production, blending, compaction in dies, and sintering to impart strength, hardness, and toughness. Compaction also may be carried out by cold or hot isostatic pressing for improved properties. Although the size and the weight of its products are limited, the PM 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 (as with low-melting-point metals).
- Control of powder shape and quality, process variables, and sintering atmospheres are important considerations in product quality. Density and mechanical and physical properties can be controlled by tooling design and by compacting pressure.
- An important PM process is powder-injection molding, which involves mixing the very fine metal powders with polymers to make them flow more easily into molds of complex shape during compaction. The polymers subsequently evaporate during sintering.
- 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. It has some competitive advantages over other methods of production, such as casting, forging, and machining.

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

ASM Handbook, Vol. 7: **Powder Metal Technologies and Applications**, ASM International, 1998.

Atkinson, H.V., and Rickinson, B.A., **Hot Isostatic Pressing**, Adam Hilger, 1991.

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., **A-Z of Powder Metallurgy**, Elsevier, 2007.

———, **Powder Metallurgy and Particulate Materials Processing**, Metal Powder Industries Federation, 2006.

———, **Powder Metallurgy of Iron and Steel**, Wiley, 1999.

———, and Bose, A., **Injection Molding of Metals and Ceramics**, Metal Powder Industries Federation, 1997.

Kang, S.-J., **Sintering: Densification, Grain Growth and Microstructure**, Butterworth-Heinemann, 2005.

Kuhn, H.A., and Ferguson, B.L., **Powder Forging**, Metal Powder Industries Federation, 1990.

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.

Powder Metallurgy Design Manual, 2nd ed., Metal Powder Industries Federation, 1995.

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 hazards are involved in PM processing? Explain their causes.

17.6. Describe what occurs to metal powders during sintering.

17.7. Why might secondary and finishing operations be performed on PM parts?

17.8. Explain the difference between impregnation and infiltration. Give some applications of each.

17.9. What is mechanical alloying? What are its advantages over the conventional alloying of metals?

17.10. 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

17.11. Why is there density variation in the compacting of powders? How is it reduced?

17.12. What is the magnitude of the stresses and forces involved in powder compaction?

17.13. Give some reasons that powder-injection molding is an important process.

17.14. How does the equipment used for powder compaction vary from those used in other metalworking operations in the preceding chapters?

17.15. Explain why the mechanical and physical properties depend on their density.

17.16. What are the effects of the different shapes and sizes of metal particles in PM processing?

17.17. Describe the relative advantages and limitations of cold and hot isostatic pressing.

17.18. How different, if any, are the requirements for punch and die materials in powder metallurgy from those for forging and extrusion operations? Explain.

17.19. The powder metallurgy process can be competitive with processes such as casting and forging. Explain why this is so.

17.20. What are the reasons for the shapes of the curves shown in Fig. 17.10 and for their relative positions on the charts?

17.21. Should green compacts be brought up to the sintering temperature slowly or rapidly? Explain your reasoning.

17.22. 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?

17.23. In Fig. 17.11e, 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?

17.24. Why do the compacting pressure and the sintering temperature depend on the type of powder metal?

QUANTITATIVE PROBLEMS

▶**17.25.** Estimate the maximum tonnage required to compact a brass slug 75 mm in diameter. Would the height of the slug make any difference in your answer? Explain your reasoning.

▶**17.26.** Refer to Fig. 17.10a. What should be the volume of loose, fine iron powder in order to make a solid cylindrical compact 25 mm in diameter and 20 mm high?

17.27. 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.

17.28. Estimate the number of particles in a 500-g sample of iron powder if the particle size is 75 μm.

17.29. Assume that the surface of a copper particle is covered by an oxide layer 0.1 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?

▶**17.30.** A coarse copper powder is compacted in a mechanical press at a pressure of 275 MPa. During sintering, the green part shrinks an additional 7%. What is the final density?

▶**17.31.** 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 90 mm in diameter and has a 15-mm hub, what is the required press force?

17.32. 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

17.33. Make sketches of PM products in which density variations (see Fig. 17.11) would be desirable. Explain why in terms of the functions of these parts.

17.34. Compare the design considerations for PM products with those for (a) casting and (b) forging. Describe your observations.

17.35. Are there applications in which you, as a manufacturing engineer, would not recommend a PM product? Explain.

17.36. Describe in detail other methods of manufacturing the parts shown in Fig. 17.1.

17.37. Using the Internet, locate suppliers of metal powders and compare the cost of the powder with the cost of ingots for five different materials.

17.38. Explain why powder-metal parts are commonly used for machine elements requiring good frictional and wear characteristics and for mass-produced parts.

17.39. 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.

17.40. Describe how the information given in Fig. 17.14 would be helpful to you in designing PM parts.

17.41. It was stated that, in the process shown in Fig. 17.19, 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.

Ceramics, Glasses, and Superconductors: Processing and Equipment

CHAPTER 18

- This chapter presents the manufacturing processes associated with 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, followed by finishing operations on ceramics, are also discussed.
- Glass manufacture involves production of continuous shapes, such as plate, tube, and bars, through drawing, rolling, or floating methods; for discrete products, the operations 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, rapid prototyping.

18.1 Introduction

The properties and various applications of ceramics and glasses were 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 parts such as floor tiles, dishes, electrical insulators, spark plugs and ball bearings.

In this chapter, the techniques that are available for processing ceramics, glasses (Fig. 18.1) and superconductors into numerous useful products are described. The methods employed for ceramics consist of crushing the raw materials; shaping them by various means; and drying, firing, and then applying finishing operations, as needed, to achieve the required dimensional tolerances and surface finish.

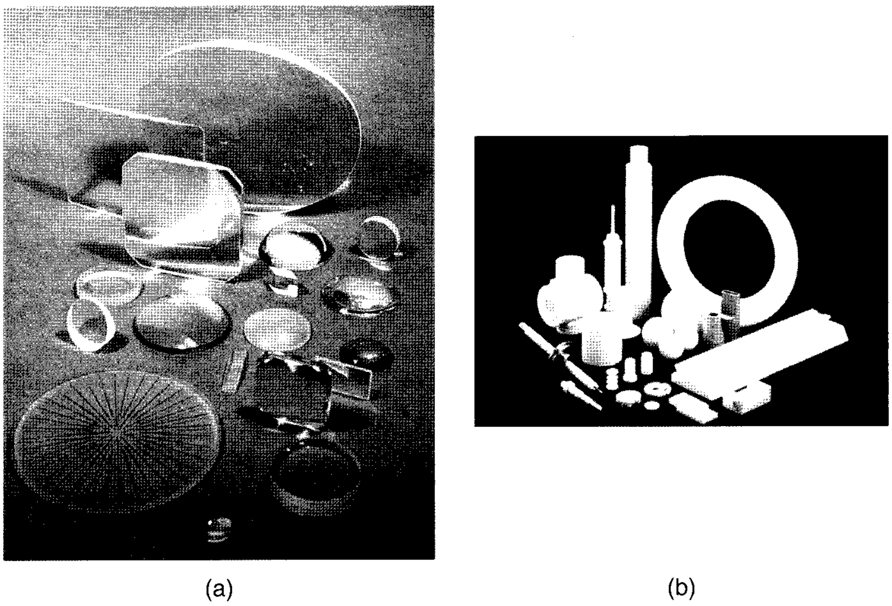

(a) (b)

FIGURE 18.1 (a) Examples of typical glass parts. (b) Examples of ceramic parts. *Source*: (a) Courtesy of Commercial Optical Manufacturing, Inc. (b) Courtesy Kyocera.

For glasses, the processes involve mixing and melting the raw materials in a furnace and shaping them in molds and by various techniques, depending on the shape and size of the part. Both discrete products (such as bottles) and continuous products (such as flat glass, rods, tubing, and fibers) can be produced. Glasses also are strengthened by thermal and chemical means, as well as by lamination with polymer sheets (as is done with 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 ceramics involved and their shapes. Production of some ceramic parts (such as pottery, ovenware, or floor tiles) generally does

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

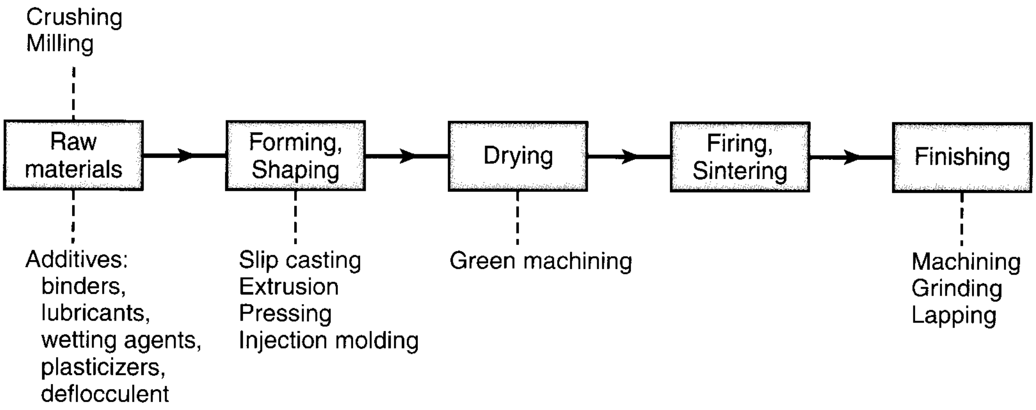


FIGURE 18.2 Processing steps involved in making ceramic parts.

not involve the same level of control of materials and processes as do high-tech components (made of such structural ceramics as silicon nitride and silicon carbide) and cutting tools (using, for example, aluminum oxide). Generally, however, the procedure involves the following steps (Fig. 18.2):

1. Crushing or grinding the raw materials into very fine particles
2. Mixing them with additives to impart certain desirable characteristics
3. Shaping, drying, and firing the material.

Ceramics may be subjected to additional processing, such as machining and grinding, for better control of dimensions and surface finish.

The first step in processing ceramics is the *crushing* (also called *comminution* or *milling*) of the raw materials. Crushing generally is 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 also prevents the suspension of fine particles in the air. The particles then may be *sized* (passed through a sieve), filtered, and washed.

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 to help remove the part from the mold
- **Wetting agent**: to improve mixing
- **Plasticizer**: to make the mix more plastic and formable
- **Agents**: to control foaming and sintering.
- **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). Water is added to make the mixture more pourable and less viscous. Typical deflocculents are Na_2CO_3 and Na_2SiO_3 in amounts of less than 1%.

The three basic shaping processes for ceramics are casting, plastic forming, and pressing.

18.2.1 Casting

The most common casting process is **slip casting** (also called **drain casting**), as illustrated in Fig. 18.3. A **slip** is a suspension of *colloidal* (small particles that do not settle) ceramic particles in an *immiscible liquid* (insoluble in each other), which is

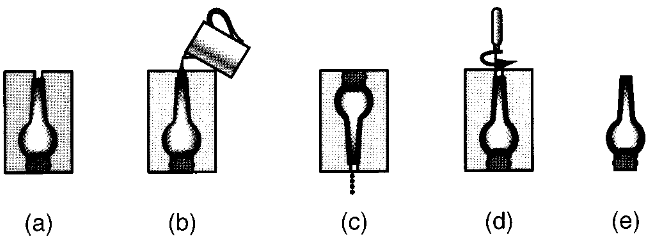


FIGURE 18.3 Sequence of operations in slip casting a ceramic part. After the slip has been poured, the part is dried and fired in an oven to give it strength and hardness. *Source*: After F.H. Norton.

generally water. The slip is poured into a porous mold, typically made of plaster of paris. Molds also may consist of several components.

The slip must have sufficient fluidity and low enough viscosity 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 properly, as air entrapment can be a significant problem 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 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 tool in Fig. 18.3d), the mold is opened, and the part is removed.

Large and complex parts (such as plumbing ware, art objects, and dinnerware) 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. Molds also may consist of multiple components.

For solid-ceramic parts, the slip is supplied continuously into the mold to replenish the absorbed water; 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 and is associated with the light-green tint in decorative ceramic slip casts at this stage. While the ceramic parts are still green, they may be machined to produce certain features or to give dimensional accuracy to the parts. Because of the delicate nature of the green compacts, however, machining usually is done manually or with simple tools. For example, the flashing in a slip casting may be removed gently with a fine wire brush, or holes can be drilled in the mold. Detailed work (such as the tapping of threads) generally is not done on green compacts because warpage (due to firing) makes such machining not viable.

Doctor-blade Process. Thin sheets of ceramics [less than 1.5 mm thick] can be made by a casting technique called the *doctor-blade process* (Fig. 18.4). The slip is cast over a moving plastic belt while its thickness is controlled by a blade.

Ceramic sheets also may be produced by other methods, including (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 various methods, such as extrusion, injection molding, or molding and jiggering (Fig. 18.5). Plastic forming tends to orient the layered structure of clay along the direction of material flow and, hence, 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 by a screw-type piece of equipment. The cross section of the extruded product is constant, and there are limitations to wall thickness for hollow extrusions. The extruded products may be subjected to additional shaping operations. Tooling costs are low, and production rates are high.

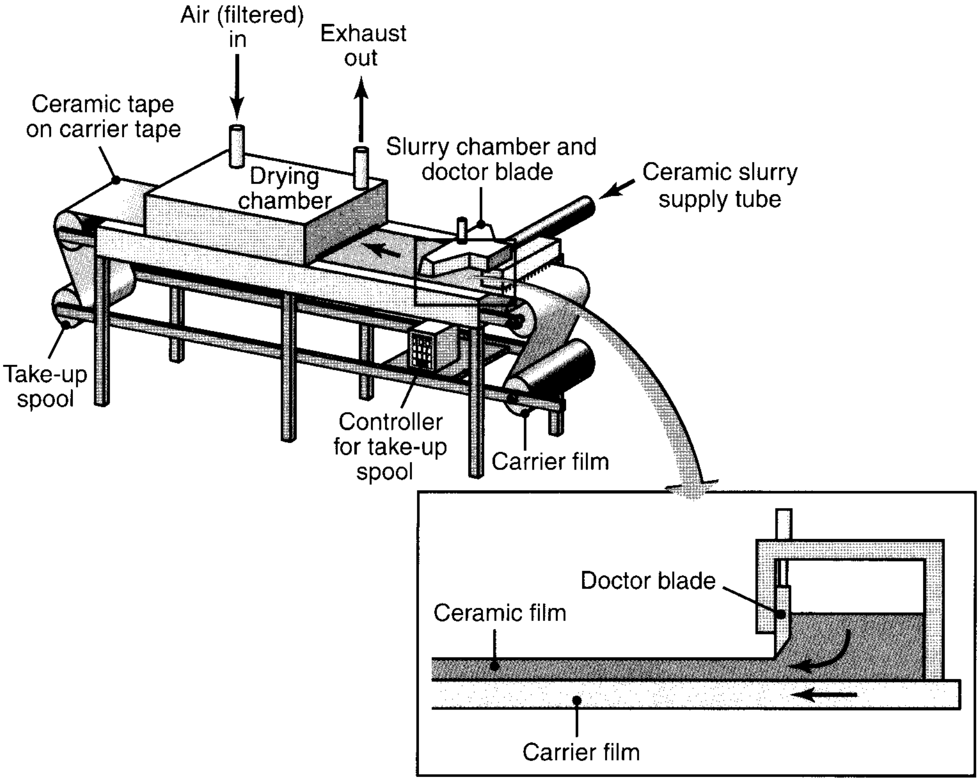


FIGURE 18.4 Production of ceramic sheets through the doctor-blade process.

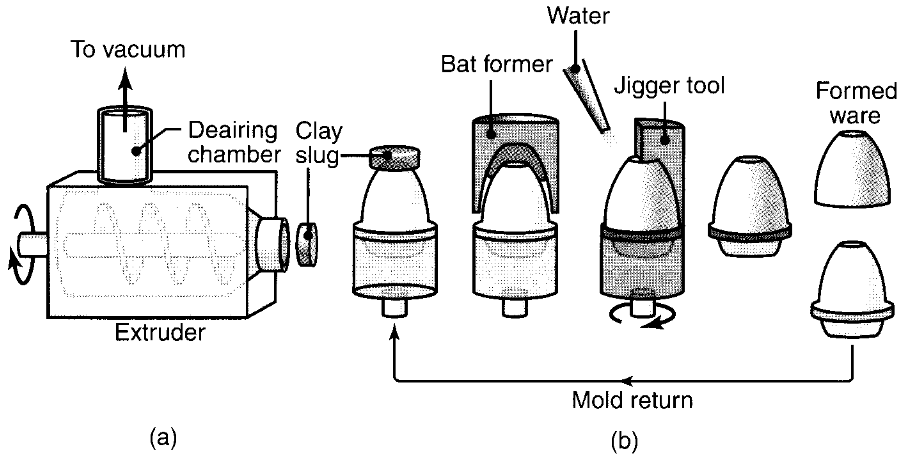


FIGURE 18.5 (a) Extruding and (b) jiggering operations. *Source*: After R.F. Stoops.

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 generally is below 4%, but it may be as high as 12%. Organic and inorganic binders (such as stearic acid, wax, starch, and polyvinyl alcohol) usually are added to the mixture; these additives also act as lubricants. This process has the same high production rates and close control of dimensional tolerances as does powder metallurgy.

The pressing pressure ranges from 35 to 200 MPa. Modern presses used for dry pressing are highly automated. Dies usually are made of carbides or of hardened steel. They must have high wear resistance in order to withstand the abrasive ceramic particles; hence, they can be expensive.

Density can vary significantly in dry-pressed ceramics (as in PM compaction) because of friction among the particles and at the mold walls. Density variations cause warping during the firing, which is particularly severe for parts having high length-to-diameter ratios. The recommended maximum ratio is 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.

Wet Pressing. In *wet pressing*, the part is formed in a mold while under high pressure in a hydraulic or mechanical press. This process generally is used to make intricate shapes. Moisture content usually ranges from 10 to 15%. Production rates are high; however, (a) part size is limited, (b) dimensional control is difficult to achieve because of shrinkage during drying, and (c) tooling costs can be high.

Isostatic Pressing. Used extensively in powder metallurgy, *isostatic pressing* also is used for ceramics in order to obtain a uniform density distribution throughout the part during compaction. For example, automotive spark-plug insulators are made by this method at room temperature, while silicon-nitride vanes for high-temperature applications (see Fig. 8.1) are made by *hot isostatic pressing*.

Jiggering. A series of steps is needed to make ceramic plates. First, clay slugs are extruded and formed into a *bat* over a plaster mold. Then they are jiggered on a rotating mold (see Fig. 18.5b). *Jiggering* is a motion in which the clay bat is formed by means of templates or rollers. The part then is dried and fired. The jiggering process is confined to axisymmetric parts and has limited dimensional accuracy. The operation is automated for improved productivity.

Injection Molding. *Injection molding* is used extensively for the precision forming of ceramics in high-technology applications, such as for rocket-engine components. The raw material is mixed with a binder, such as a thermoplastic polymer (polypropylene, low-density polyethylene, or ethylene vinyl acetate) or wax. The binder usually is removed by pyrolysis (inducing chemical changes by heat); the part is then sintered by firing.

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. Thicker sections require careful control of the materials used and of the processing parameters in order to avoid defects, such as internal voids and cracks—especially those due to shrinkage.

Hot Pressing. In this process (also called *pressure sintering*), the pressure and the heat are applied simultaneously, thereby reducing porosity and making the part denser and stronger. Graphite commonly is used as a punch and die material, and protective atmospheres usually are employed during 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.

18.2.4 Drying and Firing

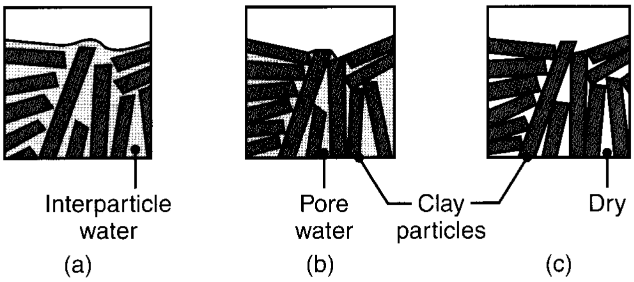


FIGURE 18.6 Shrinkage of wet clay caused by the removal of water during drying. Shrinkage may be as much as 20% by volume. *Source*: After F.H. Norton.

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 because of the tendency for the part to warp or crack from variations in moisture content and in thickness. Control of atmospheric humidity and of ambient temperature 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 consequently, the moisture gradient across the thickness of the part is lower than that in a dry environment. This low moisture gradient, in turn, prevents a large, uneven gradient in shrinkage from the surface to the interior during drying.

A ceramic part that has been shaped by any of the methods described thus far is in the *green* state. The part 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 its relative softness.

Firing (also called **sintering**) involves heating the part to an elevated temperature in a controlled environment. Some shrinkage occurs during firing. Firing gives the ceramic part its strength and hardness. This improvement in properties results from (a) the development of a strong bond between the complex oxide particles in the ceramic and (b) reduced porosity. A more recent technology (although not yet commercialized) involves the **microwave sintering** of ceramics in furnaces operating at more than 2 GHz. Its cost-effectiveness depends on the availability of inexpensive furnace insulation.

Nanophase ceramics (described in Section 8.2.5) can be sintered at lower temperatures than those used for conventional ceramics. They are easier to fabricate, because they can be compacted at room temperature to high densities, hot pressed to theoretical density, and formed into net-shaped parts without using 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) improve its surface finish and dimensional tolerances, and (c) remove any surface flaws. Although they 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 a 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, which are described in detail in various sections in Part IV of this book:

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 some of these processes. The effect of the finishing operation on the properties of the product also must be considered. For instance, because of notch sensitivity, the finer the finish, the higher the part's strength and load-carrying capacity—particularly its fatigue strength (see Fig. 2.29). Ceramic parts also may undergo static fatigue, as described for glass in Section 18.5.

To improve their appearance and strength and to make them impermeable, ceramic products often are coated with a **glaze** or **enamel**, which forms a glassy coating after firing.

EXAMPLE 18.1 Dimensional Changes During the Shaping of Ceramic Components

A solid, cylindrical ceramic part is to be made with a final length, L, of 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. 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, we can write

$$\frac{(L_d - L)}{L_d} = 0.06,$$

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 fired volume of the part. Because the linear shrinkage during firing is 6%, we can determine the dried volume, V_d, of the part as

$$V_d = \frac{V_f}{(1 - 0.06)^3} = 1.2V_f.$$

Hence,

$$\frac{V_a}{V_d} = \frac{0.97}{1.2} = 0.81, \text{ or } 81\%.$$

Therefore, the porosity, P_d, of the dried part is 19%.

18.3 Forming and Shaping of Glass

Glass is processed by melting and then shaping it, either in molds, with tools, or by blowing. Glass shapes produced include flat sheets and plates, rods, tubing, glass fibers, and discrete products such as bottles, light bulbs, and headlights. 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 (which induce 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, such as window glass, glass doors, and tabletops.
2. **Rods** and **tubing** used for chemicals, neon lights, and decorative artifacts.
3. **Discrete products** such as bottles, vases, headlights, and television tubes.
4. **Glass fibers** to reinforce composite materials (Section 9.2.1) and for use in fiber optics.

All glass forming and shaping processes begin with molten glass, typically in the range from 1000° to 1200°C. The glass has the appearance of a red-hot, viscous syrup and is supplied from a melting furnace or tank.

18.3.1 Flat-sheet and Plate Glass

Flat-sheet glass can be made by the float glass method or by drawing or rolling it from the molten state (all three methods are continuous processes):

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 then moves at a temperature of about 650°C over rollers into another chamber (*lehr*), where it solidifies. *Float glass* has smooth (*fire-polished*) surfaces, so further grinding or polishing is 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 sheets or plates involves passing the molten glass through a pair of rolls (Fig. 18.8a) in an arrangement similar to an old-fashioned clothes wringer. The solidifying glass is squeezed between these two rolls (forming it into a sheet) and then moved 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. The surfaces of the glass may be embossed with a pattern by using textured roller surfaces. In this way, the glass surface becomes a replica of the roll surface. Thus, glass sheet produced by drawing or rolling has a rough surface appearance.

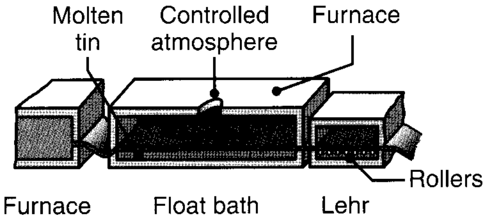


FIGURE 18.7 The float method of forming sheet glass.

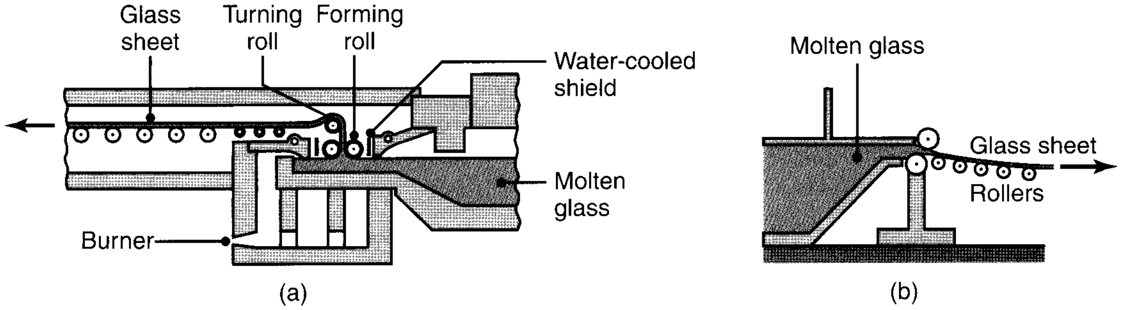


FIGURE 18.8 (a) Drawing process for drawing sheet glass from a molten bath. (b) Rolling process. *Source*: After W.D. Kingery.

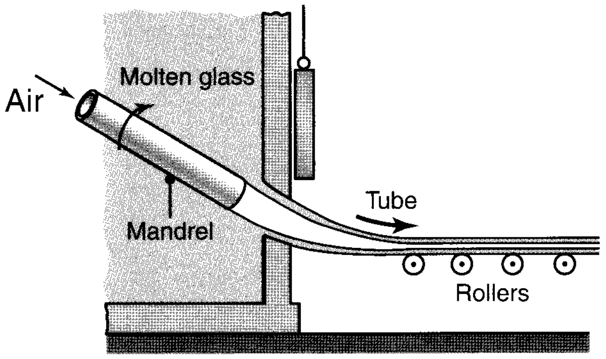


FIGURE 18.9 Manufacturing process for glass tubing. Air is blown through the mandrel to keep the tube from collapsing. Glass tubes for fluorescent bulbs are made by this method.

18.3.2 Tubing and Rods

Glass tubing is manufactured by the process shown in Fig. 18.9. Molten glass is wrapped around a rotating (cylindrical or cone-shaped) hollow mandrel and is drawn out by a set of rolls. Air is blown through the mandrel to prevent the glass tube from collapsing. These machines may be horizontal, vertical, or slanted downward. This is the method used in making the glass tubes for fluorescent bulbs, with machines (such as the Corning Ribbon Machine) capable of producing bulbs at rates of 2000 per minute. Glass rods are made in a similar manner, but air is not blown through the mandrel. The drawn product becomes a solid glass rod.

An alternative process for making tubes involves *extrusion* of a strip of glass, which is then wrapped obliquely around a rotating mandrel. The molten glass blends across adjacent layers, and the resultant tube is drawn off the mandrel in a continuous process.

18.3.3 Discrete Glass Products

Several processes are used to make discrete glass objects.

Blowing. Hollow and thin-walled glass items (such as bottles, vases, and flasks) are made by *blowing*—a process that is similar to the 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 usually is coated with a parting agent (such as oil or emulsion) to prevent the glass from sticking to the mold. Blowing may be followed by a second blowing operation for finalizing product shape, called the *blow and blow* process.

The surface finish of products made by the blowing process is acceptable for most applications, such as bottles and jars. It is difficult to control the wall thickness of the product, but this process is economical for high-rate production. Incandescent light bulbs (see Fig. I.2) are made in highly automated blowing machines at a rate of greater than 2000 bulbs per minute.

Pressing. In the *pressing* process, a gob of molten glass is placed into a mold and pressed into a confined cavity with a plunger. Thus, the process is similar to closed-die forging. 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 being pressed, the solidifying glass acquires the shape of the mold-plunger cavity. Because of the confined environment, the product has a higher dimensional accuracy than can be obtained with blowing. Pressing in one-piece molds cannot be used for (a) shapes of products from which the plunger cannot be retracted or (b) thin-walled items. For example, split molds are used for bottles, while, for thin-walled items, pressing can be combined with blowing. In the latter process, known as *press and blow*, the pressed part is subjected to air pressure (hence the term blow), which further expands the glass into the mold.

Centrifugal Casting. Also known in the glass industry as **spinning** (Fig. 18.13), this process is similar to that used for metals. The centrifugal force pushes the molten glass against the mold wall, where it solidifies. Typical products made are large lenses for research telescopes and some architectural shapes.

Sagging. Shallow dish-shaped or lightly embossed glass parts can be made by the *sagging* process. A sheet of glass is placed over a mold and heated. The glass sags by its own weight and takes the shape of the mold. The process is similar to the

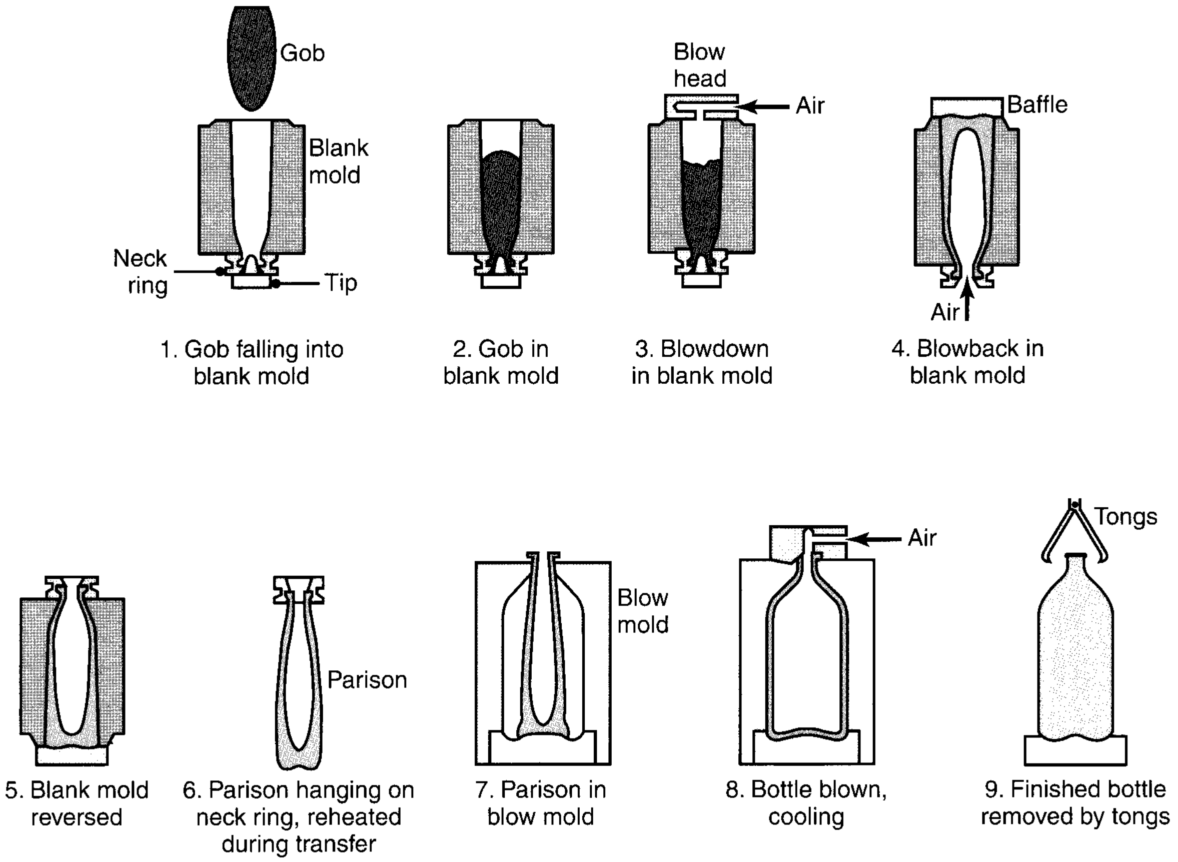


FIGURE 18.10 Steps in manufacturing an ordinary glass bottle. *Source*: After F.H. Norton.

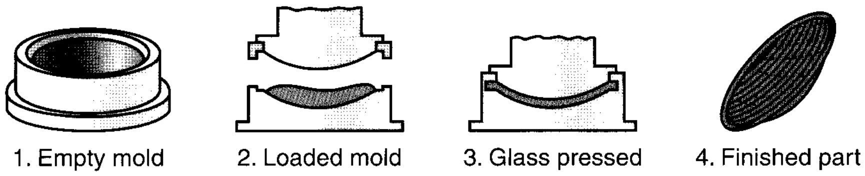


FIGURE 18.11 Manufacturing a glass item by pressing glass into a mold. *Source*: Courtesy of Corning Glass Works.

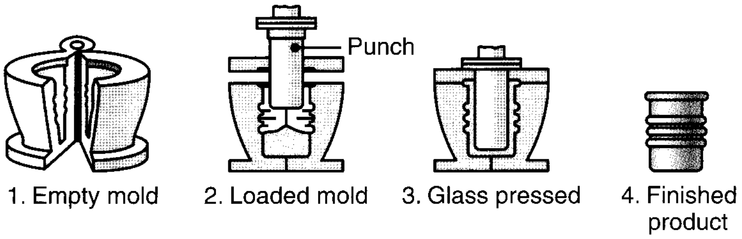


FIGURE 18.12 Pressing glass into a split mold. *Source*: After E.B. Shand.

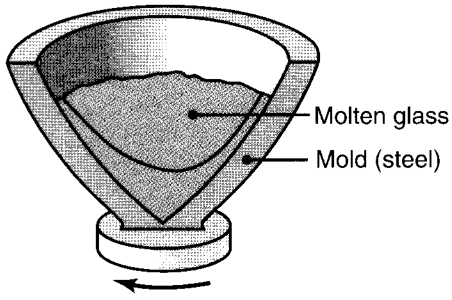


FIGURE 18.13 Centrifugal casting of glass. Large telescope lenses and television-tube funnels are made by this process. *Source*: Courtesy of Corning Glass Works.

thermoforming of thermoplastics (Section 19.6), but no pressure or vacuum is involved. Typical applications 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. thus, 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).

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. In order to protect their surfaces, fibers subsequently are coated with chemicals. Short fibers (*chopped*) are produced by subjecting long fibers to compressed air or steam as they leave the orifice.

Glass wool (short glass fibers)—used as a thermal insulating material or for acoustic insulation—is made by a *centrifugal spraying* process in which molten glass is ejected (spun) from a rotating head. The diameter of the fibers is typically in the range from 20 to 30 μm.

18.4 Techniques for Strengthening and Annealing Glass

Glass can be strengthened by a number of processes, and discrete glass products may be subjected to annealing and to other finishing operations to impart desired properties and surface characteristics.

Thermal Tempering. In this process (also called *physical tempering* or *chill tempering*), the surfaces of the hot glass are cooled rapidly by a blast of air (Fig. 18.14). As a result, the surfaces shrink and (at first) tensile stresses develop on the surfaces. As the bulk of the glass begins to cool, it contracts. The already solidified surfaces of the glass are then forced to contract, and consequently, they develop residual compressive surface stresses, 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.

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 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 a large number of pieces when broken. The broken pieces are not as sharp and hazardous as those from ordinary window glass.

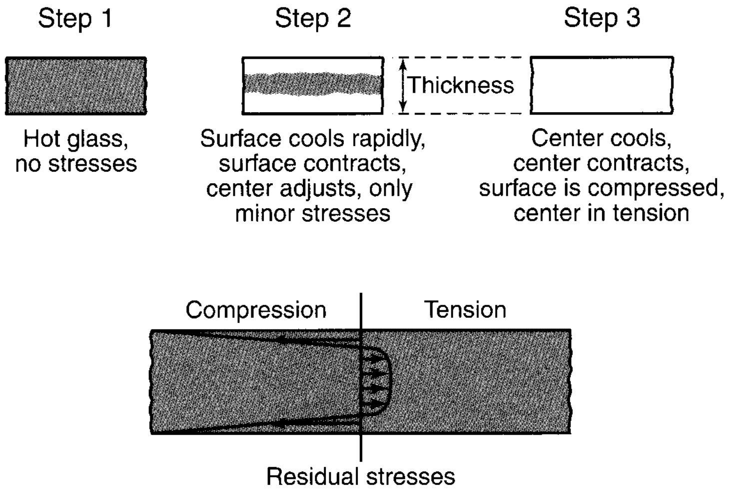


FIGURE 18.14 (a) The stages involved in inducing compressive surface residual stresses for improved strength. (b) Residual stresses in a tempered glass plate.

Chemical Tempering. In this process, the glass is heated in a bath of molten KNO_3, K_2SO_4, or $NaNO_3$, depending on the type of glass. 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. This condition is similar to that created by forcing a wedge between two bricks in a brick wall.

The time required for chemical tempering is about one hour longer than that for thermal tempering. Chemical tempering may be performed at various temperatures. At low temperatures, part distortion is minimal; therefore, complex shapes can be treated. 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.

Laminated Glass. Laminated glass, a product of another strengthening method called *laminate strengthening*, it consists of two pieces of flat glass with a thin sheet of tough plastic in between. When laminated glass is cracked, its pieces are held together by the plastic sheet—a phenomenon commonly observed in a shattered automobile windshield.

Traditionally, flat glass for glazing windows and doors has been strengthened with wire netting (such as *chicken wire*—with a hexagonal mesh) embedded in the glass during its production. When a hard object strikes the surface, the glass shatters, but the pieces are held together because of the toughness of the wire, which has both strength and ductility.

Bulletproof Glass. Laminated glass has considerable ballistic impact resistance and can prevent the full penetration of solid objects because of the presence of a tough polymer film in between the two layers of glass. *Bulletproof glass* (used in some automobiles, armored bank vehicles, and buildings) is a more challenging design, due to the very high speed and energy level of the bullet and the small size and the shape of the bullet tip, representing a small contact area and high localized stresses. Depending on the caliber of the weapon, bullet speeds range from about 350 to 950 m/s.

Bulletproof glass (also called *bullet-resistant glass*) ranges in thickness from 7 to 75 mm. The thinner plates are designed for resistance to handguns, and thicker plates are for rifles. Although there are several variations, bulletproof glass basically consists of glass laminated with a polymer sheet (usually polycarbonate). 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 sheets commonly are used for bulletproof glass. As a material widely used for safety helmets, windshields, and guards for machinery, polycarbonate is a tough and flexible polymer. Combined with a thick glass, it can stop a bullet, although the glass develops a circular shattered region. Proper bonding of these layers over the glass surface is also an important consideration, as there usually is more than one round fired during such encounters. Also, in order to maintain the transparency of the glass and minimize distortion, the index of refraction of the glass and the polymer must be nearly identical.

If a polymer sheet is only on one side of the glass, it is known as a *one-way* bulletproof glass. In a vehicle, the polymer layer is on the inside surface of the glass. An external bullet will not penetrate the window, because the bullet will strike the glass first, shattering it. The glass absorbs some of the energy of the bullet, thus slowing it down. The remaining energy is dissipated in the polymer sheet, which then stops the bullet. This arrangement allows someone inside the vehicle to fire back. A bullet from inside penetrates the polymer sheet and forces the glass to break outwards, allowing the bullet to go through. Thus, a one-way glass stops a bullet fired from outside but allows a bullet to be fired from inside.

A more recent design for bulletproof glass consists of two adjacent layers of thermoplastic polymer sheet over the same surface of the glass and is based on a somewhat different principle:

- The outermost layer (the side where the bullet enters) is an acrylic sheet (polymethylmethacrylate, PMMA). This sheet *dulls* the tip of the bullet—thus slowing down the bullet's speed and its ability to penetrate easily because of the now-blunt tip. In addition, the acrylic film has high weather resistance, making it suitable as the outer layer, which is exposed to the elements.
- The next layer is a polycarbonate sheet. Because it has high toughness, the polycarbonate layer stops the bullet, which has already been dulled by penetrating the acrylic sheet first. The glass shatters in the same manner as in other designs.

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 from these stresses, it is **annealed** by a process similar to the stress-relief annealing of metals. The glass is heated to a certain temperature and then cooled gradually. Depending on the size, the thickness, and the 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 smoothed by (a) grinding (as seen in glass tops for desks and shelves) or (b) holding a torch against the edges (**fire polishing**), which rounds them by localized softening of the glass and surface tension.

In all finishing operations on glass and other brittle materials, care should be exercised to ensure that there is no surface damage, especially stress raisers such as rough surface finish and scratches. Because of notch sensitivity, even a single scratch can cause premature failure of the part, 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. The potential consequences of part failure are always a significant factor in designing ceramic and glass products. The knowledge of limitations (such as poor tensile strength, sensitivity to internal and external defects, and low impact toughness) is important. On the other hand, these limitations have to be balanced against such desirable characteristics as hardness, scratch resistance, compressive strength at room and elevated temperatures, and a wide range of diverse physical properties.

The control of processing parameters and of the quality and level of impurities in the raw materials is important. As in all design decisions, there are priorities and limitations, and several factors should be considered simultaneously, including the number of parts needed and the costs of tooling, equipment, and labor.

Dimensional changes, warping, the possibility of cracking during processing, and service life are significant factors in selecting methods for shaping these materials. When a ceramic or glass component is part of a larger assembly, its compatibility with other components is another important consideration. Particularly important are the type of loading and thermal expansion (such as in seals and windows with metal frames). Recall from Table 3.1 the wide range of coefficients of thermal expansion for various metallic and nonmetallic materials. Thus, when a plate glass fits tightly in a metal window frame, temperature variations (such as the sun shining on only a portion of the window) can cause thermal stresses that may lead to cracking—a phenomenon often observed in some tall buildings. A common solution is placing rubber seals between the glass and the frame to avoid stresses due to differences in thermal expansion.

As noted in Sections 8.3.1 and 8.4.2, ceramics and glasses undergo a phenomenon called *static fatigue*, whereby they can suddenly break under a static load after a period of time. Although this phenomenon does not occur in a vacuum or in dry air, provisions must be made to avoid such failure. A general guide is that, in order for a glass item to withstand a load of 1000 hours or longer, the maximum stress that can be applied is about one-third of the maximum stress that it can withstand during the first second of loading.

18.6 Processing of Superconductors

Although *superconductors* (see 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 superconductor:

- 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.
- Ceramics, called **high-temperature superconductors** (HTSC), include various copper oxides. Here, "high" temperature 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 fundamental 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 powder, mixing it, and grinding it in a ball mill (see Fig. 17.6b) down to a grain size of 0.5 to 10 μm
2. Forming the powder into the desired shape
3. Heat treating the product.

The most common forming process is the **oxide-powder-in-tube** (OPIT) method. The powder first is packed into silver tubes (because silver has the highest electrical conductivity of any metal; see Table 3.1) and sealed at both ends. The tubes then are worked mechanically by such processes as swaging, drawing, extrusion, isostatic pressing, and rolling. The final shape may be wire, tape, coil, or bulk.

Other principal methods of processing are (a) coating silver wire with superconducting material, (b) depositing superconductor films by laser ablation, (c) forming by the doctor-blade process, (d) explosive cladding, and (e) chemical spraying. The formed 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

Significant progress has been made in recent years in understanding high-temperature superconducting materials and their potential use as electrical conductors. Two bismuth-based oxides are superconducting ceramic materials of choice for various military and commercial applications, such as electrical propulsion for ships and submarines, shallow-water and ground minesweeping systems, transmission cable generators, and superconducting magnetic energy storage (SMES). A number of 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 desirable properties. The following example illustrates this method for the production of high-temperature superconducting multifilament tapes:

1. First, a composite billet is produced using a silver casing and ceramic powder. The casing is an annealed high-purity silver that is filled with the bismuth-ceramic powder in an inert atmosphere. A steel ram is used to compact the casing in several increments to a 30% relative density. In order to minimize density gradients (such as those shown in Fig. 17.11), about 1 gram 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 deformation processing.
2. The billet is extruded and drawn to reduce its diameter and increase the powder density. Billets are drawn to a final diameter of 1.63 mm on a draw bench. It takes 12 passes with a 20.7% reduction per pass to perform the task. The dies have a semicone angle of 8°, and the drawing speed is approximately 1.4 m/min. A semisoluble oil and zinc-stearate spray are used as lubricants.
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.

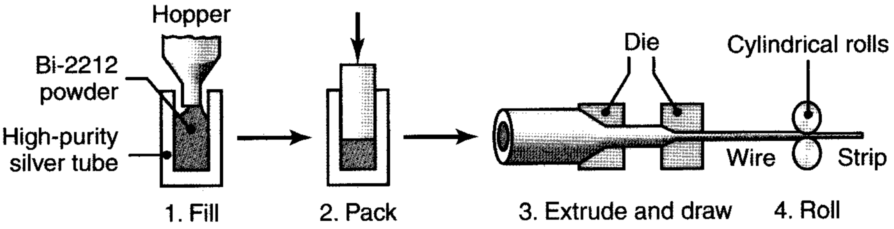


FIGURE 18.15 Schematic illustration of the powder-in-tube process.

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 inherent brittleness, ceramics are processed with due consideration of distortion and cracking. The control of raw-material quality and processing parameters also is important.

- Glass products are made by several shaping processes that are 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 general lack of tensile strength and toughness and their sensitivity to external and internal defects. Warping and cracking during production are important considerations.
- Manufacturing superconductors into useful products is challenging because of the anisotropy and inherent brittleness of the materials involved. Although new processes are being developed, the basic process consists of packing the powder into a silver tube and deforming it plastically 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

Barsoum, M.W., **Fundamentals of Ceramics**, Taylor & Francis, 2003.

Buchanan, R.C. (ed.), **Ceramic Materials for Electronics: Processing, Properties and Applications**, 3rd ed., Dekker, 2004.

Concise Encyclopedia of Advanced Ceramics, The MIT Press, 1991.

Engineered Materials Handbook, Vol. 4: **Ceramics and Glasses**, ASM International, 1991.

German, R.M., and Bose, A., **Injection Molding of Metals and Ceramics**, Metal Powder Industries Federation, 1997.

Ghosh, A., Hiremath B., and Halloran, J. (eds.), **Design for Manufacturability of Ceramic Components**, American Ceramic Society, 1995.

Harper, C.A. (ed.), **Handbook of Ceramics, Glasses, and Diamonds**, McGraw-Hill, 2001.

Hiremath, B., Gupta, T., and Nair, K.M. (eds.), **Ceramic Manufacturing Practices and Technologies.** American Ceramic Society, 1997.

Holand, W., and Beall, G.H., **Design and Properties of Glass-Ceramics**, American Chemical Society, 2001.

King, A.G., **Ceramics Processing and Technology**, Noyes Publishing, 2001.

Mutsuddy, B.C., **Ceramic Injection Molding**, Chapman & Hall, 1995.

Rahaman, M.N., **Ceramics Processing**, CRC Press, 2007.

Reed, J.S., **Principles of Ceramics Processing**, 2nd ed., Wiley, 1995.

Reimanis, I., Henager, C., and Tomsia, A. (eds.), **Ceramic Joining**, American Ceramic Society, 1997.

Richerson, D.W., **Modern Ceramic Engineering: Properties, Processing, and Use in Design**, 3rd ed., Marcel Dekker, 2005.

Sabia, R., Greenhurt, V., and Pantano, C. (eds.), **Finishing of Advanced Ceramics and Glasses**, American Ceramic Society, 1999.

Shelby, J.E., **Introduction to Glass Science and Technology**, Royal Society of Chemistry, 2005.

Somiya, S., Aldinger, F., Spriggs, R.M., Uchino, K., and Koumoto, K. (eds.), **Handbook of Advanced Ceramics: Materials, Applications, Processing and Properties.** Academic Press, 2004.

Weimer, A. (ed.), **Carbide, Nitride and Boride Materials: Synthesis and Processing**, Chapman & Hall, 1997.

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. How is glass tubing produced?

18.10. What is the difference between physical and chemical tempering of glass?

18.11. What is the structure of laminated glass? Bulletproof glass?

18.12. How are glass fibers made? What are their sizes?

QUALITATIVE PROBLEMS

18.13. Inspect various products; noting their shape, color, and transparency, identify those that are made of (a) ceramic, (b) glass, and (c) glass ceramics.

18.14. Describe the differences and similarities in processing metal powders vs. ceramics.

18.15. Which property of glasses allows them to be expanded to large dimensions by blowing? Can metals undergo such behavior? Explain.

18.16. Explain why ceramic parts may distort or warp during drying. What precautions should be taken to avoid this situation?

18.17. What properties should plastic sheets have to be used in laminated glass? Why?

18.18. 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.

18.19. Are any of the processes used for making discrete glass products similar to ones described in preceding chapters? Describe them.

18.20. Injection molding is a process that is used for powder metals, polymers, and ceramics. Explain why is this so.

18.21. Explain the phenomenon of static fatigue and how it affects the service life of a ceramic or glass component.

18.22. 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.

18.23. Is there any flash that develops in slip casting? How would you propose to remove such flash?

QUANTITATIVE PROBLEMS

18.24. Using Example 18.1, calculate (a) the porosity of the dried part if the porosity of the fired part is to be 5% and (b) the initial length, L_0, of the part if the linear shrinkages during drying and firing are 8 and 7%, respectively.

18.25. What would be the answers to Problem 18.24 if the quantities given were halved?

18.26. 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

18.27. List same 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.

18.28. Consider some ceramic products with which you are familiar, and outline a sequence of processes that you think were used to manufacture them.

18.29. 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.

18.30. How different, if any, are the design considerations for ceramics from those for other materials? Explain.

18.31. 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?

18.32. Give examples of designs and applications in which static fatigue should be taken into account.

18.33. 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.

18.34. 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.

18.35. One method of producing superconducting wire and strip is by compacting powders of these materials, placing them into a tube, and drawing them through dies or rolling them. Describe your thoughts concerning the steps and possible difficulties encountered at each stage of this process.

CHAPTER 19

Plastics and Composite Materials: Forming and Shaping

EXAMPLES:

CASE STUDY:

- This chapter describes the manufacturing processes involved in producing polymers and composite materials for consumer and industrial products.
- Extrusion is widely used to produce rods, tubing, and pellets for further processing and is also used to make base stock for the production of plastic bottles, sheet, and film.
- The chapter discusses a variety of molding operations for producing discrete parts, including injection and reaction-injection molding, transfer molding, rotational molding, and compression molding.
- The manufacturing processes associated with shaping composite materials are then described, including compression and vacuum molding, contact molding, pultrusion, and filament winding.
- The production of prepregs for these processes is also discussed.
- 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, powder metallurgy, and machining.

19.1 Introduction

The processing of plastics and elastomers involves operations similar to those used in the forming and shaping of metals, described in preceding chapters. The 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 the design and manufacture of various processing equipment to make numerous consumer and industrial products in large quantities. 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 attendant 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

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

require much less force and energy to process. Plastics in general can be molded, cast, formed, and machined into complex shapes in few operations, with relative ease, and at high production rates (Table 19.1). They also can be joined by various means (Section 32.6) and coated (generally for improved appearance) by various techniques (described in Chapter 34). Plastics are shaped into discrete products or as sheets, plates, rods, and tubing that may then be formed by secondary processes into a variety of discrete products. The types and properties of polymers and the shape and complexity of components that can be produced are influenced greatly by their method of manufacture and processing parameters.

Plastics usually are shipped to manufacturing plants as pellets, granules, or powders and are melted (for thermoplastics) just before the shaping process. Liquid plastics that cure into solid form are used especially in the making of thermosets and reinforced-plastic parts. With increasing awareness of our environment, raw materials also may consist of reground or chopped plastics obtained from recycling centers. As expected, however, product quality is not as high for such materials.

In this chapter, we follow the outline shown in Fig. 19.1, which describes the basic processes and economics of forming and shaping plastics and reinforced plastics. We also describe processing techniques for metal-matrix and ceramic-matrix composites, which have become increasingly important in various applications with critical requirements. We begin with *melt-processing* techniques (starting with extrusion) and continue on to molding processes—both categories involving the application of external pressure during processing.

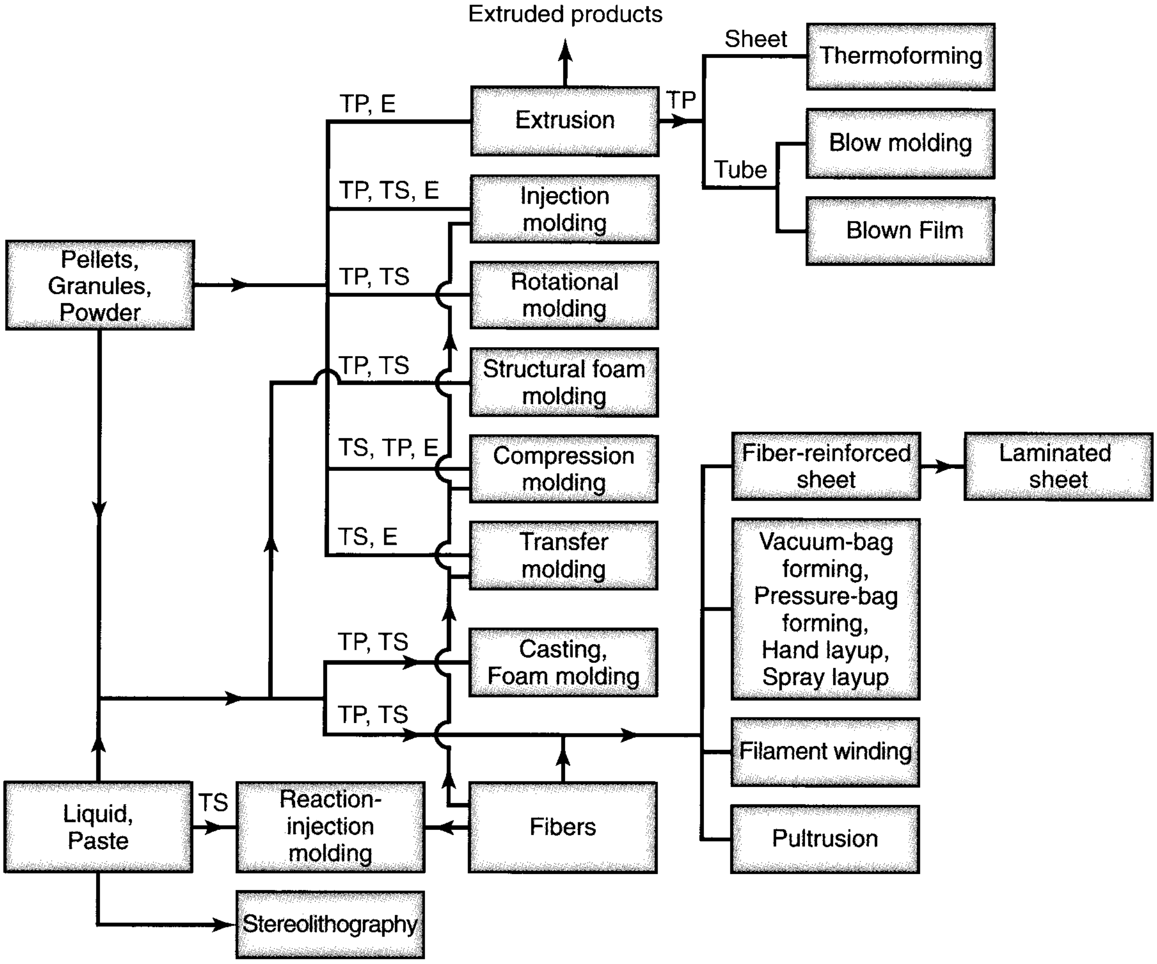


FIGURE 19.1 Outline of forming and shaping processes for plastics, elastomers, and composite materials. (TP = Thermoplastic; TS = Thermoset; E = Elastomer.)

19.2 Extrusion

In *extrusion*, which produces the largest volume of plastics, 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* (Fig. 19.2). The barrel is equipped with a helical screw that blends the pellets and conveys them down the barrel towards the die. The barrel heaters and the internal friction from the mechanical action of the screw heat the pellets and liquifies them. The screw action also builds up pressure in the barrel.

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 causes melting to begin.
3. *Metering* or *pumping section*: Where additional shearing (at a high rate) and melting occur, with pressure building up at the die.

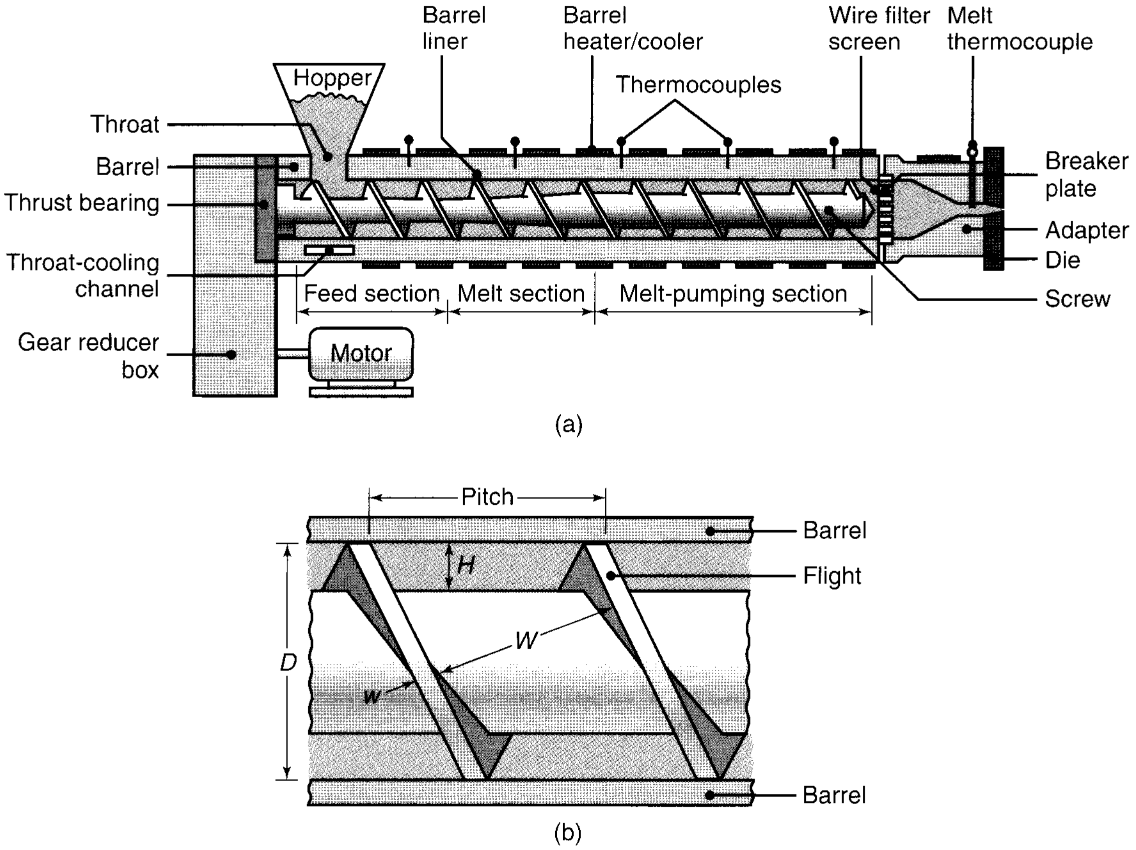


FIGURE 19.2 (a) Schematic illustration of a typical screw extruder. (b) Geometry of an extruder screw. Complex shapes can be extruded with relatively simple and inexpensive dies.

The lengths of these individual sections can be changed to accommodate the melting characteristics of different types of plastics. The molten plastic is forced through a die in a process similar to that of extruding metals. A metal-wire filter screen (Fig. 19.2a) usually is placed just before the die to filter out unmelted or congealed resin. This screen also helps build up back pressure in the barrel and is replaced periodically. Between the screen and the die is a *breaker plate*, which has several small holes in it and helps improve mixing of the polymer prior to its entering the die. The extruded product is then cooled, generally by exposing it to blowing air or by passing it through a water-filled channel (trough).

Controlling the rate and uniformity of cooling is important to minimize product shrinkage and distortion. In addition to single-screw extruders, other designs include *twin* (two parallel screws side by side) and *multiple screws* for polymers that are difficult to extrude (see also *reciprocating screw*, Section 19.3).

A typical helical screw is shown in Fig. 19.2b and indicates 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, and 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. The L/D ratio in typical commercial extruders ranges from 5 to 30, and barrel diameters are generally in the range from 25 to 200 mm.

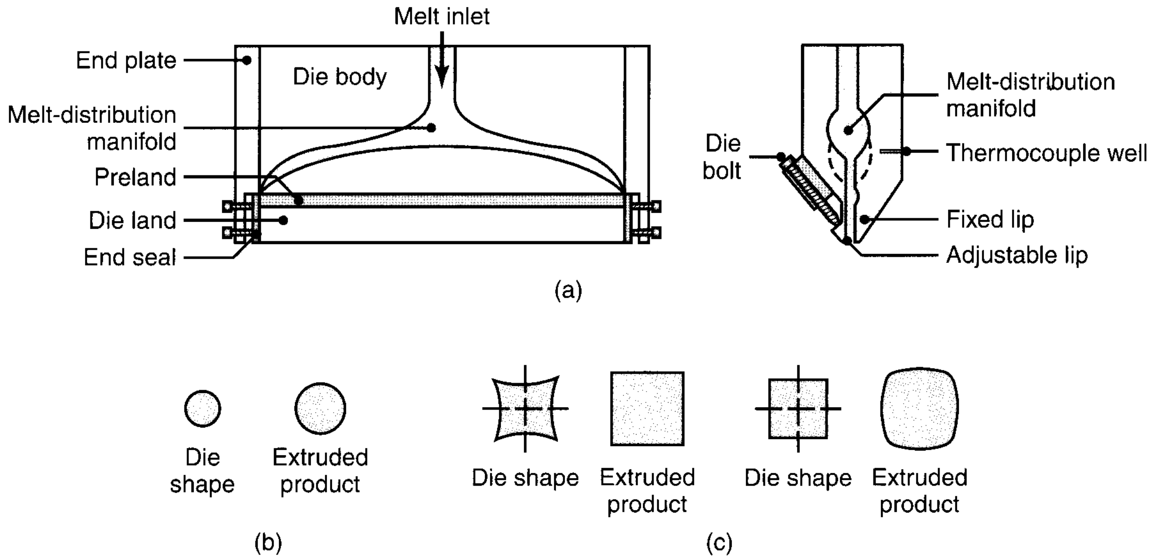


FIGURE 19.3 Common extrusion die geometries: (a) coat-hanger die for extruding sheet; (b) round die for producing rods; and (c) and (d) nonuniform recovery of the part after it exits the die. *Source*: (a) *Encyclopedia of Polymer Science and Engineering*, 2nd ed., Copyright © 1985. Reprinted by permission of John Wiley & Sons, Inc.

Because it has a direct bearing on the quality of the product extruded and on the design of the extruder and the die, the mechanics of this operation has been studied extensively. Several relationships have been established between the dimensions shown in Fig. 19.2b, the screw rotational speed, and the viscosity of the polymer to describe what are known as *extruder characteristics* and *die characteristics*. These characteristics then determine such quantities as the pressure and flow rate at any location during the operation of the extruder. (Further details are beyond the scope of this book and are available in various references cited in the bibliography at the end of this chapter.)

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 section can be extruded with relatively inexpensive tooling. Some common die profiles are shown in Fig. 19.3b. Note that some of the profiles are not intuitive, but this is attributable to the polymer usually undergoing much greater and uneven shape recovery than is encountered in metal extrusion. Furthermore, since the polymer will *swell* at the exit of 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 and coiled or cut into desired lengths.

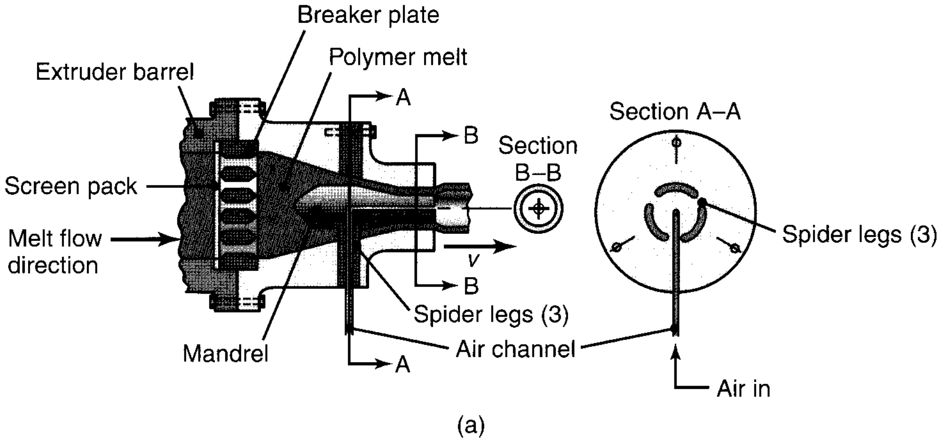
The control of processing parameters such as extruder-screw rotational speed, barrel-wall temperatures, die design, and rate of cooling and drawing speeds are important in order to ensure product integrity and uniform dimensional accuracy. *Defects* observed in extruding plastics are similar to those observed in metal extrusion (described in 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 in metals). Other surface defects are *bambooing* and *sharkskin effects*—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. Machinery costs can be 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 products.

Plastic Tubes and Pipes. These are produced in an extruder with a *spider die*, as shown in Fig. 19.4a (see also Fig. 15.8 for details). Woven fiber or wire reinforcements also may be fed through specially designed dies in this operation for the production of reinforced hoses that need to withstand higher pressures. The extrusion of tubes is also a necessary first step for related processes, such as extrusion blow molding and blown film.



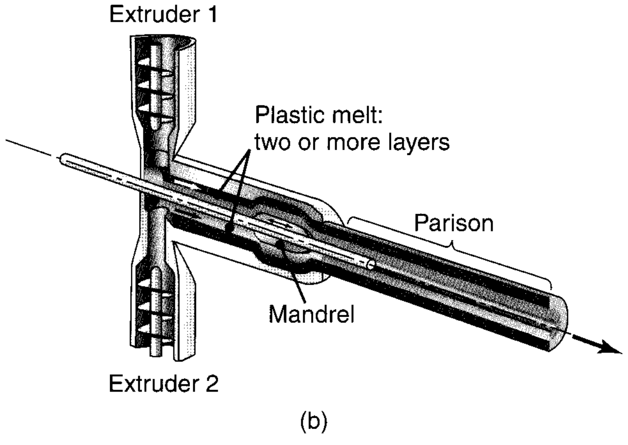


FIGURE 19.4 Extrusion of tubes. (a) Extrusion using a spider die (see also Fig. 15.8) and pressurized air. (b) Coextrusion for producing a bottle.

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 tubes.

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 function. Coextrusion commonly is performed in shapes such as flat sheets, films, and tubes, and is used especially in food packaging where different layers of polymers have different functions, such as (a) providing inertness for food, (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 also are extruded and coated with plastic by this process. The wire is fed into the die opening at a controlled rate with the extruded plastic in order to produce a uniform coating. *Plastic-coated wire paper clips* also are made by this process. To ensure proper insulation, extruded electrical wires are checked continuously for their resistance as they exit the die; they also are marked automatically with a roller to identify the specific type of wire.

Polymer Sheets and Films. These can be produced by using a specially designed flat-extrusion die, such as that shown in Fig. 19.3a. Also known as the *coat-hanger die*, it is designed to distribute the polymer melt evenly throughout the width. The polymer is extruded by forcing it through the die, after which the extruded sheet is taken up—first on water-cooled rolls and then by a pair of rubber-covered pull-off rolls. Generally, polymer *sheet* is considered to be thicker than 0.5 mm, and *film* is thinner than 0.5 mm.

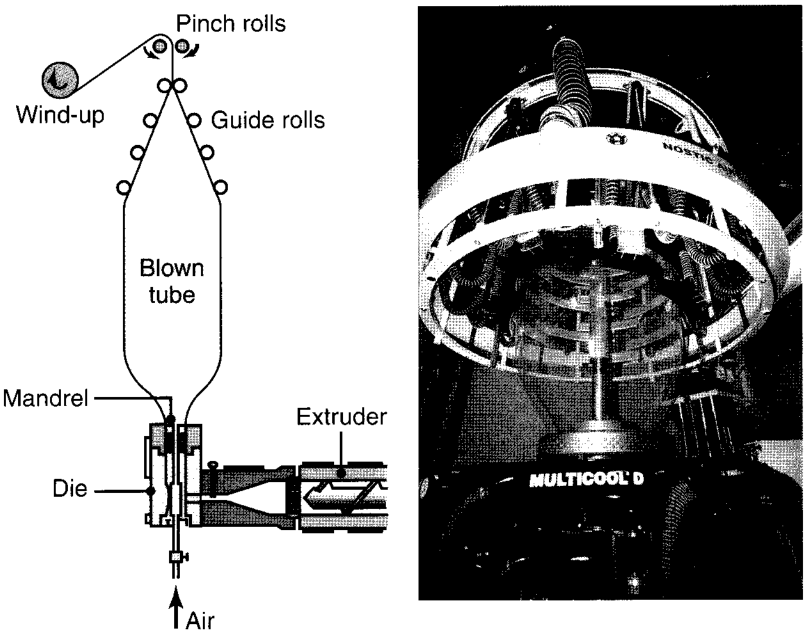
Thin Polymer Films. Common *plastic bags* and other thin polymer film products are made from **blown film**, which in turn is made from a thin-walled tube produced by an extruder (Fig. 19.5). In this process, a tube is extruded continuously vertically upward and then expanded into a balloon shape by blowing air through the center of the extrusion die until the desired film thickness is reached. Because of the molecular orientation of thermoplastics (Section 7.3), a *frost line* develops on the balloon and its transparency is reduced.

The balloon usually is cooled by air from a cooling ring around it, which can also act as a barrier to further expansion of the balloon, thus controlling its dimensions. The cooled bubble is then *slit* lengthwise, becoming *wrapping film*, or it is pinched and cut off, becoming a *plastic bag*. The width of the film produced after slitting can be on the order of 6 m or more.

The ratio of the blown diameter to the extruded tube diameter is known as the *blow ratio*, which is about 3:1 in Fig. 19.5. Note that, as described in Section 2.2.7, the polymer has to have a high strain-rate sensitivity exponent, m, to successfully be blown by this process without tearing.

Plastic Films. Plastic films, especially polytetrafluoroethylene (PTFE; trade name: *Teflon*), can be produced by *shaving* the circumference of a solid, round plastic billet with specially designed knives in a manner similar to producing veneer from a large piece of round wood. The process is called **skiving** (see also Section 24.4).

Pellets. Used as raw material for other plastic-processing methods described in this chapter, pellets also are made by extrusion. A small-diameter, solid rod is extruded continuously and then *chopped* into short lengths (*pellets*). With some



(a) (b)

FIGURE 19.5 (a) Schematic illustration of the production of thin film and plastic bags from tube—first produced by an extruder and then blown by air. (b) A blown-film operation. This process is well developed, producing inexpensive and very large quantities of plastic film and shopping bags. *Source*: (b) Courtesy of Wind Moeller & Hoelscher Corp.

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

Assume that a typical plastic shopping bag made by blown film has a lateral dimension (width) of 400 mm. (a) What should be 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 2.5 times the extrusion-die diameter. Taking the maximum value of 2.5, we calculate the die diameter as $255/2.5 = 100$ mm.

b. Note in Fig. 19.5a that, after extrusion, the balloon is being pulled upward by the pinch rolls. Thus, in addition to diametral stretching 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, these fibers are used in a wide variety of consumer and industrial products, including clothing, carpeting, fabrics, rope, and packaging.

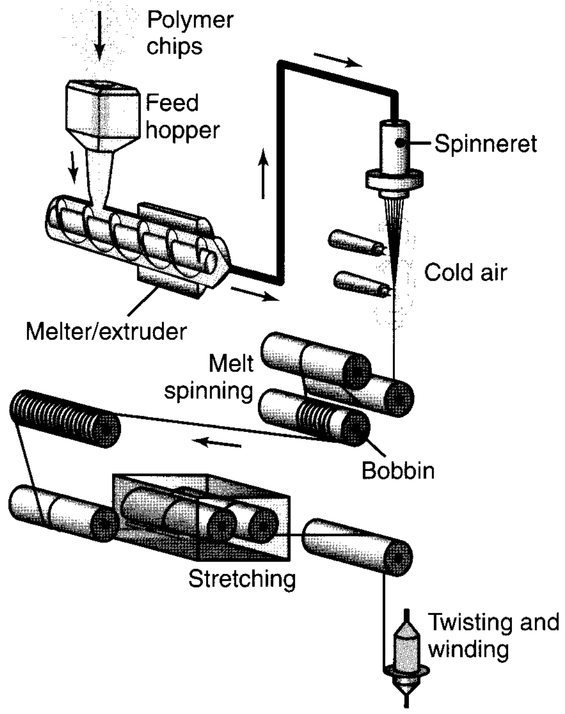


FIGURE 19.6 The melt-spinning process for producing polymer fibers. The fibers are used in a variety of applications, including fabrics and as reinforcements for composite materials. In the stretching box the right roll rotates faster than the left roll.

Most synthetic fibers used in reinforced plastics are polymers that are extruded through the tiny holes of a device called a **spinneret** (resembling a showerhead) to form continuous filaments of semisolid polymer. The extruder forces the polymer through the spinneret, which may have from one to several hundred holes. If the polymers are thermoplastics, they first are melted in the extruder, as described in 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 production rates and with very high reliability.

As the filaments emerge from the holes in the spinneret, the liquid polymer is first converted to a rubbery state and then solidified. This process of extrusion and solidification of continuous filaments is called **spinning**. The term *spinning* also is used for the production of natural textiles (such as cotton or wool), where short pieces of fiber are twisted into yarn. There are four methods of spinning fibers: melt, wet, dry, and gel spinning.

1. **In melt spinning** (shown in Fig. 19.6), the polymer is melted for extrusion through the spinneret and then solidified directly by cooling. A typical spinneret for this operation has about 50 holes around 0.25 mm in diameter and is about 5 mm thick. The fibers that emerge from the spinneret are cooled by forced-air convection and are simultaneously pulled, so that their final diameter becomes much smaller than the spinneret opening. Polymers (such as nylon, olefin, polyester, and PVC) are produced in this matter. Because of the important applications of nylon and polyester fibers, melt spinning is the most important fiber-manufacturing process.

   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 and thus provide additional thermal insulation, while other cross sections have specific applications.
2. **Wet spinning** is the oldest process for fiber production and is used for polymers that have been dissolved in a solvent. The spinnerets are submerged in a chemical bath. As the filaments emerge, they precipitate in the bath, producing a fiber that is then wound onto a bobbin. The term "wet spinning" refers to the use of a precipitating liquid bath, resulting in wet fibers that require drying before they can be used. Acrylic, rayon, and aramid fibers can be produced by this process.
3. **Dry spinning** is used for thermosets carried by a solvent. However, instead of precipitating the polymer by dilution as in wet spinning, solidification is achieved by evaporating the solvent in a stream of air or inert gas. The filaments do not come in contact with a precipitating liquid, thus eliminating the need for drying. Dry spinning may be used for the production of acetate, triacetate, polyether-based elastane, and acrylic fibers.
4. **Gel spinning** is a special process used to obtain high strength or special fiber properties. The polymer is not melted completely or dissolved in liquid, but the molecules are bound together at various points in liquid-crystal form. This operation produces strong interchained forces in the resulting filaments

that can significantly increase the tensile strength of the fibers. In addition, the liquid crystals are aligned along the fiber axis by the strain encountered during extrusion. The filaments emerge from the spinneret with an unusually high degree of orientation relative to each other—further enhancing their strength. This process also is called *dry–wet spinning*, because the filaments first pass through air and then are cooled further in a liquid bath. Some high-strength polyethylene and aramid fibers are produced by gel spinning.

A necessary step in the production of most fibers is the application of significant *stretching* to induce orientation of the polymer molecules in the fiber direction. This orientation is the main reason for the high strength of the fibers, compared with the polymer in bulk form. The stretching can be done while the polymer is still pliable—just after extrusion from the spinneret—or it can be performed as a cold-drawing operation. The strain induced can be as high as 800%.

Graphite fibers are produced from different polymer fibers by *pyrolysis*. In this operation, controlled heat in the range from 1500° to 3000°C is applied to the polymer fiber (typically polyacrylonitrile, PAN) to drive off all elements except the carbon. The fiber is under tension in order to develop 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.)

19.3 Injection Molding

Injection molding is similar to hot-chamber die casting (Fig. 19.7; see also Section 11.3.5). The pellets or granules are fed into the heated cylinder, and the melt

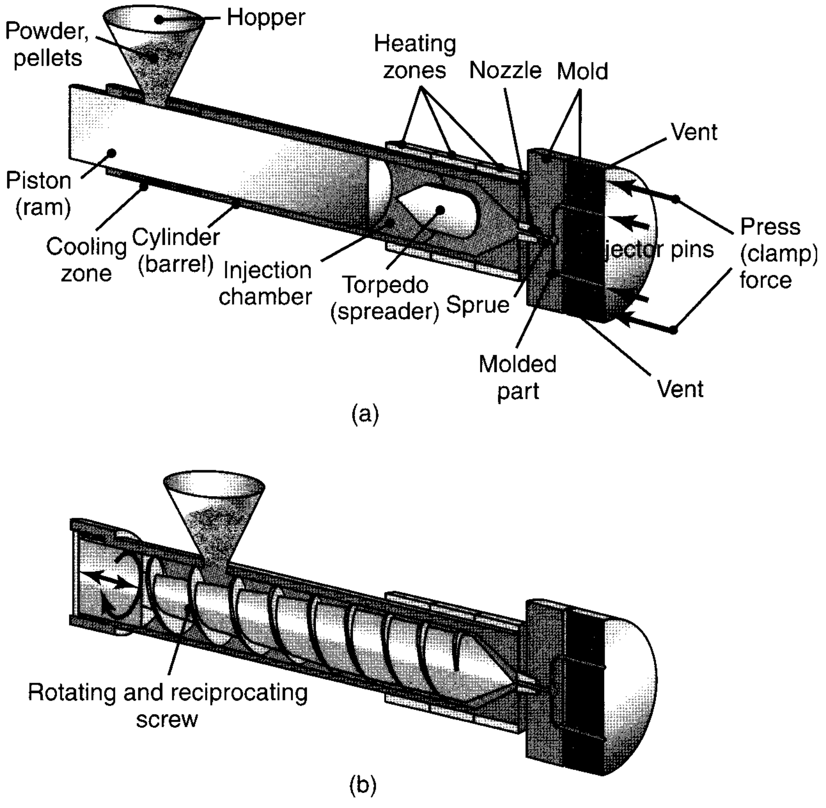
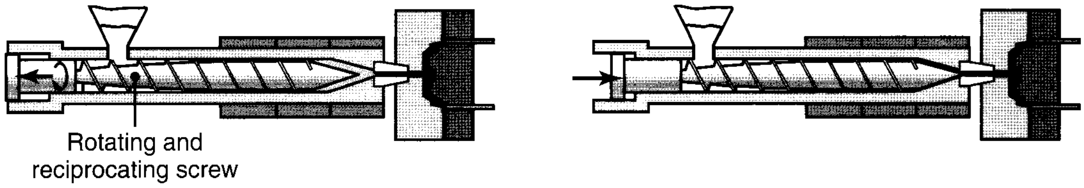
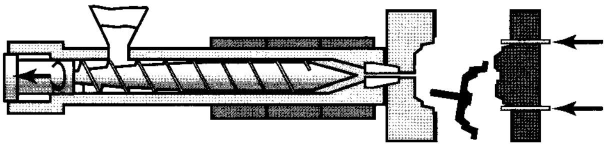


FIGURE 19.7 Schematic illustration of injection molding with (a) a plunger and (b) a reciprocating rotating screw.



1. Build up polymer in front of sprue bushing; pressure pushes the screw backwards. When sufficient polymer has built up, rotation stops.

2. When the mold is ready, the screw is pushed forward by a hydraulic cylinder, filling the sprue bushing, sprue, and mold cavity with polymer. The screw begins rotating again to build up more polymer.



3. After polymer is solidified/cured, the mold opens, and ejector pins remove the molded part.

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).

is forced into the mold either by a hydraulic *plunger* or by the *rotating screw* system 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.

Modern machines are of the *reciprocating* or *plasticating screw type* (Fig. 19.7b), with 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 usually range from 70 to 200 MPa.

Some injection-molded products are shown in Fig. 19.9. Other products include cups, containers, housings, tool handles, knobs, toys, plumbing fixtures, telephone receivers, and electrical and communication-equipment components. 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.

After the part has cooled sufficiently (for thermoplastics) or cured (for thermosets), the molds are opened and ejectors are used to remove the part from the mold. The molds are then closed, and the process is repeated automatically. Elastomers also are injection molded into discrete products by these processes. Because the material is molten when injected into the mold, complex shapes with good dimensional accuracy can be obtained. However, because of uneven cooling of the part in the mold, residual stresses develop.

Molds with moving and unscrewing mandrels also are used in injection molding, as they allow the molding of parts having 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

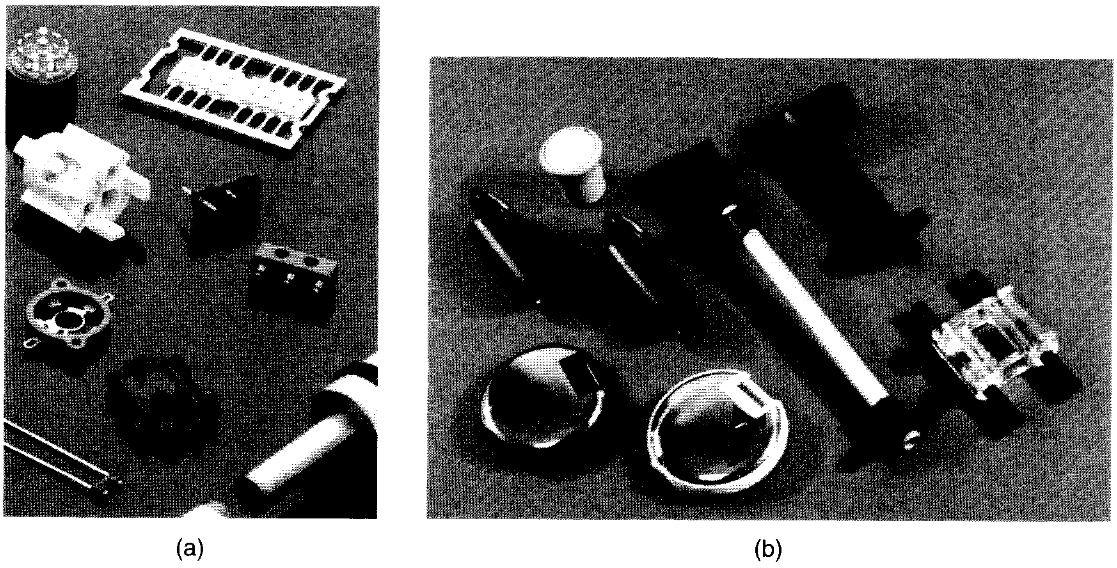


FIGURE 19.9 Typical products made by injection molding, including examples of insert molding. *Source*: (a) Courtesy of Plainfield Molding, Inc. (b) Courtesy of Rayco Mold and Mfg. LLC.

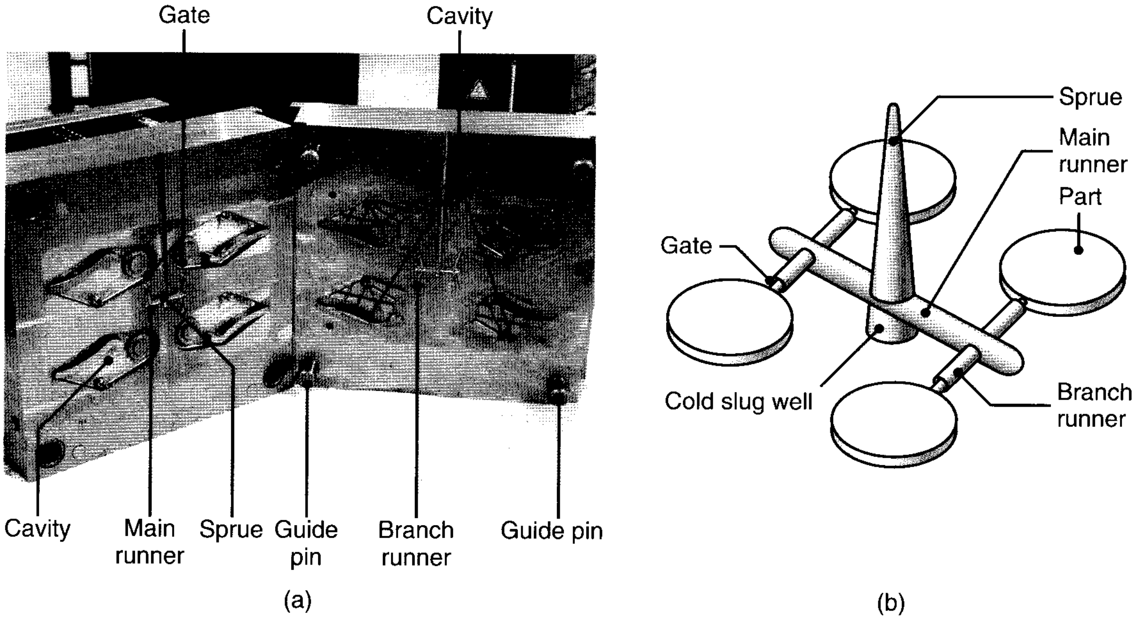


FIGURE 19.10 Illustration of mold features for injection molding. (a) Two-plate mold with important features identified. (b) Schematic illustration of the features in a mold. *Source*: Courtesy of Tooling Molds West, Inc.

dies), cores, cavities, cooling channels, inserts, knockout pins, and ejectors. There are three basic types of molds:

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.

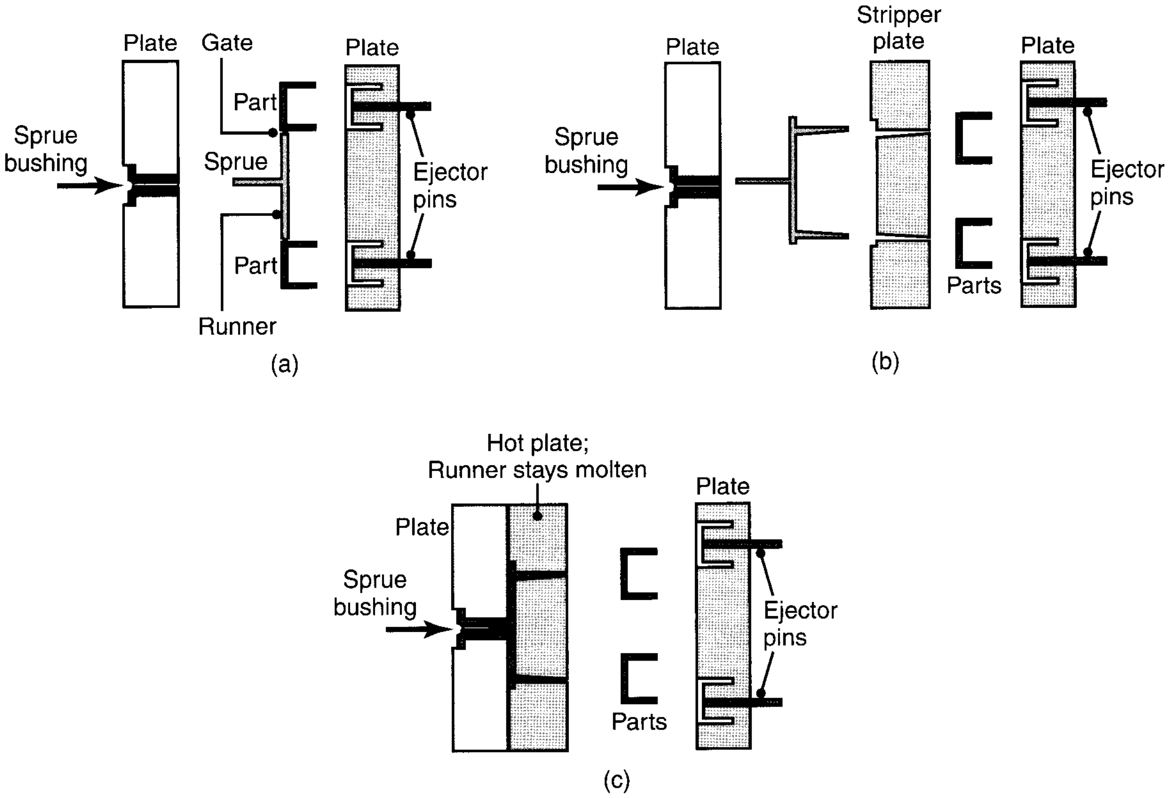


FIGURE 19.11 Types of molds used in injection molding: (a) two-plate mold; (b) three-plate mold; and (c) hot-runner mold.

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 *coinjection* or *sandwich molding*) allows the forming of parts with a combination of various colors and shapes. An example is the molding of automobile rear-light covers made of different materials and colors, such as red, amber, and white. Also, for some parts, printed film can be placed in the mold cavity, so they need not be decorated or labeled after molding.

Insert molding involves metallic components (such as screws, pins, and strips) that are placed in the mold cavity prior to injection and then 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 and faucet parts.

Overmolding. This is a process for making products (such as hinge joints and ball-and-socket joints) in one operation and without the need for postmolding assembly. Two different plastics usually are used to ensure that no bonds will form between the molded halves of the joint, as otherwise motion would be impeded.

In **ice-cold molding**, the same type of plastic is used to form both components of the joint. The operation is carried out in a standard 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 the second molded component 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.

Process Capabilities. Injection molding is a high-rate production process and permits good dimensional control. Although most parts generally weigh from 100 to 600 g, they can be much heavier, such as automotive-body panels and exterior components. Typical cycle times range from 5 to 60 seconds, although they can be several minutes for thermosetting materials.

Injection molding is a versatile process capable of producing complex shapes with good dimensional accuracy. As in other forming processes, mold design and the control of material flow in the die cavities are important factors in the quality of the product and thus in avoiding defects. Because of the basic similarities to metal casting regarding material flow and heat transfer, *defects* observed in injection molding are somewhat of the same nature, as outlined next.

- For example, in Fig. 10.13g, 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 full 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 due to die wear, a *flash* will form in a manner similar to flash formation in impression-die forging (see Figs. 14.5 and 19.17c).
- A defect known as *sink marks* (or pull-in) similar to that shown in Fig. 19.31c also is observed in injection-molded parts.
- Methods of avoiding defects consist of the proper control of temperatures, pressures, and mold design modifications using simulation software.

Much progress has been made in the analysis and design of molds and material flow in injection molding. *Modeling techniques* and *simulation software* have been developed for studying optimum gating systems, mold filling, mold cooling, and part distortion. *Software programs* now are available to expedite the design process for molding parts with good dimensions and characteristics. The programs take into account such factors as injection pressure, temperature, heat transfer, and the condition of the resin.

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 generally is supplied by hydraulic means, although electrical means (which weigh less and are quieter than hydraulic machines) also are used. Modern machines are equipped with microprocessors in a control panel and monitor all aspects of the 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. The largest machine in operation has a capacity of 45 MN, and it can produce parts weighing 25 kg. The cost of a 1-MN machine ranges from about $60,000 to about $90,000 and of a 2.7-MN machine from about $85,000 to about $140,000. Die

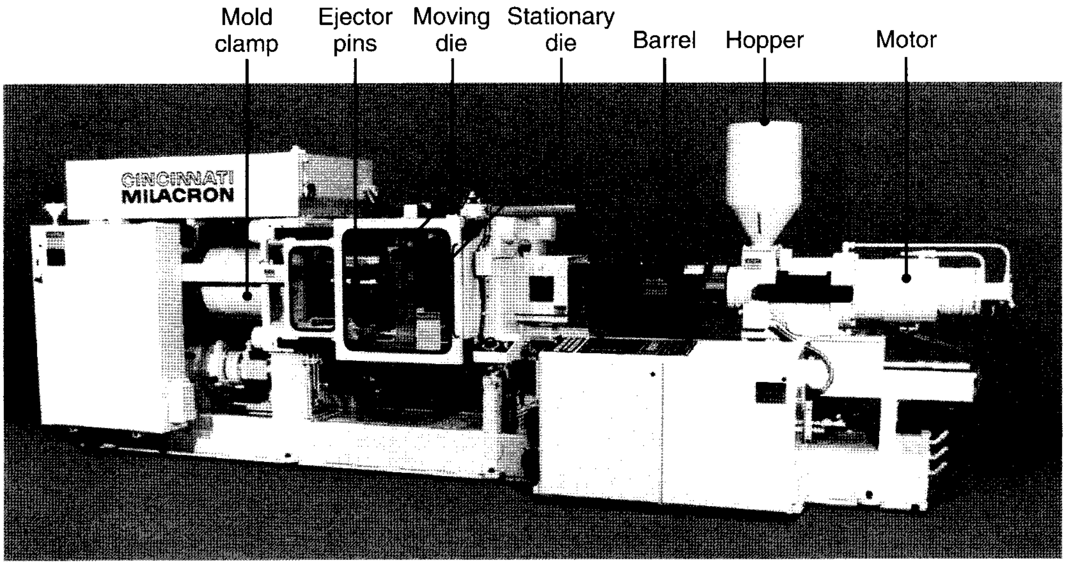


FIGURE 19.12 A 2.2-MN injection-molding machine. The tonnage is the force applied to keep the dies closed during the injection of molten plastic into the mold cavities and hold it there until the parts are cool and stiff enough to be removed from the die.
Source: Courtesy of Cincinnati Milacron, Plastics Machinery Division.

costs typically range from \$20,000 to \$200,000. Consequently, high-volume production is essential to justify such high expenditure.

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 made in one cycle (see also Fig. 11.21). Mold costs can be on the order of \$100,000 for large ones. Mold life may be on the order of 2 million cycles for steel molds, but it can be about only 10,000 cycles for aluminum molds.

EXAMPLE 19.2 Force Required in Injection Molding

A 2.2-MNn injection-molding machine is to be used to make spur gears 110 mm in diameter and 2.5 mm thick. The gears have a fine-tooth profile. How many gears can be injection molded in one set of molds? Does the thickness of the gears influence your answer?

Solution Because of the fine detail involved (fine gear teeth), let's assume that the pressures required in the mold cavity will be on the order of 100 MPa. The cross-sectional (projected) area of the gear is $\pi(110)^2/4 = 9500\text{mm}^2$. If we assume that the parting plane of the two halves of the mold is in the middle of the gear, the force required is $(9500)(100) = 0.95$ MN.

Since the capacity of the machine is 2.2 MN, we have 22 MN of clamping force available. Hence, the mold can accommodate two cavities and produce two gears per cycle. Because it does not influence the cross-sectional area of the gear, the thickness of the gear does not directly influence 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 and two or more reactive fluids are forced at high speed into a mixing chamber at a pressure of 10 to 20 MPa and then into the mold cavity (Fig. 19.13). Chemical reactions take place rapidly in the mold, and the polymer solidifies. Typical polymers are polyurethane,

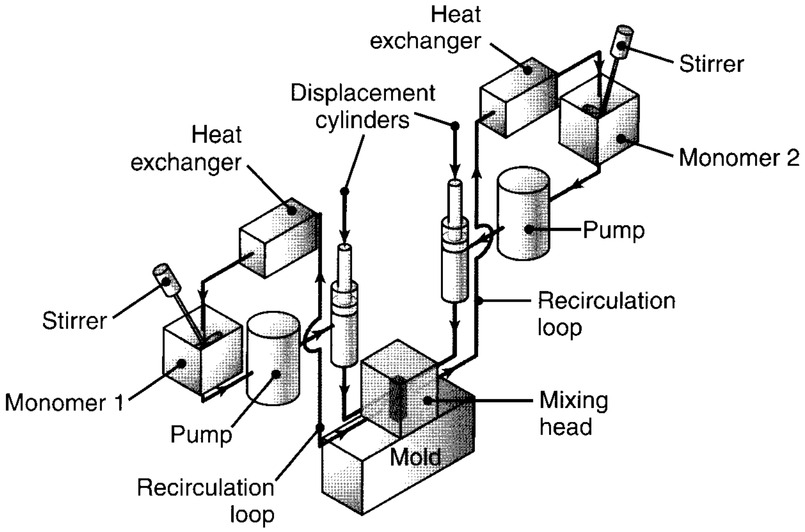


FIGURE 19.13 Schematic illustration of the reaction-injection molding process. Typical parts made are automotive-body panels, water skis, and thermal insulation for refrigerators and freezers.

nylon, and epoxy. Cycle times may range up to about 10 minutes, depending on the materials, part size, and shape.

Major applications of this process include automotive parts (such as bumpers and fenders, steering wheels, and instrument panels), thermal insulation for refrigerators and freezers, water skis, and stiffeners for structural components. Parts made may range up to about 50 kg. Reinforcing fibers (such as glass or graphite) also may be used to improve the product's strength and stiffness. Depending on the number of parts to be made and the part quality required, molds can be made of common materials, such 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 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. 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 some operations, the extrusion is continuous and the molds move with the tubing. The molds close around the tubing, sealing off one end, breaking the long tube into individual sections, and moving away as air is injected into the tubular piece. The part is then cooled and ejected from the mold. Corrugated-plastic pipe and tubing are made by continuous blow molding in which the pipe or tubing is extruded horizontally and blown into moving molds.

In **injection blow molding**, a short tubular piece (**parison**) is injection molded (Fig. 19.14b) into cool dies. (Parisons may be made and stored for later use.) The

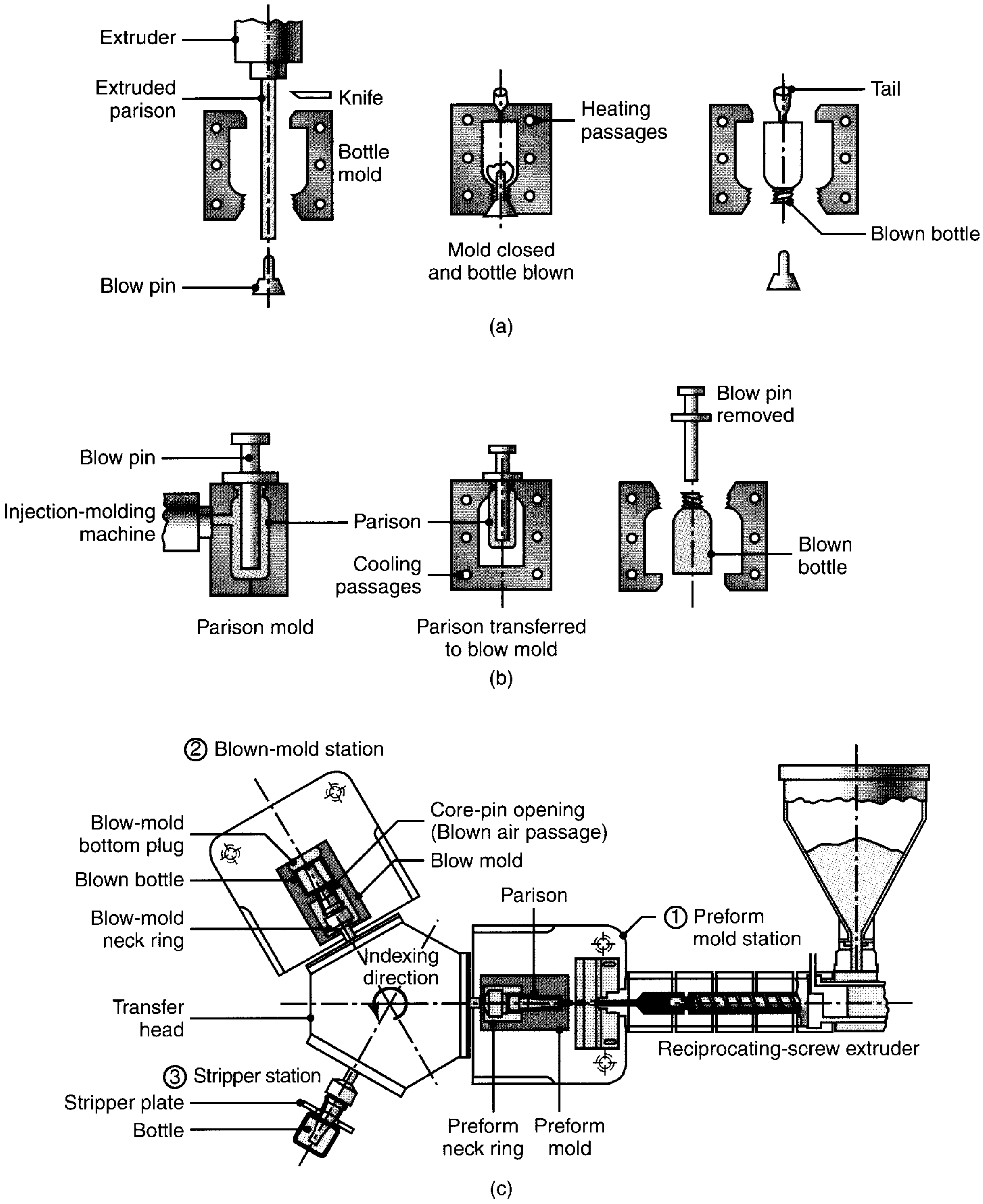


FIGURE 19.14 Schematic illustrations of (a) the extrusion blow-molding process for making plastic beverage bottles; (b) the injection blow-molding process; and (c) a three-station injection blow-molding machine for making plastic bottles.

dies then open, and the parison is transferred to a blow-molding die by an indexing mechanism (Fig. 19.14c). Hot air is injected into the parison, expanding it to the walls of the mold cavity. Typical products made are plastic beverage bottles (typically made of polyethylene or polyetheretherketone, PEEK) and small, hollow containers. A related process is **stretch blow molding**, in which the parison is expanded and elongated simultaneously, subjecting the polymer to biaxial stretching and thus enhancing its properties.

Multilayer blow molding involves the use of coextruded tubes or parisons and thus permits the production of a multilayer structure (see Fig. 19.4b). A typical example of such a product is plastic packaging for food and beverages, having such characteristics as odor and permeation barrier, taste and aroma protection, scuff resistance, the capability of being printed, and the ability to be filled with hot fluids. Other applications of this process are for containers in the cosmetics and the pharmaceutical industries.

19.5 Rotational Molding

Most thermoplastics and some thermosets can be formed into large, hollow parts by *rotational molding*. In this process, a thin-walled metal mold is made in two pieces (split-female mold) and is designed to be rotated about two perpendicular axes (Fig. 19.15). For each part cycle, a premeasured quantity of powdered plastic material is placed inside the warm mold. (The powder is obtained from a polymerization process that precipitates a powder from a liquid.) Then the mold is heated (usually in a large oven) and is rotated continuously about the two principal axes.

This action tumbles the powder 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 large variety of parts are made by rotational molding, such as storage tanks of various sizes, trash cans, boat hulls, buckets, housings, large hollow toys, carrying cases, and footballs. Various metallic or plastic inserts or components also may be molded integrally into the parts made by this process.

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 and toys.

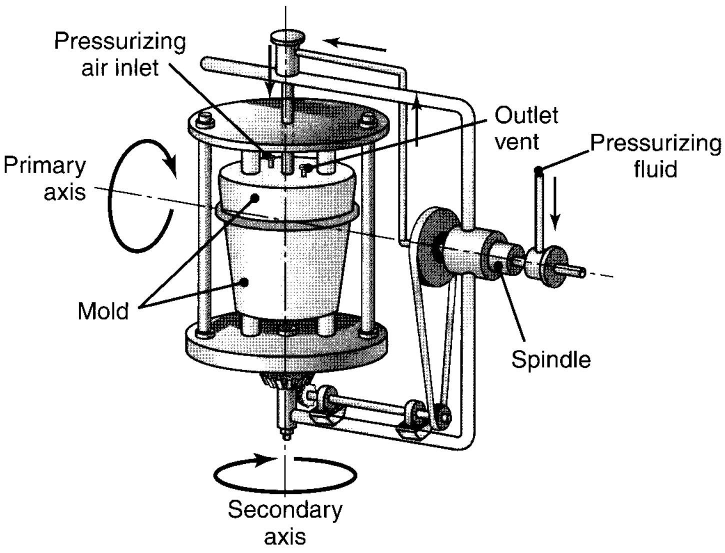


FIGURE 19.15 The rotational molding (rotomolding or rotocasting) process. Trash cans, buckets, and plastic footballs can be made by this process.

Process Capabilities. Rotational molding can produce parts with complex, hollow shapes with wall thicknesses as small as

0.4 mm. Parts as large as 1.8 m × 1.8 m × 3.6 m with a volume as large as 80,000 liters have been produced. The outer surface finish of the part is a replica of the surface finish of the inside mold walls. 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.

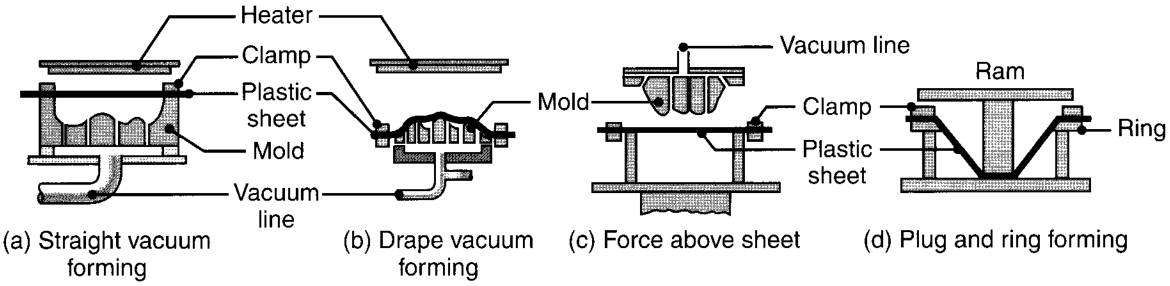
19.6 Thermoforming

Thermoforming is a process for forming thermoplastic sheets or films over a mold through the application of heat and pressure (Fig. 19.16). In this process, a sheet is (a) 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 (b) forced against the mold surfaces through the application of a vacuum or air pressure. The sheets used in thermoforming are available as a coiled strip or as lengths and widths of various sizes. They also are available filled with various materials for making parts with specific applications.

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 formed, the pressure difference caused by a vacuum usually is sufficient for forming. However, thicker and more complex parts require air pressure, which may range from about 100 to 2000 kPa, depending on the type of material and thickness of the sheet. Mechanical means, such as the use of *plugs*, also may be employed to help form the parts. 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 a combination of *drawing* and *stretching* operations (much like in some sheet-metal forming), the material must exhibit high, uniform elongation; otherwise, it will neck and tear. Thermoplastics have high capacities for uniform elongation by virtue of their high strain-rate sensitivity exponent, m, as described in Section 2.2.7.

Molds for thermoforming usually are made of aluminum because high strength is not required; hence, tooling is relatively inexpensive. Thermoforming



(a) Straight vacuum forming (b) Drape vacuum forming (c) Force above sheet (d) Plug and ring forming

FIGURE 19.16 Various thermoforming processes for a thermoplastic sheet. These processes commonly are used in making advertising signs, cookie and candy trays, panels for shower stalls, and packaging.

molds have small through-holes in order to aid vacuum forming. These holes typically are less than 0.5 mm in diameter; otherwise, they would leave marks on the parts formed. Defects encountered in thermoforming include (a) tearing of the sheet during forming, (b) 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 material, premeasured volume of powder, or viscous mixture of liquid-resin and filler material is placed directly into a heated mold cavity that 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.

Pressures range from about 10 to 150 MPa. As seen in Fig. 19.17, there is a flash formed, which subsequently is removed by trimming or by some other means. Typical parts made are dishes, handles, container caps, fittings, electrical and electronic components, washing-machine agitators, and housings. Fiber-reinforced parts with chopped fibers also are formed exclusively by this process.

Compression molding is used mainly with thermosetting plastics, with the original material being in a partially polymerized state. However, thermoplastics and elastomers are also processed by compression molding. Curing times range from about 0.5 to 5 minutes, depending on the material and on part thickness and shape. The thicker the material, the longer it will take to cure.

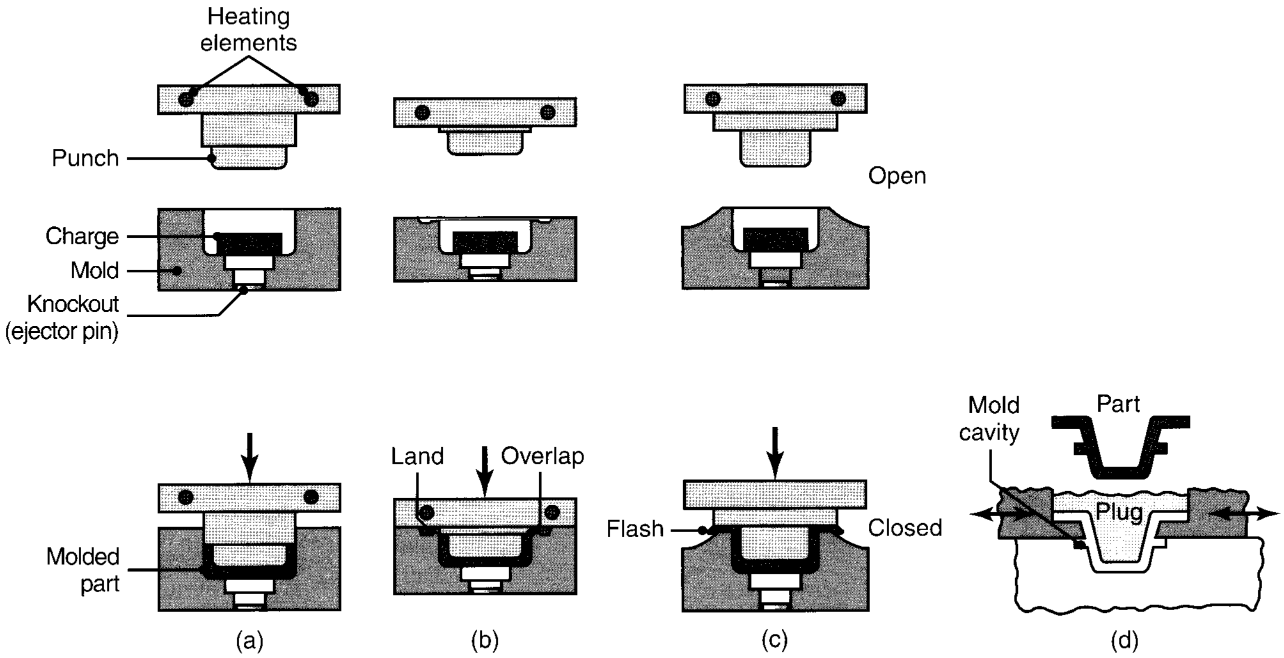


FIGURE 19.17 Types of compression molding—a process similar to forging: (a) positive, (b) semipositive, and (c) flash, in which the flash is later trimmed off. (d) Die design for making a compression-molded part with external undercuts.

Process Capabilities. Three types of compression molds are available:

- *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 used in injection molding. They typically are made of tool steels and may be chrome plated or polished for an improved surface finish of the molded product.

19.8 Transfer Molding

Transfer molding represents a further development of compression molding. The uncured thermosetting resin is placed in a heated transfer pot or chamber (Fig. 19.18), and after the material is heated, it is injected into heated closed molds. Depending on the type of machine used, a ram, plunger, or rotating-screw feeder forces the material to flow through the narrow channels into the mold cavity at pressures up to 300 MPa. This viscous flow generates considerable heat, which raises the temperature of the material and homogenizes it. Curing takes place by cross-linking. Because the resin is in a molten state as it enters the molds, the complexity of the parts and the dimensional control approach those of injection molding.

Process Capabilities. Typical parts made by transfer molding are electrical connectors and electronic components, rubber and silicone parts, and the encapsulation of microelectronic devices. The process is suitable particularly for intricate shapes with varying wall thicknesses. The molds tend to be more expensive than those for

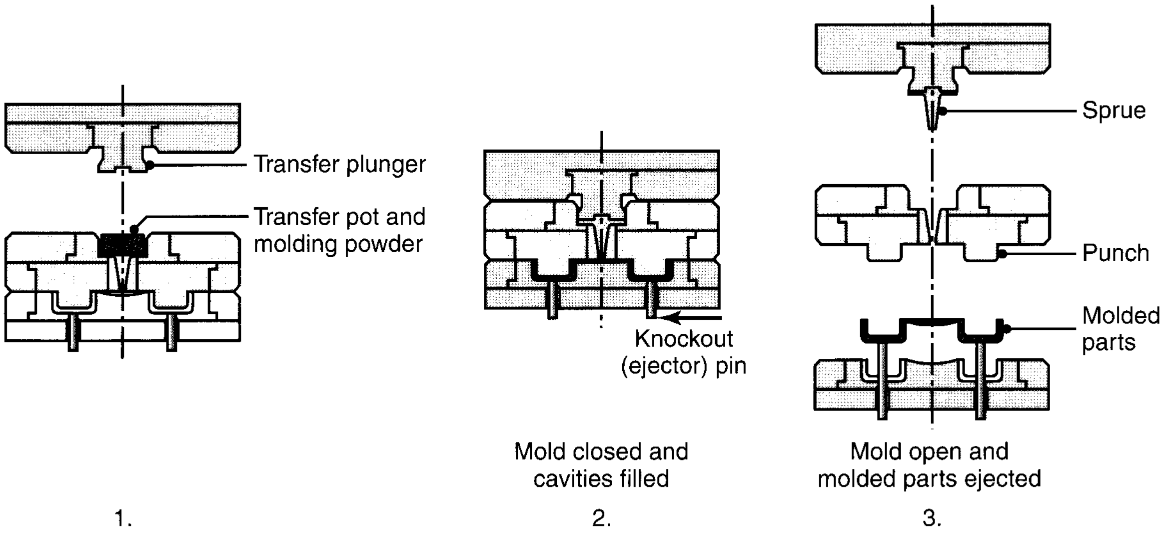


FIGURE 19.18 Sequence of operations in transfer molding for thermosetting plastics. This process is suitable particularly for intricate parts with varying wall thickness.

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 (epoxies, phenolics, polyurethanes, and polyester) can be *cast* into a variety of shapes using either rigid or flexible molds (Fig. 19.19). Compared with other methods of processing plastics, casting is a slow, but simple and inexpensive, process. However, 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.

In the basic 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 formed after polymerization takes place at ambient pressure. Degassing may be necessary for product integrity. Intricate shapes can be produced using *flexible molds*, which are then peeled off (in a manner similar to using rubber gloves) and reused. As with metals, thermoplastics may be *cast continuously*, with the polymer carried over continuous stainless-steel belts and polymerized by external heat.

Centrifugal Casting. This process, similar to centrifugal metal casting (Section 11.3.6), is used with thermoplastics, thermosets, and reinforced plastics with short fibers.

Potting and Encapsulation. As a variation of casting that is important, particularly to the electrical and electronics industry, *potting* and *encapsulation* involve casting the plastic material (typically a liquid resin, such as expoxy) 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 and fixes it in position. 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, studs, and similar parts) may be encapsulated partially by dipping them in a hot thermoplastic using polymers of various colors.

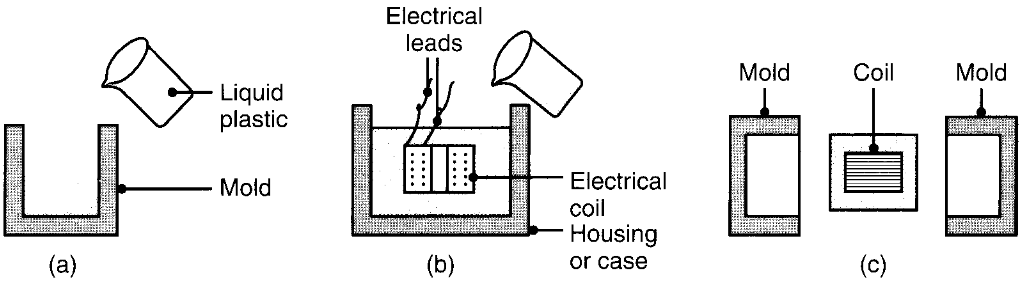


FIGURE 19.19 Schematic illustration of (a) casting, (b) potting, and (c) encapsulation processes for plastics and electrical assemblies, where the surrounding plastic serves as a dielectric.

19.10 Foam Molding

Products such as styrofoam cups, food containers, 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. As is readily seen upon close inspection, these products have a **cellular structure**. The structure 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 can be 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 exposed to heat, usually by steam. As a result, 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 (which can then be cut to size). The bead size selected also depends on the minimum wall thickness of the product: The smaller the size, the thinner the part. The beads can be colored prior to expansion; thus, the part becomes integrally colored. Both thermoplastics and thermosets can be used for foam molding, but thermosets are in a liquid-processing form and hence are in a condition similar to that of polymers in reaction-injection molding.

A common method of foam molding is to use *preexpanded polystyrene beads*, in which the beads are expanded partially by steam (hot air, hot water, or an oven also can be used) in an open-top chamber. The beads then are 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 in the manner described previously.

Structural Foam Molding. This is a molding process used to make plastic products with a *solid outer skin* and a *cellular core structure*. Typical products made are furniture components, computer and business-machine housings, and moldings (replacing more expensive wood moldings). In this process, thermoplastics are mixed with a blowing agent (usually an inert gas such as nitrogen) and injection molded into cold molds of desired shapes. The rapid cooling against the cold-mold surfaces produces a skin that is rigid [which can be as much as 2 mm thick] and a core of the part that is cellular in structure. The overall part density can be as low as 40% of the density of the solid plastic. Thus, with a rigid skin and a less dense bulk, molded parts have a high stiffness-to-weight ratio (see also Fig. 3.2).

Polyurethane Foam Processing. Products such as furniture cushions and insulating blocks are made by this process. Basically, the operation starts with the mixing of 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 to provide sound and thermal insulation. Various low-pressure and high-pressure machines are available, having computer controls for proper mixing. The mixture solidifies with 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 that have been used in the cold working of metals (such as rolling, closed-die forging, coining, deep drawing, and rubber forming—all described in Part III) also can be used to form thermoplastics at room temperature (*cold forming*). Typical materials formed are polypropylene, polycarbonate, ABS, and rigid PVC. Important considerations regarding this process 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 formed part).

The advantages of the cold forming of plastics over other methods of shaping are as follows:

- Strength, toughness, and uniform elongation are increased.
- Plastics with high molecular weights can be used to make parts with superior properties.
- Forming speeds are not affected by part thickness because (unlike other plastic-processing methods) there is no heating or cooling involved. Cycle times generally are shorter than in molding processes.

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). Thus, the forming operation takes place while the polymer is still in a solid state. The main advantages of solid-phase forming 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

We have described the properties, characteristics, and applications of elastomers and rubbers in Section 7.9. Recall 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 to be processed into various shapes is basically a compound of rubber and various additives and fillers. The additives include *carbon black*—an important element that enhances 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 lower their viscosity; the mixture subsequently is *vulcanized,* using sulfur as the vulcanizing agent. This compound is then ready for further processing (such as calendering, extrusion, and various molding processes), which may also include reinforcements in such forms as fibers and fabric. During processing, the part becomes cross-linked, imparting the desirable properties that we all associate with rubber products ranging from rubber boots to pneumatic tires.

Elastomers can be shaped by a variety of processes that also are used for shaping thermoplastics. Thermoplastic elastomers commonly are shaped by extrusion or injection molding—extrusion being the more economical and the faster process. They also can be formed by blow molding or thermoforming. Thermoplastic polyurethane, for example, can be shaped by all conventional methods. It also can be blended with thermoplastic rubbers, polyvinyl chloride compounds, ABS, and nylon to obtain 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 are used in conjunction with extrusion to impart greater strength. Examples of extruded elastomer products are tubing, hoses, moldings, and inner tubes. Injection-molded elastomer products cover a broad range of applications, such as numerous components for automobiles and appliances.

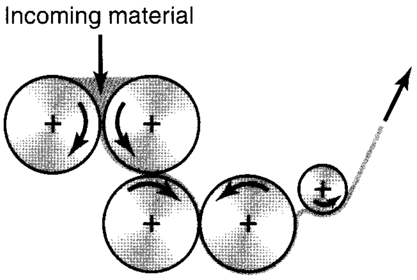


FIGURE 19.20 Schematic illustration of calendering. Sheets produced by this process subsequently are used in thermoforming. The process also is used in the production of various elastomer and rubber products.

Rubber and some thermoplastic sheets are formed by the **calendering** process (Fig. 19.20), wherein a warm mass of the compound is fed into a series of rolls and is **masticated**. The thickness produced is typically 0.3 to 1 mm, but can be made less by stretching the material. It then is stripped off at speeds on the order of 2 m/s to form a sheet, which may be as wide as 3 m. The calendered rubber then may be processed into various products, such as tires and belts for machinery. The rubber or thermoplastics also may be formed over both surfaces of a tape, paper, fabric, or various soft or rigid plastics, thus making them permanently *laminated*. Roll surfaces may be textured to produce a rubber sheet with various patterns and designs.

Discrete rubber products, such as gloves, are made by *dipping* a metal form (such as in the shape of a hand for making gloves) repeatedly into a liquid compound that adheres to the form. A typical compound is *latex*, which is a milklike sap obtained from the inner bark of a tropical tree. The compound is then vulcanized (cross-linked), usually in steam, and stripped from the form and becomes a discrete product.

19.13 Processing Polymer-matrix Composites

As described in Chapter 9, polymer-matrix composites (PMCs) (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, reinforced plastics require special methods to shape them into consumer and industrial products (Fig. 19.21).

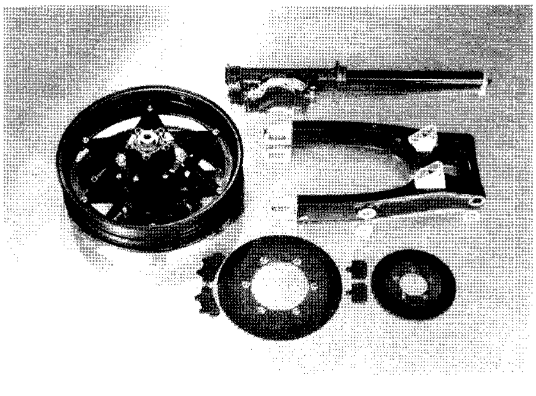


FIGURE 19.21 Reinforced-plastic components for a Honda motorcycle. The parts shown are front and rear forks, a rear swing arm, a wheel, and brake disks.

Polymer-matrix composites can be fabricated by various methods. Fabrication to ensure reliable properties in composite parts and structures, particularly over the long range of their service life, can be challenging because of the presence of two or more types of materials. The matrix and the reinforcing fibers in the composite have—*by design*—very different properties and characteristics and consequently have different responses to the methods of processing (Section 9.2).

The several steps involved in manufacturing reinforced plastics and the time and care required make the processing costs very substantial, and generally, they are not competitive with traditional materials and shapes. This situation has necessitated the careful assessment and integration of design and manufacturing processes (concurrent engineering) in order to take advantages of the unique properties of these composites. This is done while minimizing manufacturing costs and maintaining long-range product integrity, reliability, and production rate. An important safety and environmental concern in reinforced plastics is the dust generated during processing. For example, airborne carbon fibers are known to remain in the work area long after the fabrication of parts has been completed.

19.13.1 Fiber Impregnation

In order to obtain 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.

Prepregs. In a typical procedure for making fiber-reinforced plastic *prepregs* (meaning preimpregnated with resin), the continuous fibers are aligned and subjected to a surface treatment to enhance the 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* (Fig. 19.23a), such as the horizontal stabilizer for the F-14 fighter aircraft.

Typical products 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 typically encountered by military aircraft (including fatigue strength under hot or wet conditions).

Because the process of laying prepreg tapes is a time-consuming and labor-intensive operation, special and highly automated *computer-controlled tape-laying machines* have been built for this purpose (Fig. 19.23b). The prepreg tapes automatically are 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 and quickly for a variety of parts by computer control and with high repeatability.

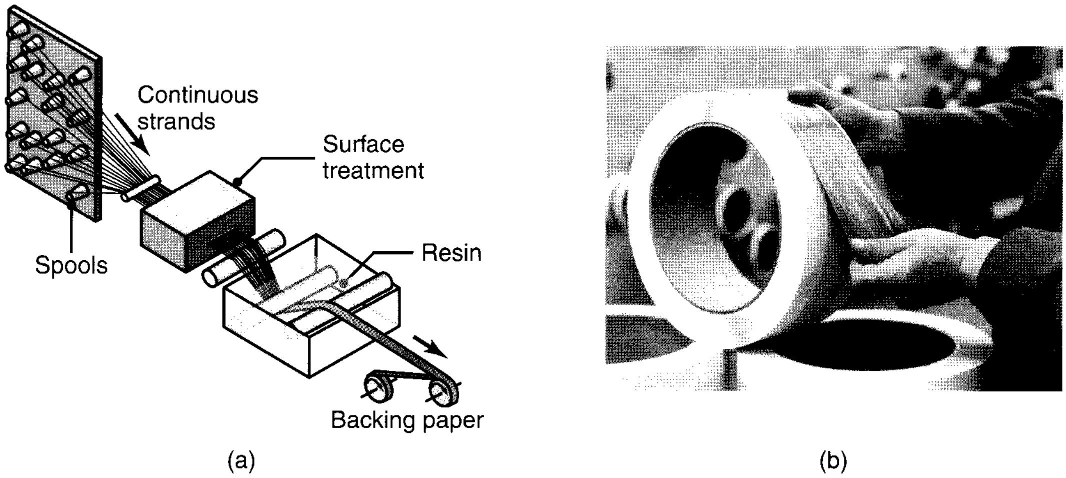


FIGURE 19.22 (a) Manufacturing process for polymer-matrix composite tape. (b) Boron-epoxy prepreg tape. These tapes are then used in making reinforced plastic parts and components with high strength-to-weight and stiffness-to-weight ratios, particularly important for aircraft and aerospace applications and sports equipment. *Source*: (a) After T. W. Chou, R.L. McCullough, and R.B. Pipes. (b) Courtesy of Avco Specialty Materials/Textron.

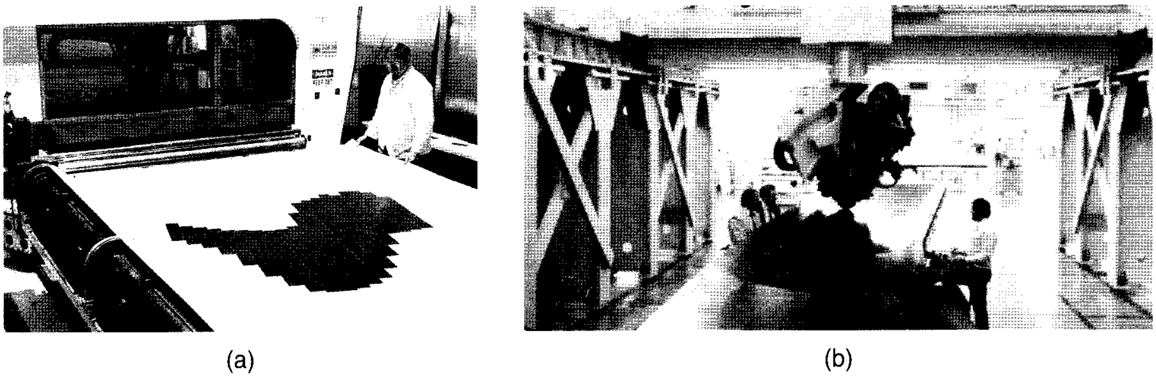

(a) (b)

FIGURE 19.23 (a) Single-ply layup of boron-epoxy tape for the horizontal stabilizer for an F-14 fighter aircraft. (b) A 10-axis computer-numerical-controlled tape-laying system. This machine is capable of laying up 75- and 150-mm wide tapes on contours of up to ±30° and at speeds of up to 0.5 m/s. *Source*: (a) Courtesy of Grumman Aircraft Corporation. (b) Courtesy of The Ingersoll Milling Machine Company.

Sheet-molding Compound. In making *sheet-molding compound* (SMC), continuous strands of reinforcing fiber are first chopped into short fibers (Fig. 19.24) 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 rollers.

The product is then gathered into rolls (or 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 takes one day.

The molding compounds should be stored at a temperature sufficiently low to delay curing. They have a limited shelf life (usually around 30 days) and must be processed within this period. Alternatively, the resin and the fibers can be mixed together only at the time they are placed into the mold.

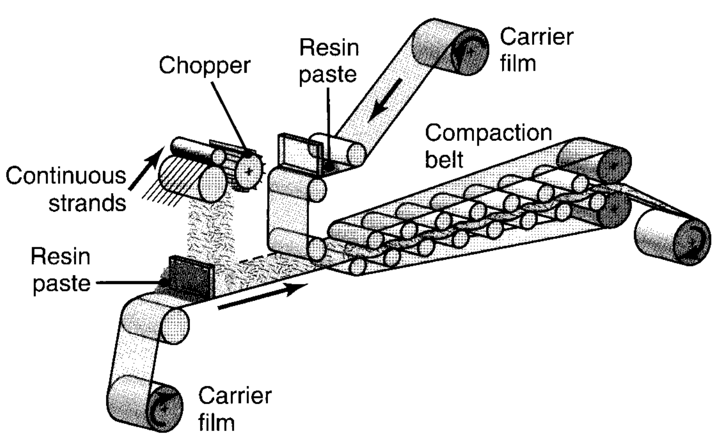


FIGURE 19.24 Schematic illustration of the manufacturing process for producing fiber-reinforced plastic sheets. The sheet still is viscous at this stage and later can be shaped into various products. *Source*: After T.-W. Chou, R.L. McCullough, and R.B. Pipes.

Bulk-molding Compound. *Bulk-molding compounds* (BMC) are in the shape of billets (hence the term bulk) and generally are up to 50 mm (2 in.) in diameter. They are made in the same manner as SMCs and extruded to produce a bulk form. When processed into products, BMCs have flow characteristics that are similar to those of dough; thus, they also are called *dough-molding compounds* (DMC).

Thick-molding Compound. *Thick-molding compounds* (TMC) combine a characteristic of BMCs (lower cost) with one of SMCs (higher strength). They are usually 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, which is a viscous, sticky mixture of polymers, fibers, and additives. Generally, it is molded into the shape of a log, which subsequently is cut or sliced into the desired shape. Fiber lengths generally 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. These compounds 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 prior to placing the SMC in the mold.

Vacuum-bag Molding. In this process (Fig. 19.25a), prepregs are laid in a mold to form the desired shape. The pressure required to shape the product and to develop good bonding is obtained 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. 19.25b). A flexible bag is placed over the resin and reinforcing fiber mixture, and pressure is applied over the mold at a range typically from 200 to 400 kPa. If higher heat and pressure are needed to produce parts with higher density and fewer voids, the entire assembly is put into an *autoclave* (a chamber under heat and pressure).

Care should be exercised to maintain fiber orientation if specific directional properties are desired. In 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 (*release cloth* or *bleeder cloth*) are placed on top of the prepreg sheets.

The molds can be made of metal (usually aluminum), but they more often 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.

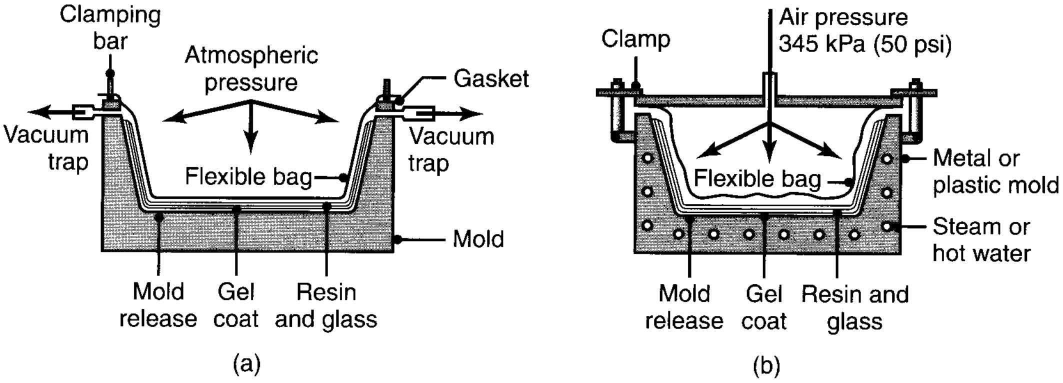


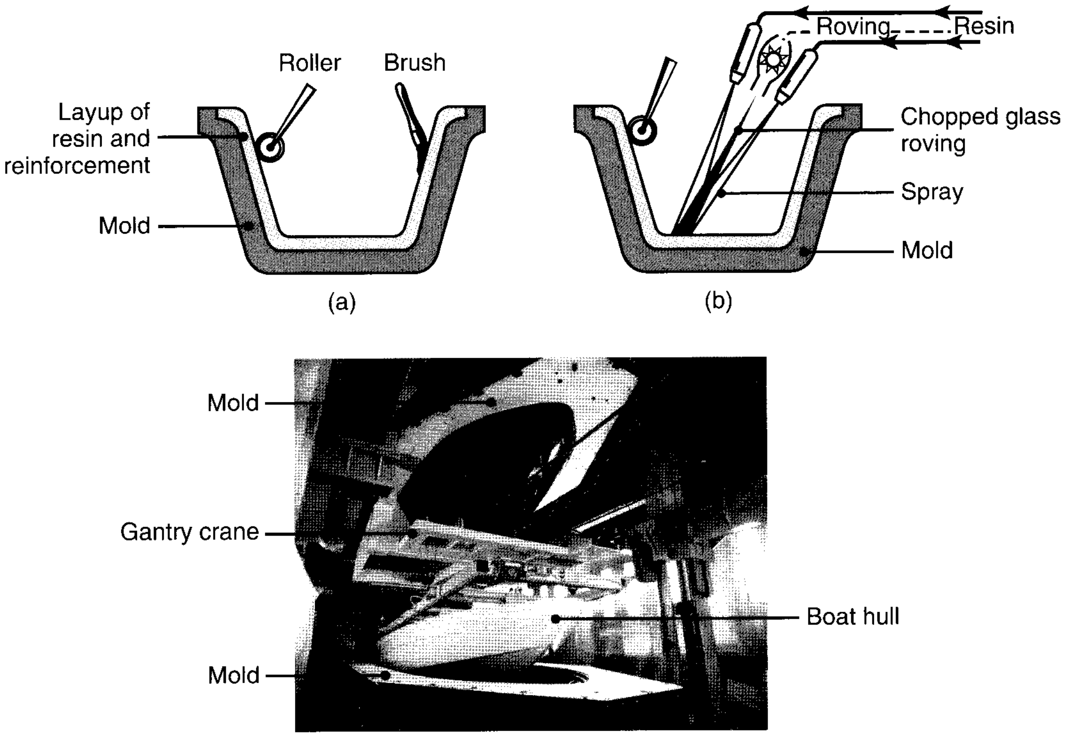
FIGURE 19.25 Schematic illustration of (a) vacuum-bag forming and (b) pressure-bag forming. These processes are used in making discrete reinforced plastic parts.
Source: After T.H. Meister.

Contact Molding. This is a series of processes that use a single male or female mold made of such materials as reinforced plastics, wood, metal, or plaster (Fig. 19.26); it is also referred to as *open-mold processing*. This 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 are 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 placed in proper order (resins and reinforcements), brushed, and shaped in the mold by hand with a roller (Fig. 19.26a). The squeezing action of the roller expels any trapped air bubbles and compacts the part. The reinforcements placed in the mold may consist of various shapes, including prepregs, and their orientation in the final product can be controlled.

In **spray layup**, molding is done by spraying the materials into the mold. As seen in Fig. 19.26b, both the resin and the chopped fibers are sprayed over the mold surfaces. Rolling the deposited materials (as in hand layup) to remove any porosity may be necessary in this process. Since the chopped fibers have random orientations, directional properties cannot be imparted in products made by spray layup.

Both hand layup and spray layup are relatively slow, have high labor costs, and require significant time and labor in finishing operations. Also, the choice of materials that can be used is limited. However, either process is simple to perform,



(a) (b)

(c)

FIGURE 19.26 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 Genmar Holdings, Inc.

and the tooling is inexpensive. Note also that only the mold-side surface of the formed part is smooth, from being in contact with the mold surfaces.

Resin-transfer Molding. This process is based on transfer molding (Section 19.8). A resin is mixed with a catalyst and is forced by a piston-type, positive-displacement pump into the mold cavity that is filled with a fiber reinforcement. The process is a viable alternative to hand layup, sprayup, and compression molding for low- or intermediate-volume production.

Transfer/injection Molding. This is an automated operation that combines compression-molding, injection-molding, and transfer-molding processes. This combination has the good surface finish, dimensional stability, and mechanical properties obtained in compression molding and the high-automation capability and low cost of injection molding and transfer molding.

19.13.3 Filament Winding, Pultrusion, and Pulforming

Filament Winding. *Filament winding* is a process in which the resin and fibers are combined at the time of curing (Fig. 19.27a) in order to develop a composite structure. Axisymmetric parts (such as pipes and storage tanks), and even 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. This process can be modified by wrapping the mandrel with a prepreg material.

The products made by filament winding are very strong because of their highly reinforced structure. Parts as large as 4.5 m in diameter and 20 m long have been made by this process. Filament winding also has been used for strengthening cylindrical or spherical pressure vessels (Fig. 19.27b) made of materials such as aluminum and titanium, where the presence of a metal inner lining makes the part impermeable. Filament winding also can be used directly over solid-rocket propellant forms. Seven-axis computer-controlled machines have been developed that automatically dispense several unidirectional prepregs to also make nonsymmetric parts, such as aircraft engine ducts, fuselages, propellers, blades, and struts.

Pultrusion. Long parts with various uniform cross sections (such as rods, profiles, flat strips, and tubing) are made continuously by the *pultrusion* process. This sequence

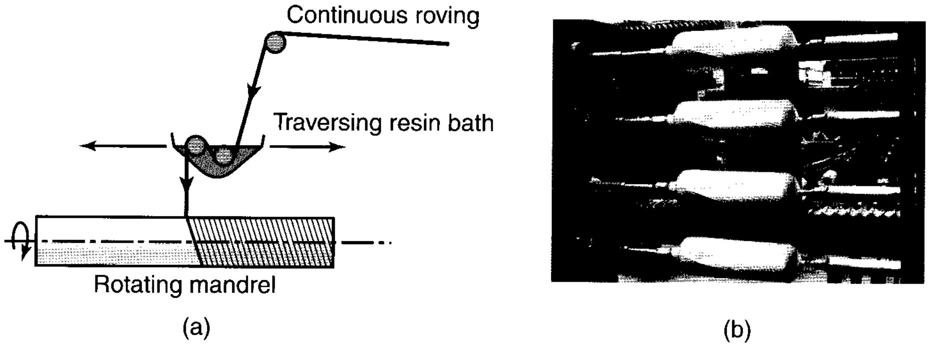


FIGURE 19.27 (a) Schematic illustration of the filament-winding process; (b) fiberglass being wound over aluminum liners for slide-raft inflation vessels for the Boeing 767 aircraft. The products made by this process have a high strength-to-weight ratio and also serve as lightweight pressure vessels. *Source*: Courtesy of Brunswick Corporation.

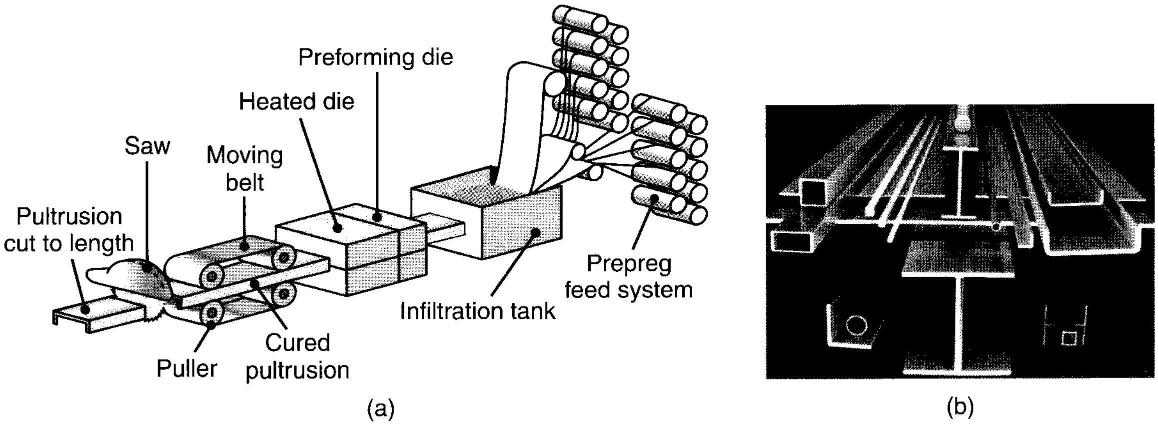


FIGURE 19.28 (a) Schematic illustration of the pultrusion process. (b) Examples of parts made by pultrusion. The major components of fiberglass ladders (used especially by electricians) are made by this process. They are available in different colors, but are heavier because of the presence of glass fibers. *Source*: Courtesy of Strongwell Corporation.

of operations is shown in Fig. 19.28. The continuous reinforcement, glass roving, or fabric (typically made of calcium aluminosilicate glass fiber—E type; see Section 9.2.1) is supplied through several bobbins. The bundle is pulled first through a thermosetting polymer bath (usually polyesters), then through a performing die, and finally through a heated steel die.

The product is cured during its travel through the heated die, which has a length of up to 1.5 m and a speed slow enough to allow sufficient time for the polymer to set. (Note that this is an operation similar to continuously baking bread and cookies or making resin-bonded grinding wheels.) The exiting material is then cut into desired lengths. Typical products made by pultrusion (which may contain up to about 75% reinforcing fiber) are golf clubs, ski poles, fishing poles, driveshafts, and such structural members as ladders, walkways, and handrails. Cross sections as large as 1.5 m × 0.3 m have been made by this process.

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 product. The dies recirculate and shape the products successively. Commonly made products are hammer handles reinforced by glass fibers and curved automotive leaf springs.

EXAMPLE 19.3 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 of polyurea/polyurethanes; and (c) compression-molding of sheet-molding compound (SMC) with resin-transfer-molded polyester and vinylester.

Typical examples of parts made for automobiles are (a) body panels and other large exterior components made by injection molding; (b) front fenders and rear quarter panels made of polyphenylene-ether/ nylon or thermoplastric polyester; (c) outer door panels made of polycarbonate/ ABS; and (d) 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, then assembled mechanically to a steel frame.

Large exterior-body parts also are made of reaction-injection molded (RIM) polyurethane, although polyureas have become important for body panels and bumpers. Thermoset fascias are made of reinforced RIM polyurethane and (more recently) new 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 (RTM).

Environmental and recycling considerations in material and process selection for automobiles have become increasingly 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 previously 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 19.1 Manufacture of Head Protector™ Tennis Racquets

Competitive tennis is a demanding sport, and as a result, there is a strong demand to produce exceptionally lightweight and stiff racquets to improve performance. A tennis racquet consists of a number of regions, as shown in Fig. 19.29. Of particular interest is the "sweet spot"; when the tennis ball is struck at the sweet spot, the player has optimum control, power, and vibration. A number of innovative racquet-head designs have been developed over the years to maximize the size of the sweet spot, and a stiff composite (typically with high-modulus graphite fibers in an epoxy matrix; see Chapter 9) is used in the manufacture of the racquet head. The orientation of the fibers varies in different locations of the racquet; as an example, the main tube for the racquet consists of carbon-epoxy prepreg oriented at ±30° from layer to layer.

The advantages to such materials are clear, in that stiff racquets indicate that higher forces can be applied to the ball. However, the use of these advanced materials has led to an increased frequency of tennis elbow, a painful condition associated with the 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

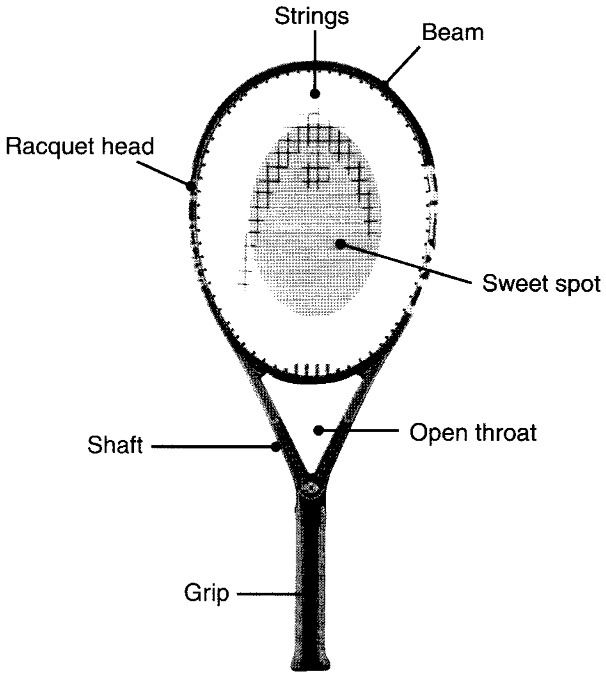


FIGURE 19.29 A Head Protector™ tennis racquet. *Source*: Courtesy of Head Sport AG.

vibration of the racquet encountered with every stroke, especially when balls are struck away from the sweet spot.

An innovative tennis-racquet design, the Protector™ (produced 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 (see 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 (i.e., left, right, front, and back). The module consists of about 50 PZT fibers (each approximately 0.3 mm in diameter) sandwiched between two polyamide layers

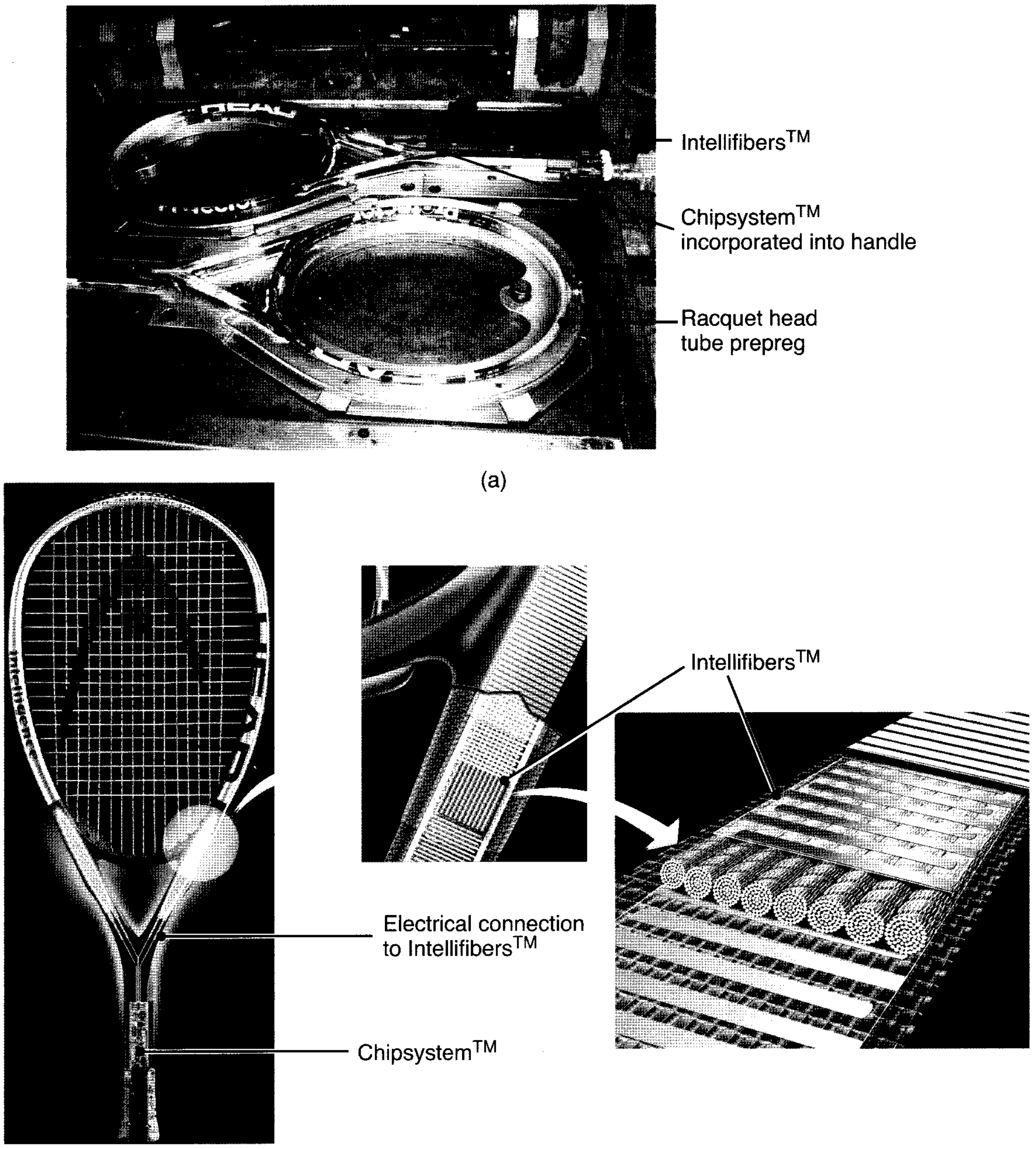


FIGURE 19.30 (a) The composite Head Protector™ racquet immediately after molding; (b) a completed Head Protector™ racquet, highlighting the incorporation of piezoelectric Intellifibers™. *Source*: Courtesy of Head Sport AG.

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 at low current. The energy is stored in coils on the printed circuit board (Chipsystem™) incorporated in the racquet handle in real time 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 can damp vibrations only within a range of its design frequency.

The manufacture of a Protector™ tennis racquet involves a number of steps. First, a carbon-epoxy prepreg is produced, as described in Section 19.13.1. 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. A polyamide sleeve (or bladder) is then placed over a rod, and the prepreg is rolled over the sleeve. 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 by drilling holes into the preform during finishing.

Prior to molding, all the components are assembled onto a template, and 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 to form the handle. The PZT fibers are incorporated as the outer layer in the racquet in the throat area, and the printed electrodes are connected 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. 19.25b). A racquet as it appears directly after molding is shown in Fig. 19.30a.

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. 19.30b. 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.

19.14 Processing Metal-matrix and Ceramic-matrix Composites

Metal-matrix composites (MMCs) 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 is used to force the liquid-metal matrix into a preform (usually shaped out of wire or sheet and made of reinforcing fibers).
- **Solid-phase processing** consists basically of *powder-metallurgy* techniques, including cold and hot isostatic pressing. Proper mixing is important in order to obtain a homogeneous distribution of the fibers. An example of this technique 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 powder-metallurgy techniques generally are heat treated for optimum properties.

EXAMPLE 19.4 Metal-matrix Composite Brake Rotors and Cylinder Liners

Some brake rotors currently are being made of composites consisting of an aluminum-based matrix reinforced with 20% silicon-carbide particles. The particles are stirred into molten aluminum alloys, and the mixture is cast into ingots. The ingots then are remelted and cast into shapes (such as brake rotors and drums) by casting processes such as green-sand, bonded-sand, investment, permanent-mold, and squeeze casting. These rotors (a) are about one-half the weight of those made of gray cast iron, (b) conduct heat three times faster, (c) add the stiffness and wear-resistance characteristics of ceramics, and (d) reduce noise and vibration because of internal damping.

To improve the wear- and heat resistance of cast-iron cylinder liners in aluminum engine blocks, aluminum-matrix liners also are being developed. 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 more recent techniques such as melt infiltration, controlled oxidation, and hot-press sintering, are used to make ceramic-matrix composites (CMCs). Although the more recent processes are largely in the experimental stage, they continue to be developed to improve the properties and performance of these composites.

- **Slurry infiltration** is the most common process for making ceramic-matrix composites. It involves the preparation of a fiber preform that is first hot pressed and then impregnated with a combination of slurry (which contains the matrix powder), a carrier liquid, and an organic binder. High strength, toughness, and uniform structure are obtained by this process, 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 the forming and shaping of plastics are similar to those for casting metals. 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 after its life cycle. (Some of these issues were described in Sections I.4 and I.6 in the General Introduction and Section 7.8.)

Outlined here are the general design guidelines for the production of plastic and composite-material parts:

1. As we have seen, processes for plastics have inherent flexibility to produce a wide variety of part shapes and sizes. 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. In considering possible process substitutions, it is necessary to bear in mind that the materials involved and their characteristics are quite different—each having its own properties that are important to a particular application.
2. Compared with metals, plastics have much lower stiffness and strength. Therefore, section sizes and shapes should 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. Large, flat surfaces can be stiffened by such simple means as prescribing curvatures on parts. For example, observe the stiffness of very thin, but gently curved, slats in venetian blinds. Such simple design aspects can be ascertained by inspecting the differences between the hoods of garden tractors—for example, one made of sheet metal and the other of plastic. Reinforcement with fibers or particles also are effective in achieving lightweight and stiff designs.
3. 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 designs of the part and the die should be such that they will not present difficulties concerning proper shape generation (Fig. 19.31), dimensional control, and surface finish. Because of low stiffness and thermal effects, dimensional tolerances (especially for thermoplastics) are not as small as in metalworking processes. Dimensional tolerances are much smaller in injection molding, for example, than they are in thermoforming. As in the casting of metals and alloys, the control of material flow in the mold cavities is important. The effects of molecular orientation during the processing of the polymer also should be considered, especially in extrusion, thermoforming, and blow molding.

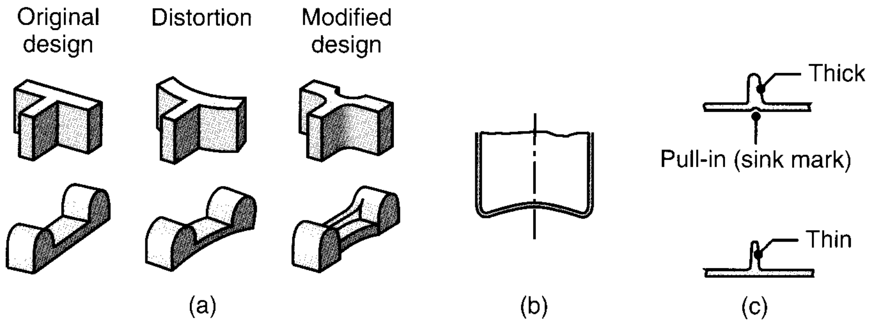


FIGURE 19.31 Examples of design modifications to eliminate or minimize distortion in plastic parts: (a) suggested design changes to minimize distortion; (b) stiffening the bottoms of thin plastic containers by doming—a technique similar to the process used to shape the bottoms of aluminum beverage cans; 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.

4. Large variations in cross-sectional areas and section thicknesses, as well as abrupt changes in geometry, should be avoided to achieve the desired shape. Note, for example, that the *sink marks* (pull-in) shown in the top piece in Fig. 19.31c are due to the fact that thick sections solidify last. Furthermore, contraction in larger cross sections tends to cause *porosity* in plastic parts, as it does in metal casting (see Fig. 12.2), thus affecting product quality. By contrast, a lack of stiffness may make it more difficult to remove thin parts from molds after shaping them.
5. The low elastic moduli of plastics further requires that shapes be selected properly for improved stiffness of the component (Fig. 19.31b), 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 molds and dies. Generally, the recommended part thickness ranges from about 1 mm for small parts to about 3 mm for large parts.
6. Physical properties (especially a high coefficient of thermal expansion) are important considerations. Improper part design or assembly can lead to distortion (warping) and uneven shrinking (Fig. 19.31a). Plastics can be molded easily around metallic parts and inserts; however, their interfacial strength and compatibility with metals when so assembled is an important consideration.
7. 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 the thermal cycling of the part during processing. The magnitude and direction of residual stresses (however produced) are important factors, such as in stress cracking. Furthermore, these stresses can relax over a period of time and cause distortion of the part during its service life.
8. A major design advantage of reinforced plastics is the directional nature of the strength of the composite (for example, see Fig. 9.7). External forces applied to the material are transferred by the resin 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 and weak in others.

To achieve strength in two principal directions, individual unidirectional layers are laid at the corresponding angles to each other, as is done in 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

As in all other manufacturing processes, design and manufacturing decisions are based ultimately on performance and cost, including the cost of equipment, tooling, labor, and production. The final selection of a process or a sequence of processes also depends greatly on production volume. 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 economical production quantities.

It can be seen from this table that equipment costs and tool and die costs are related somewhat, and for processes such as thermoforming, casting, and rotational

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.

molding, these costs are relatively low. The most expensive are injection-molding machines, followed by compressing-molding and transfer-molding machines. Tool and die costs are also high. 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 die casting.

Larger dies may be considered in order to accommodate a larger number of 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. Therefore, a detailed analysis must be made 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 apply to the other plastic-processing methods described throughout this chapter.

For the production of composite materials, equipment and tooling costs for most molding operations are generally high, production rates are low, and economical production quantities vary widely. High equipment and tooling costs can be acceptable if the production run is large.

SUMMARY

- Thermoplastics can be shaped by a variety of processes, including extrusion, molding, casting, and thermoforming, as well as by some of the processes used in metalworking. 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 their physical properties, such as high thermal expansion and generally low resistance to temperature. Inspection techniques are available to determine the integrity of these products.
- Because of their expanding use in critical applications, the processing of metal-matrix and ceramic-matrix composites has undergone important developments to ensure product integrity and reduce 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.
- Because of the wide variety of low-cost materials and manufacturing techniques available, the economics of processing plastics and composite materials is an important consideration, particularly when compared with that of metal components. The relevant factors include machinery and die costs, 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

Baird, D.G., and Collias, D.I., **Polymer Processing: Principles and Design**, Wiley, 1998.

Beaumont, J.P., **Runner and Gating Design Handbook: Tools for Successful Injection Molding**, Hanser Gardner, 2008.

Belcher, S.L., **Practical Guide to Injection Blow Molding**, CRC Press, 2007.

Berins, M.L. (ed.), **Plastics Engineering Handbook**, 5th ed., Chapman & Hall, 1994.

Buckley, C.P., Bucknall, C.B., and McCrum, N.G., **Principles of Polymer Engineering**, 2nd ed., Oxford, 1997.

Campbell, F., **Manufacturing Processes for Advanced Composites**, Elsevier, 2004.

Campbell, P.D.O., **Plastic Component Design**, Industrial Press, 1996.

Chanda, J.M., and Roy, S.K., **Plastics Technology Handbook**, 3rd ed., Marcel Dekker, 1998.

Dave, R.S., and Loos, A.C., **Processing of Composites**, Hanser Gardner, 2000.

Gerdeen, J.C., Lord, H.W., and Rorrer, R.A.L., **Engineering Design with Polymers and Composites**, CRC Press, 2006.

Gramman, P.J., Osswald, T.A., Rios, A.C, and Davis, B.A., **Compression Molding**, Hanser Gardner, 2004.

Hans, G., and Unger, P., **Injection Molds: 130 Proven Designs**. 4th ed., Hanser Gardner, 2007.

———, *Rheology and Processing of Polymeric Materials:* Vol. 2: **Polymer Processing**, Oxford University Press, 2008.

Harper, C., **Handbook of Plastics, Elastomers, and Composites**, 3rd ed., McGraw-Hill, 1996.

Hernandez-Ortiz, J.P., **Polymer Processing: Modeling and Simulation**, Hanser-Gardner 2006.

Johnson, P.S., **Rubber Processing: An Introduction**, Hanser Gardner, 2001.

Karszes, W.M., **Extrusion Coating: Equipment and Materials**, Tappi Press, 1997.

Kazmer, D.O., **Injection Mold Design Engineering**, Hanser Gardner, 2008.

Lee, N.C., **Blow Molding Design Guide**, Hanser Gardner, 1998.

Mazumdar, S.K., **Composites Manufacturing: Materials, Products and Process Engineering**, CRC Press, 2001.

Mallick, P.K. (ed.), **Composites Engineering Handbook**, Marcel Dekker, 1997.

———, **Fiber-reinforced Composites: Materials, Manufacturing and Design**, 3rd ed., Dekker, 2008.

Micheli, W., **Extrusion Dies for Plastics and Rubber: Design and Engineering Computations**, Hanser Gardner, 2004.

Mills, N., **Mechanical Design of Polymer Foam Products**, Chapman & Hall, 1997.

Modern Plastics Encyclopedia, McGraw-Hill, annual.

Olmsted, B.A., and Davis, M., **Practical Injection Molding**, CRC Press, 2001.

Osswald, T.A., Turng, L.-S., and Gramann, P., **Injection Molding Handbook**, 2nd ed., Hanser Gardner, 2008.

Potter, K., **Resin Transfer Moulding**, Chapman & Hall, 1997.

Quinn, J.A., **Composites Design Manual**, Technomic, 1999.

Rauwendaal, C., **Polymer Extrusion**, 4th ed., Hanser Gardner, 2001.

Rosato, D.V., **Plastics Processing Data Handbook**, 2nd ed., Chapman & Hall, 1997.

Rosen, S. R., **Thermoforming: Improving Process Performance**, Society of Manufacturing Engineers, 2002.

Schwartz, M.M., **Composite Materials**, Vol. 2: **Processing, Fabrication, and Applications**, Prentice Hall, 1999.

Starr, T.F., **Pultrusion for Engineers**, CRC Press, 2000.

Strong, A.B., **Plastics: Materials and Processing**, 3rd ed., Prentice Hall, 2005.

———, **Fundamentals of Composites Manufacturing**, 2nd ed., Society of Manufacturing Engineers, 2008.

Tadmore, Z., and Gogos, C.G., **Principles of Polymer Processing**, 2nd ed., Wiley, 2006.

Wilkinson, A.N., and Ryan, A., **Polymer Processing and Structure Development**, Kluwer, 1999.

REVIEW QUESTIONS

19.1. What are the forms of raw materials for processing plastics into products?

19.2. How are injection-molding machines rated?

19.3. What is (a) a parison, (b) a plastisol, and (c) a prepreg?

19.4. How is thin plastic film produced?

19.5. List several common products that can be made by thermoforming.

19.6. What similarities and differences are there between compression molding and closed-die forging?

19.7. Explain the difference between potting and encapsulation.

19.8. Describe the advantages of cold-forming plastics over other plastic-processing methods.

19.9. What are the characteristics of filament-wound products? Explain why they are desirable.

19.10. Describe the methods that can be used to make tubular plastic products.

19.11. What is pultrusion? Pulforming?

19.12. How are plastic sheet and plastic film produced?

19.13. What process is used to make foam drinking cups?

19.14. If a polymer is in the form of a thin sheet, is it a thermoplastic or thermoset? Why?

19.15. How are polymer fibers made? Why are they much stronger than bulk forms of the polymer?

19.16. What are the advantages of coextrusion?

QUALITATIVE PROBLEMS

19.17. Describe the features of a screw extruder and its functions.

19.18. Explain why injection molding is capable of producing parts with complex shapes and fine detail.

19.19. Describe the advantages of applying the traditional metal-forming techniques, described in Chapters 13 through 16, to making (a) thermoplastic and (b) thermoset products.

19.20. Explain the reasons that some plastic-forming processes are more suitable for certain polymers than for others. Give examples.

19.21. Describe the problems involved in recycling products made from reinforced plastics.

19.22. Can thermosetting plastics be used in injection molding? Explain.

19.23. Inspect some plastic containers, such as those containing talcum powder, and note that the integral lettering on them is raised rather than depressed. Explain.

19.24. 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.

19.25. Explain why operations such as blow molding and film-bag making are performed vertically.

19.26. Comment on the principle of operation of the tape-laying machine shown in Fig. 19.23b.

19.27. Typical production rates are given in Table 19.2. Comment on your observations and explain why there is such a wide range.

19.28. What determines the cycle time for (a) injection molding, (b) thermoforming, and (c) compression molding? Explain.

19.29. Does the pull-in defect (sink marks) shown in Fig. 19.31c also occur in metal-forming and casting processes? Explain.

19.30. What determines the intervals at which the indexing head in Fig. 19.14c rotates from station to station?

19.31. Identify processes that would be suitable for small production runs on plastic parts, of, say, 100.

19.32. 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.

QUANTITATIVE PROBLEMS

▶ **19.33.** 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.

▶ **19.34.** 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.

▶ **19.35.** 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.

SYNTHESIS, DESIGN, AND PROJECTS

19.36. 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.

19.37. Explain the design considerations involved in replacing a metal beverage can with one made completely of plastic.

19.38. Give examples of several parts suitable for insert molding. How would you manufacture these parts if insert molding were not available?

19.39. Give other examples of design modifications in addition to those shown in Fig. 19.31.

19.40. With specific examples, discuss the design issues involved in making products out of plastics vs. reinforced plastics.

19.41. Inspect various plastic components in a typical automobile, and identify the processes that could have been used in making them.

19.42. 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.

19.43. 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 electrician's tools, such as wire cutters and pliers requiring electrical insulation.

19.44. 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.

19.45. 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.

▶ **19.46.** 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.

19.47. 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.

19.48. 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.

19.49. Examine some common and colorful plastic poker chips and give an opinion on how they were manufactured.

19.50. 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.

Rapid-Prototyping Processes and Operations

CHAPTER 20

- This chapter describes the technologies associated with rapid prototyping, sharing the characteristics of computer integration, production without the use of traditional tools and dies, and the ability to rapidly produce a single part on demand; they all have the basic characteristics of producing individual parts layer by layer.
- The chapter discusses the (nonmetallic and metallic) materials used in rapid prototyping and describes the commercially important rapid-prototyping technologies.
- These processes include fused-deposition modeling, stereolithography, multijet modeling, polyjet modeling, three-dimensional printing, and selective laser sintering.
- The chapter ends with a description of the revolutionary practice of applying rapid-prototyping techniques to the production of tooling (rapid tooling) that can be used in other manufacturing processes.

Typical parts made: A wide variety of metallic and nonmetallic parts for product design analysis, evaluation and finished products.

Alternative processes: Machining, casting, molding, and fabricating.

EXAMPLES:

CASE STUDY:

20.1 Introduction

In the development of a new product, there is invariably a need to produce a single example, or **prototype,** of a designed part or system before allocating large amounts of capital to new production facilities or assembly lines. The main reasons for this need are that the capital cost is very high and production tooling takes considerable time to prepare. Consequently, a working prototype is needed for design evaluation and troubleshooting before a complex product or system is ready to be produced and marketed.

A typical product development process was outlined in Fig. I.3 in the General Introduction. An iterative process naturally occurs when (a) errors are discovered or (b) more efficient or better design solutions are gleaned from the study of an earlier generation prototype. The main problem with this approach, however, is that the

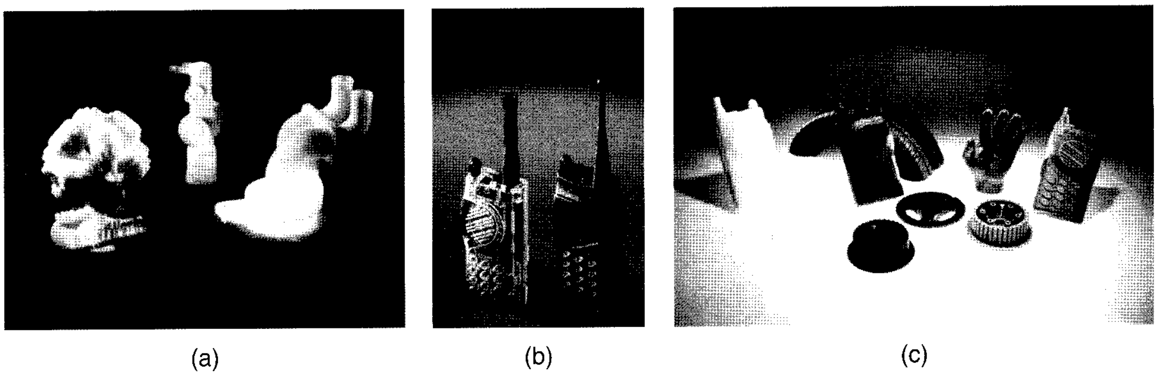


FIGURE 20.1 Examples of parts made by rapid-prototyping processes: (a) selection of parts from fused-deposition modeling; (b) stereolithography model of cellular phone; and (c) selection of parts from three-dimensional printing. *Source*: (a) Courtesy of Stratasys, Inc., (b) and (c) Courtesy of 3D Systems, Inc.

production of a prototype can be extremely time consuming. Tooling can take several months to prepare, and the production of a single complicated part by conventional manufacturing operations can be very difficult. Furthermore, during the time that a prototype is being prepared, facilities and staff still generate costs.

An even more important concern is the speed with which a product flows from concept to a marketable item. In a competitive marketplace, it is well known that products that are introduced before those of their competitors generally are more profitable and enjoy a larger share of the market. At the same time, there are important concerns regarding the production of high-quality products. For these reasons, there is a concerted effort to bring high-quality products to market quickly.

A technology that speeds up the iterative product-development process considerably is the concept and practice of **rapid prototyping** (RP)—also called **desktop manufacturing, digital manufacturing,** or **solid free-form fabrication.** Examples of rapid-prototyped parts are shown in Fig. 20.1.

EXAMPLE 20.1 Functional Rapid Prototyping

Toys are examples of mass-produced products that have universal appeal. However, some toys are actually quite complex, and the function of a computer-aided design (CAD) cannot be ensured until prototypes are produced. Figure 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, and the prototype could actually hold and squirt water. The alternative would be to produce components on CNC milling machines or fabricate them in another fashion, but this can be done only at much higher cost.

By producing a prototype, interference issues and assembly problems can be assessed and corrected if necessary. Further, from an aesthetic standpoint, the elaborate decorations on such a toy can be more effectively evaluated from a prototype than on a CAD file and can be adjusted to improve the appeal of the toy. 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.

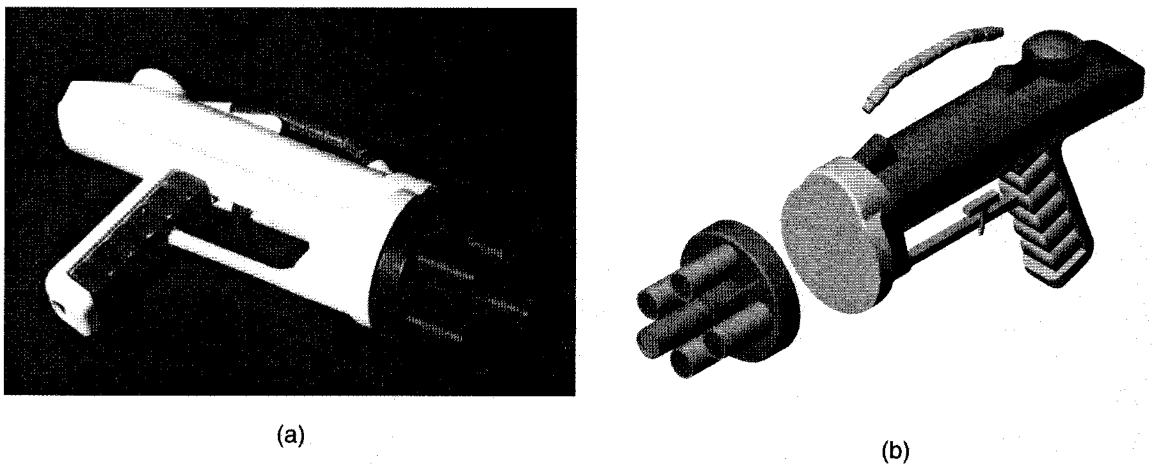


FIGURE 20.2 Rapid prototyping of a Super Soaker™ squirt gun. (a) Fully functional toy produced through fused-deposition modeling; (b) original CAD description. *Source*: Courtesy of Rapid Models and Prototypes, Inc., and Stratasys, Inc.

Developments in rapid prototyping began in the mid-1980s. The advantages of this technology include the following:

- Physical models of parts produced from CAD data files can be manufactured in a matter of hours and allow the rapid evaluation of manufacturability and design effectiveness. In this way, rapid prototyping serves as an important tool for visualization and for concept verification.
- With suitable materials, the prototype can be used in subsequent manufacturing operations to produce the final parts. Sometimes called *direct prototyping*, this approach can serve as an important manufacturing technology.
- Rapid-prototyping operations can be used in some applications to produce actual tooling for manufacturing operations (**rapid tooling**, see Section 20.5.1). Thus, one can obtain tooling in a matter of a few days.

Rapid-prototyping processes can be classified into three major groups: **subtractive**, **additive**, and **virtual**. As the names imply, subtractive processes involve material removal from a workpiece that is larger than the final part. Additive processes build up a part by adding material incrementally to produce the part. Virtual processes use advanced computer-based visualization technologies.

Almost all materials can be used through one or more rapid-prototyping operations, as outlined in Table 20.1. However, because their properties are more suitable for these operations, polymers are the workpiece material most commonly used today, followed by metals and ceramics. Still, new processes are being introduced continually. The more common materials used in rapid-prototyping operations are summarized in Table 20.2. This chapter is intended to serve as a general introduction to the most common rapid-prototyping operations, describe their advantages and limitations, and explore the present and future applications of these processes.

TABLE 20.1

Characteristics of Additive Rapid-prototyping 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)
Multijet/polyjet modeling	Liquid	Liquid layer curing	Photopolymerization	Photopolymers
Fused-deposition modeling	Liquid	Extrusion of melted polymer	Solidification by cooling	Polymers (such as ABS, polycarbonate, and polysulfone)
Ballistic-particle manufacturing	Liquid	Droplet deposition	Solidification by cooling	Polymers and wax
Three-dimensional printing	Powder	Binder-droplet deposition onto powder layer	No phase change	Ceramic, polymer, metal powder, and sand
Selective laser sintering	Powder	Layer of powder	Sintering or melting	Polymers, metals with binder, metals, ceramics and sand with binder
Electron-beam melting	Powder	Layer of powder	Melting	Titanium and titanium alloys, cobalt chrome
Laminated-object manufacturing	Solid	Deposition of sheet material	No phase change	Paper and polymers

20.2 Subtractive Processes

Making a prototype traditionally has involved a series of processes using a variety of tooling and machines, and it usually takes anywhere from weeks to months, depending on part complexity and size. This approach requires skilled operators using *material removal by machining and finishing operations* (as described in detail in Part IV)—one by one—until the prototype is completed. To speed the process, subtractive processes increasingly use computer-based technologies such as the following:

- **Computer-based drafting packages**, which can produce three-dimensional representations of parts.
- **Interpretation software**, which can translate the CAD file into a format usable by manufacturing software.
- **Manufacturing software**, which is capable of planning the operations required to produce the desired shape.
- **Computer-numerical-control** (CNC) **machinery** with the capabilities necessary to produce the parts.

When a prototype is required only for the purpose of shape verification, a soft material (usually a polymer or a wax) is used as the workpiece in order to reduce or avoid any machining difficulties. The material intended for use in the actual application also can be machined, but this operation may be more time consuming, depending on the machinability of the material. Depending on part complexity and

TABLE 20.2

Mechanical Properties of Selected Materials for Rapid Prototyping

Process	Material	Tensile strength (MPa)	Elastic modulus (GPa)	Elongation in 50 mm (%)	Notes
Stereo-lithography	Somos 7120a	63	2.59	2.3–4.1	Transparent amber; good general-purpose material for rapid prototyping
	Somos 9120a	32	1.14–1.55	15–25	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	3.3–3.5	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 rapid prototyping
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 rapid prototyping and general use
	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
Selective laser sintering	Duraform PA	43	1.6	14	White; produces durable heat- and chemical-resistant parts; suitable for snap-fit assemblies and sandcasting or silicone tooling
	Duraform GF	27	4.0	1.4	White; glass-filled form of Duraform PA 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

machining capabilities, prototypes can be produced in a few days to a few weeks. Subtractive systems can take many forms; they are similar in approach to the manufacturing cells described in Section 39.2. Operators may or may not be involved, although the handling of parts is usually a human task.

20.3 Additive Processes

Additive rapid-prototyping operations all build parts in *layers*, and as summarized in Table 20.1, they consist of *stereolithography, Multijet/polyjet modeling, fused-deposition modeling, ballistic-particle manufacturing, three-dimensional printing, selective laser sintering, electron-beam and laminated-object manufacturing.* In order to visualize the methodology used, it is beneficial to think of constructing a loaf of bread by stacking and bonding individual slices on top of each other. All of the processes described in this section *build parts slice by slice*. The main difference between the various additive processes lies in the method of producing the individual slices, which are typically 0.1 to 0.5 mm thick and can be thicker for some systems.

All additive operations require elaborate software. As an example, note the solid part shown in Fig. 20.3a. The first step is to obtain 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 rapid-prototyping machine with detailed information regarding the manufacture of the part. Fig. 20.3d shows the paths of the extruder in one slice, using the fused-deposition-modeling operation described in Section 20.3.1.

This approach requires operator input in the setup of the proper computer files and in the initiation of the production process. Following that stage, the machines generally operate unattended and provide a rough part after a few hours. The part is then subjected to a series of manual finishing operations (such as sanding and painting) in order to complete the rapid-prototyping process.

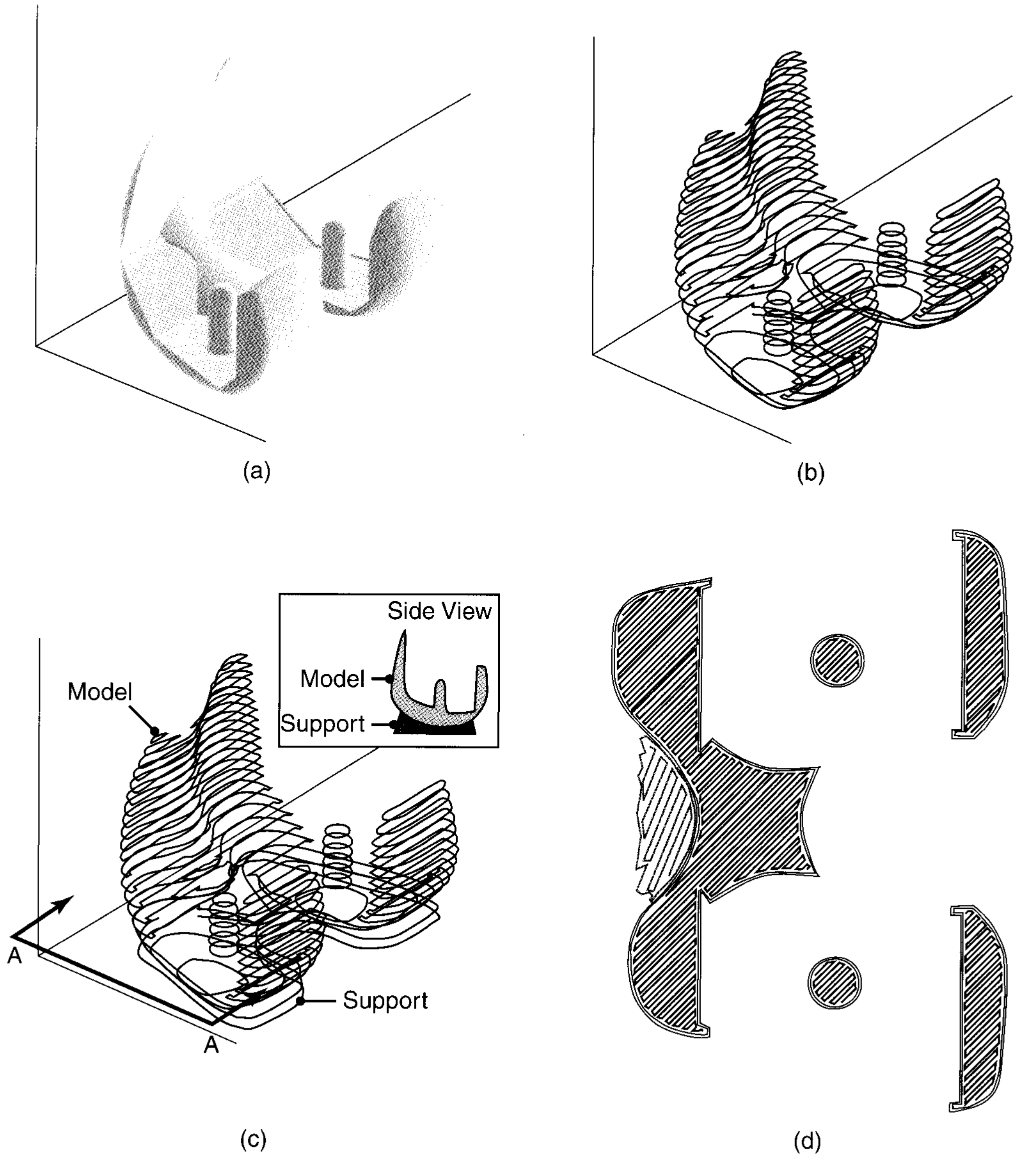
It should be recognized that the setup and finishing operations are very labor intensive and that the production time is only a portion of the time required to obtain a prototype. In general, however, additive processes are much faster than subtractive processes, taking as little as a few minutes to a few hours to produce a part.

20.3.1 Fused-deposition Modeling

In the *fused-deposition-modeling* (FDM) process (Fig. 20.4), a gantry robot–controlled extruder head moves in two principal directions over a table, which can be raised and lowered as needed. A thermoplastic filament is extruded through the small orifice of a heated die. The initial layer is placed on a foam foundation by extruding the filament at a constant rate while the extruder head follows a predetermined path (see Fig. 20.3d). When the first layer is completed, the table is lowered so that subsequent layers can be superimposed.

Occasionally, complicated parts are required, such as the one shown in Fig. 20.5a. This part is difficult to manufacture directly, because once the part has been constructed up to height a, the next slice would require the filament to be placed in a location where no material exists beneath to support it. The solution is to extrude a support material separately from the modeling material, as shown in Fig. 20.5b. Note that the use of such support structures allows all of the layers to be supported by the material directly beneath them. The support material is produced with a less dense filament spacing on a layer, so it is weaker than the model material and can be broken off easily after the part is completed.

The layers in an FDM model are determined by the extrusion-die diameter, which typically ranges from 0.050 to 0.12 mm. This thickness represents the best achievable tolerance in the vertical direction. In the x–y plane, dimensional accuracy can be as fine as 0.025 mm—as long as a filament can be extruded into the feature. A variety of polymers are available for different applications. *Flat wire metal*



(a) (b)

(c) (d)

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.

deposition uses a metal wire instead of a polymer filament, but also needs a laser to heat and bond the deposited wire to build parts.

Close examination of an FDM-produced part will indicate that a stepped surface exists on oblique exterior planes. If this surface roughness is objectionable, a heated tool can be used to smooth the surface, the surface can be hand sanded, or a coating can be applied (often in the form of a polishing wax). However, the overall tolerances are then compromised unless care is taken in these finishing operations.

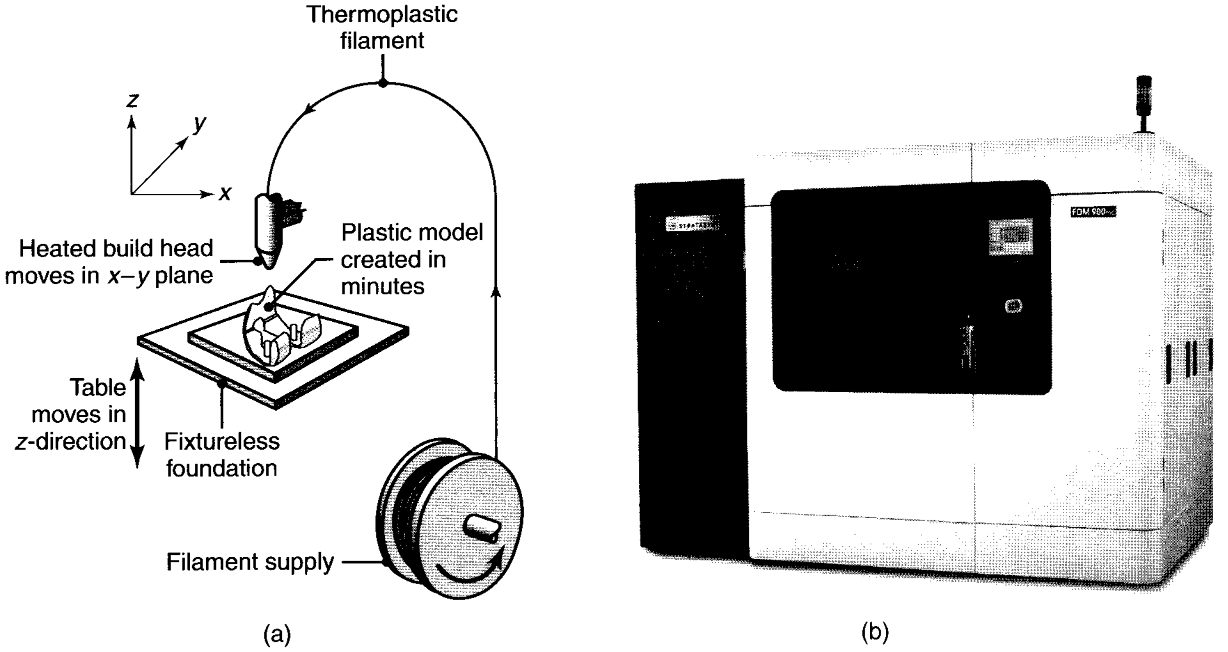


FIGURE 20.4 (a) Schematic illustration of the fused-deposition-modeling process. (b) The FDM 900mc, a fused-deposition-modeling machine. *Source*: Courtesy of Stratasys, Inc.

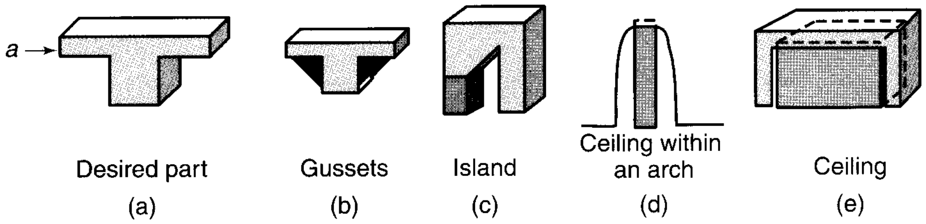


FIGURE 20.5 (a) A part with a protruding section that requires support material. (b)–(e) Common support structures used in rapid-prototyping machines. *Source*: After P.F. Jacobs, *Rapid Prototyping & Manufacturing: Fundamentals of Stereolithography*. Society of Manufacturing Engineers, 1992.

Although some FDM machines can be obtained for around \$20,000, others can cost as much as \$300,000. The main differences between them are the maximum size of the parts that can be produced and the numbers and types of materials that can be used.

20.3.2 Stereolithography

A common rapid-prototyping process—one that actually was developed prior to fused-deposition modeling—is *stereolithography* (STL). This process (Fig. 20.6) is based on the principle of *curing* (hardening) 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).

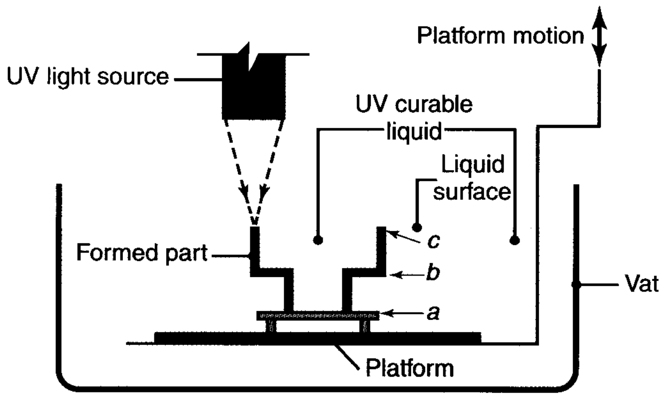


FIGURE 20.6 Schematic illustration of the stereolithography process.

At its highest position (depth *a* in Fig. 20.6), 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 produces a solid body. The platform is then lowered sufficiently to cover the cured polymer with another layer of liquid polymer, and the sequence is repeated. The process is repeated until level *b* in Fig. 20.6 is reached. Thus far, we have generated a cylindrical part with a constant wall thickness. Note that the platform is now lowered by a vertical distance *ab*.

At level *b*, the x–y movements of the beam define a wider geometry, so we now have a flange-shaped portion that is being produced over the previously formed part. After the proper 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 process, comes from the facts that the movements are three-dimensional and the process is similar to lithography (see Section 28.7), in which the image to be printed on a flat surface is ink receptive and the blank areas are ink repellent. Note also that, like FDM, stereolithography can utilize a weaker support material. In stereolithography, this support takes the form of perforated structures. After its completion, the part is removed from the platform, blotted, and cleaned ultrasonically and with an alcohol bath. Then the support structure is 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, it is around 0.0125 mm. Oblique surfaces also can be of very high quality.

Solid parts can be produced by applying special laser-scanning patterns to speed up production. For example, by spacing scan lines in stereolithography, volumes or pockets of uncured polymer can be formed within cured shells. When the part is later placed in a postprocessing oven, the pockets cure and a solid part forms. Similarly, parts that are to be investment cast 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 a day—without postprocessing 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 from \$80 per litre. The maximum part size that can be produced is $0.5 \times 0.5 \times 0.6$ m..

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 thinner layers and lower volumetric cure rates. When stereolithography is used to fabricate micromechanical systems (see Chapter 29), it is called **microstereolithography**.

20.3.3 Multijet/Polyjet Modeling

The **Multijet Modeling** (MJM) or **Polyjet process** is similar to inkjet printing, where print heads deposit the photopolymer on the build tray. Ultraviolet bulbs, alongside the jets, immediately cure and harden each layer, thus eliminating the need for any postmodeling curing that is needed in stereolithography. The result is a smooth surface of thin layers as small as 16 μm that can be handled immediately after the process is completed. Two different materials are used: One material is used for the actual model, while a second, gel-like resin is used for support, such as these shown in Fig. 20.5. Each material is simultaneously jetted and cured, layer by layer. When the model is completed, the support material is removed with an aqueous solution. Build sizes are fairly large, with an envelope of up to 500 × 400 × 200 mm. These processes have capabilities similar to those of stereolithography and use similar resins (Table 20.2). The main advantages are the capabilities of avoiding part cleanup and lengthy postprocess curing operations, and the much thinner layers produced, thus allowing for better resolution.

EXAMPLE 20.2 Coffeemaker Design

Alessi Corporation is well known for its high-end kitchen products. Although it makes products out of a wide range of materials, it is best known for its highly polished stainless-steel designs. An example is the Cupola coffeemaker, a market favorite that was to be redesigned from the bottom up while preserving the general characteristics of the established design.

Alessi engineers used Multijet modeling to produce prototypes of components of the coffeemaker, as shown in Fig. 20.7. The prototypes allowed engineers to evaluate the ease and security of mechanical assembly, but a significant effort was expended on the design of the coffeemaker's lip in order to optimize the pouring of coffee. A large number of lip prototypes were constructed and evaluated to obtain the most robust and aesthetically pleasing design. The ability to compare physical prototypes to the existing product was deemed essential to evaluating the designs. After a final design was selected from the numerous prototypes produced, it was found that a 5–6-week time savings was achieved in product development. The time savings translated into cost savings, as well as assuring timely market launch of the redesigned product.

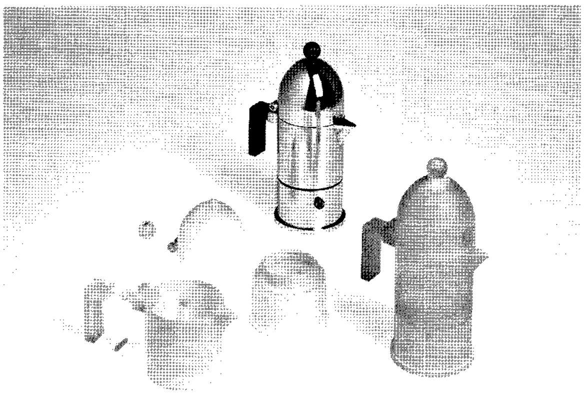


FIGURE 20.7 Coffeemaker prototypes produced through Multijet modeling and final product. *Source*: Courtesy Alessi Corporation and 3D Systems, Inc.

Source: Courtesy Alessi Corporation and 3D Systems, Inc.

20.3.4 Selective Laser Sintering

Selective laser sintering (SLS) is a process based on the sintering of nonmetallic or (less commonly) metallic powders selectively into an individual object. The basic elements in this process are shown in Fig. 20.8. The bottom of the processing

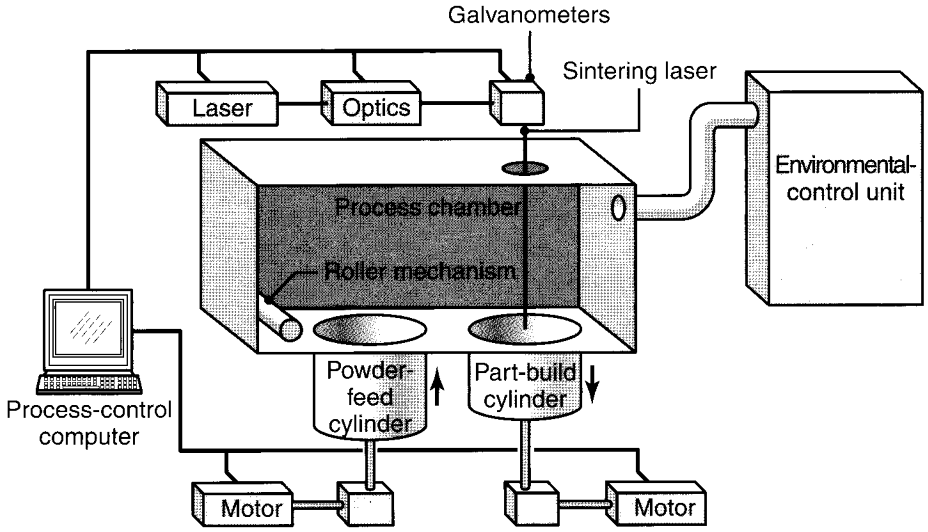


FIGURE 20.8 Schematic illustration of the selective-laser-sintering process.
Source: After C. Deckard and P.F. McClure.

chamber is equipped with two cylinders:

1. A powder-feed cylinder, which is raised incrementally to supply powder to the part-build cylinder through a roller mechanism.
2. A part-build cylinder, which is lowered incrementally as the part is being formed.

First, a thin layer of powder is deposited in the part-build cylinder. 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, yet it supports the sintered portion. Another layer of powder is then deposited; this cycle is repeated again and again until the entire three-dimensional part has been produced. The loose particles are shaken off, and the part is recovered. The part does not require further curing—unless it is a ceramic, which has to be fired to develop strength.

A variety of materials can be used in this process, including polymers (such as ABS, PVC, nylon, polyester, polystyrene, and epoxy), wax, metals, and ceramics with appropriate binders. It is most common to use polymers because of the smaller, less expensive, and less complicated lasers required for sintering. With ceramics and metals, it is common to sinter only a polymer binder that has been blended with the ceramic or metal powders. The resultant part can be carefully sintered in a furnace and infiltrated with another metal if desired.

20.3.5 Electron-beam Melting

A process similar to selective laser sintering and electron-beam welding (see Section 30.6), *electron-beam melting* 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; the part build size is limited to around 200 × 200 × 180 mm. Electron-beam melting (EBM) is up to 95% efficient from an energy standpoint (compared with 10–20% efficiency for selective laser sintering), so that the titanium powder is actually melted and fully dense parts can be produced. A volume build rate of up to 60 cm^3/hr can be obtained, with individual layer thicknesses of 0.050–0.200 mm. Hot isostatic pressing (Section 17.3.2) also can be performed on parts to improve their fatigue strength. Although applied mainly to titanium and cobalt-chrome to date, the process is being developed for stainless steels, aluminum, and copper alloys.

20.3.6 Three-dimensional Printing

In the *three-dimensional-printing* (3DP) process, a printhead deposits an inorganic binder material onto a layer of polymer, ceramic, or metallic powder, as shown in Fig. 20.9. A piston supporting the powder bed is lowered incrementally, and with each step, a layer is deposited and then fused by the binder.

Multijet modeling and polyjet processes (described in Section 20.3.3) are sometimes referred to as three-dimensional printing approaches, because they operate in a similar fashion to ink-jet printers but incorporate a third (thickness) direction. However, three-dimensional printing is most commonly associated with printing a binder onto powder.

Three-dimensional printing allows considerable flexibility in the materials and binders used. Commonly used powder materials are blends of polymers and fibers, foundry sand, and metals. Furthermore, since multiple binder printheads can be incorporated into one machine, it is possible to produce full-color prototypes by having different-color binders (see Example 20.3). The effect is a three-dimensional analog to printing photographs using three ink colors on an ink-jet printer.

A common part produced by 3DP from ceramic powder is a ceramic-casting shell (see Section 11.2.4), in which an aluminum-oxide or aluminum-silica powder is fused with a silica binder. The molds have to be postprocessed in two steps: (1) curing at around 150°C and (2) firing at 1000° to 1500°C.

The parts produced through the 3DP process are somewhat porous and therefore may lack strength. Three-dimensional printing of metal powders can also be

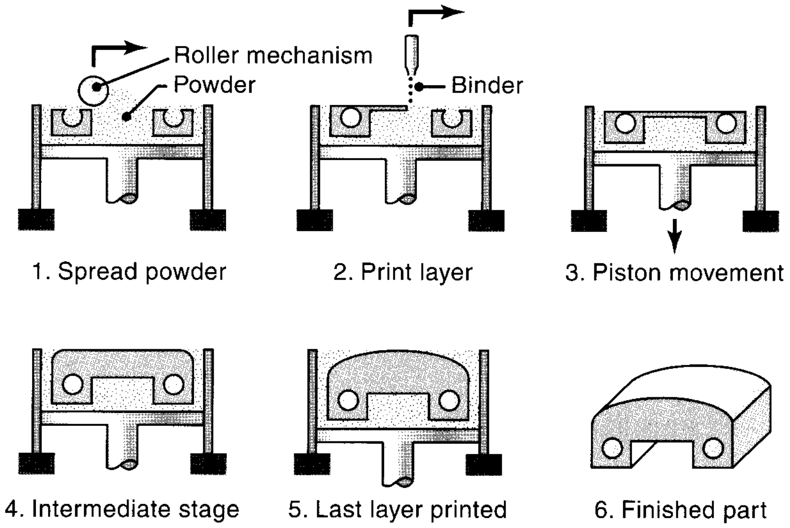


FIGURE 20.9 Schematic illustration of the three-dimensional-printing process. *Source*: After E. Sachs and M. Cima.

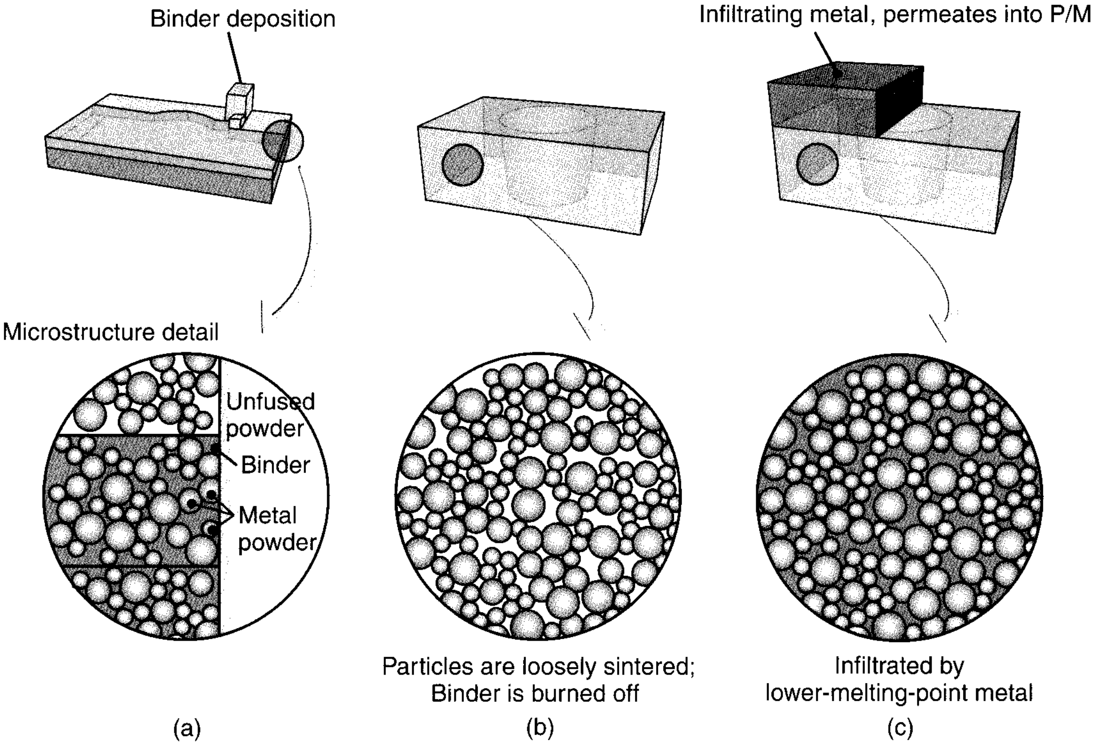


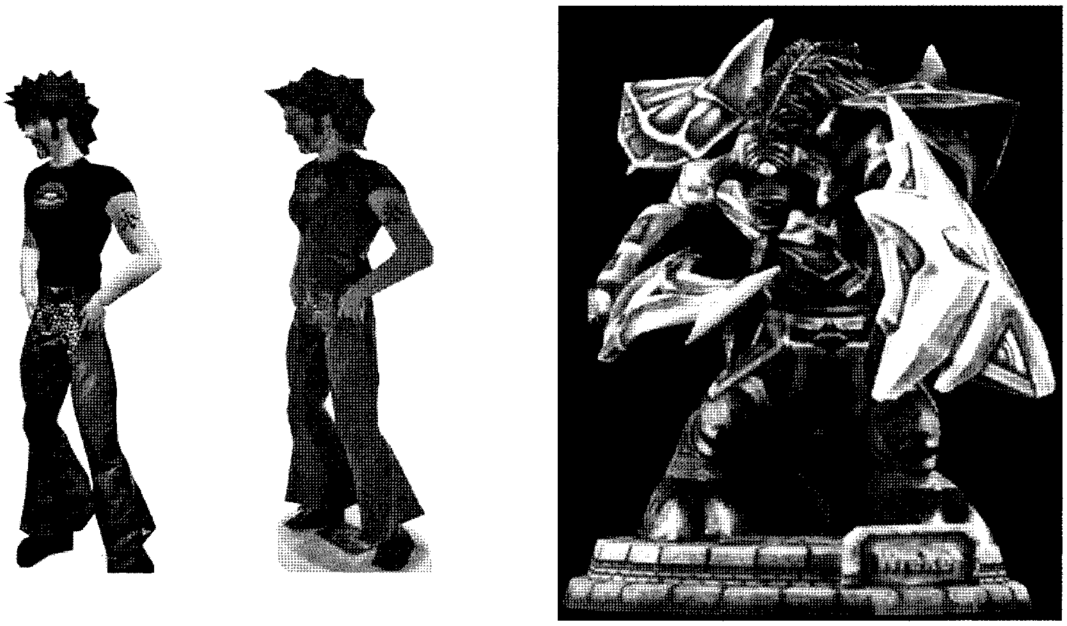
FIGURE 20.10 Three-dimensional printing using (a) part-build, (b) sinter, and (c) infiltration steps to produce metal parts. (d) An example of a bronze-infiltrated stainless-steel part produced through three-dimensional printing. *Source*: Courtesy of Kennametal Extrude Hone.

combined with sintering and metal infiltration (see Section 17.5) to produce fully dense parts, using the sequence shown in Fig. 20.10. Here, the part is produced as before by directing the binder onto powders. However, the build sequence is then followed by sintering to burn off the binder and partially fuse the metal powders, just as in powder injection molding described in Section 17.3.3. Common metals used in 3DP are stainless steels, aluminum, and titanium. 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* (see below).

In a related **ballistic-particle manufacturing** process, a stream of a material (such as plastic, ceramic, metal, or wax) is ejected through a small orifice at a surface (target) using an ink-jet type mechanism. A powder is not involved; the material deposited by the ink-jet mechanism is used to build the prototype. The ink-jet head is guided by a three-axis robot to produce three-dimensional prototypes.

EXAMPLE 20.3 Production of Second Life® Avatars

Second Life® and World of Warcraft® are examples of virtual worlds accessed through a website and are enjoyed by millions of people worldwide. To participate, users create an "avatar" that depicts their alter ego in the fictional world. Many modern computer games (such as Rock Band 2) also allow users to produce very detailed avatars, with a unique appearance and unique personalities. Avatars contain



(a) (b)

FIGURE 20.11 Rapid-prototyped versions of user-defined characters, or avatars, produced from geometric descriptions within popular websites or games. (a) Second Life® avatar, as appears on a computer screen (left) and after printing (right); (b) an avatar known as "Wreker" from World of Warcraft®. *Source*: Courtesy Z Corporation, Figure Prints and Fabjectory, Inc.

three-dimensional geometry data that describes their appearance, which can be translated to a file format suitable for rapid prototyping.

Avatars can be printed in full color to a 150-mm high figurine with Z-Corp Spectrum Z510 or ZPrinter 450 three-dimensional printers (Fig 20.11). Users can order their avatar prototypes on the web, which are then printed and shipped to the user within days.

EXAMPLE 20.4 Fuselage Fitting for Helicopters

Sikorsky Aircraft Company needed to produce a limited number of the fuselage fittings shown in Fig. 20.12a. Sikorsky wanted to produce the forging dies by means of three-dimensional-printing technologies. A die was designed using the CAD part description. Forging allowances were incorporated and flashing accommodated by the die design.

The dies were printed using a three-dimensional printer produced by ProMetal and are shown in Fig. 20.12b. The dies were made by producing 0.178-mm layers with stainless-steel powder as the workpiece media. The total time spent in the 3DP machine was just under 45 hours. This was followed by curing of the binder (10 hours, plus 5 hours for cooldown), sintering (40 hours, plus 17 hours for cooldown), and infiltration (27 hours, plus 15 hours for cooldown). The dies then were finished and positioned in a die holder, and the part was forged in an 800-ton hydraulic press with a die temperature of around 300°C. An as-forged part is shown in Fig. 20.12c and requires trimming of the flash before it can be used. The dies were produced in just over six days—compared with the many months required for conventional die production.

Source: Courtesy of Kennametal Extrude Hone.

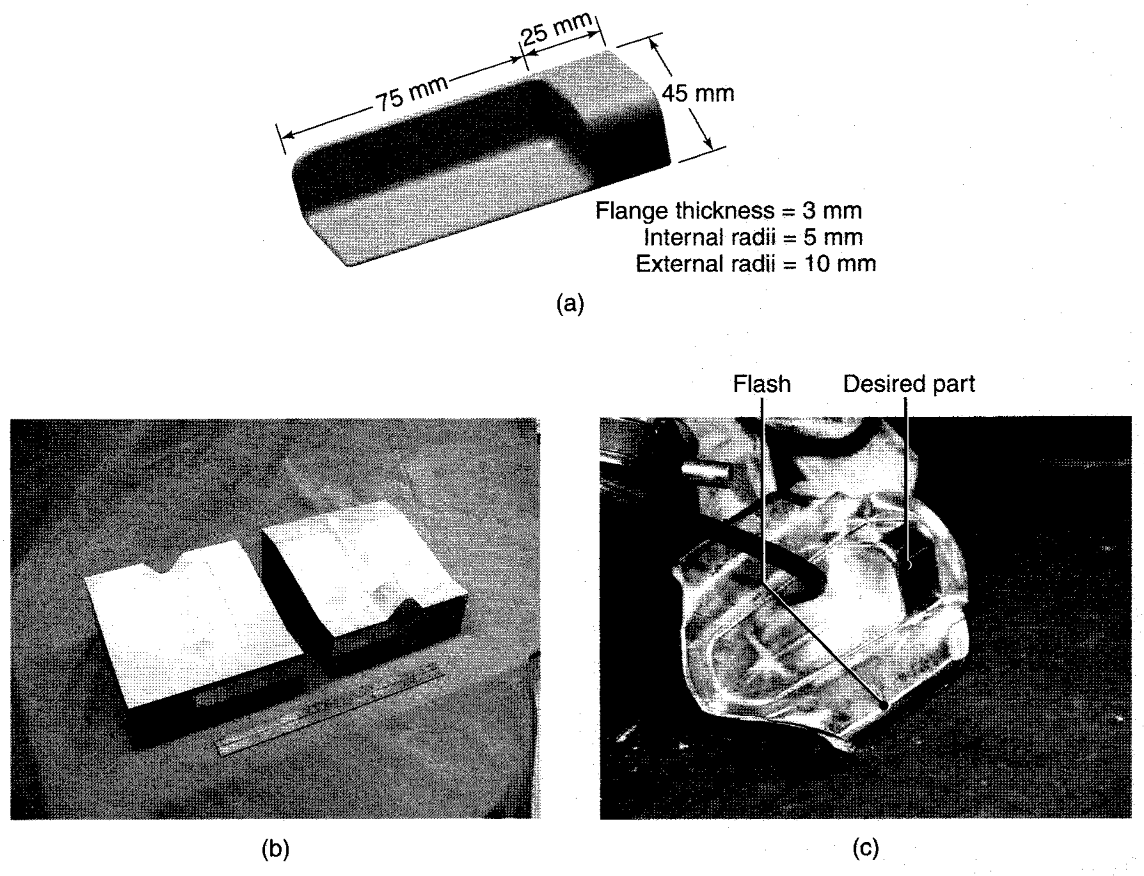


FIGURE 20.12 A fitting required for a helicopter fuselage. (a) CAD representation with added dimensions. (b) Dies produced by three-dimensional printing. (c) Final forged workpiece. *Source*: Courtesy of Kennametal Extrude Hone.

20.3.7 Laminated-object Manufacturing

Lamination implies a laying down of layers that are bonded adhesively to one another. Several variations of *laminated-object manufacturing* (LOM) are available. The simplest and least expensive versions of LOM involve using control software and vinyl cutters to produce the prototype. Vinyl cutters are simple CNC machines that cut shapes from vinyl or paper sheets. Each sheet then has a number of layers and registration holes, which allow proper alignment and placement onto a build fixture. Figure 20.13 illustrates the manufacture of a prototype by laminated-object manufacturing with manual assembly. Such LOM systems are highly economical and are popular in schools and universities because of the hands-on demonstration of additive manufacturing and production of parts by layers.

LOM systems can be elaborate; the more advanced systems use layers of paper or plastic with a heat-activated glue on one side to produce parts. The desired shapes are burned into the sheet with a laser, and the parts are built layer by layer (Fig. 20.14). On some systems, the excess material must be removed manually once the part is completed. 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 (with squares printed on it, similar to graph paper).

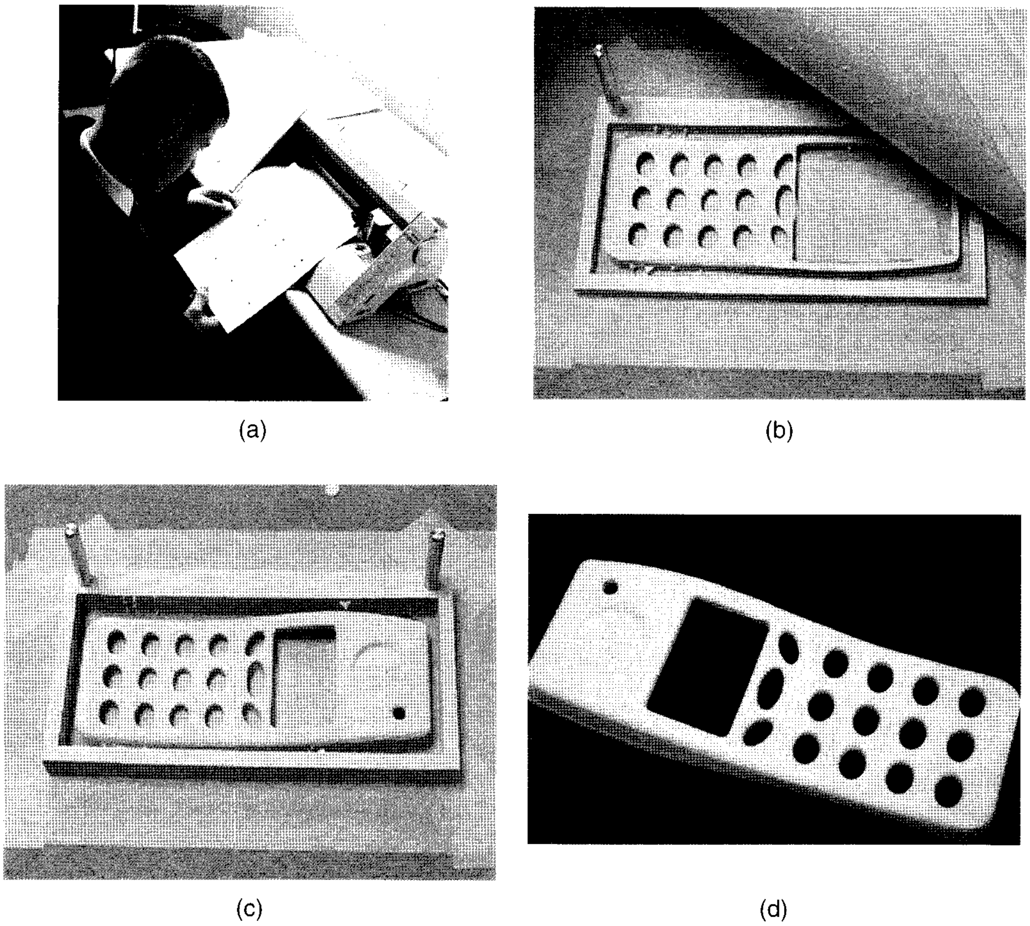


FIGURE 20.13 Production of a prototype through laminated-object manufacturing. (a) Layers are obtained from a vinyl cutter; (b) layers are manually stacked to form the part; (c) completed laminated assembly; (d) final part prototype. *Source*: Courtesy P. Barraclough, Boxford Ltd.

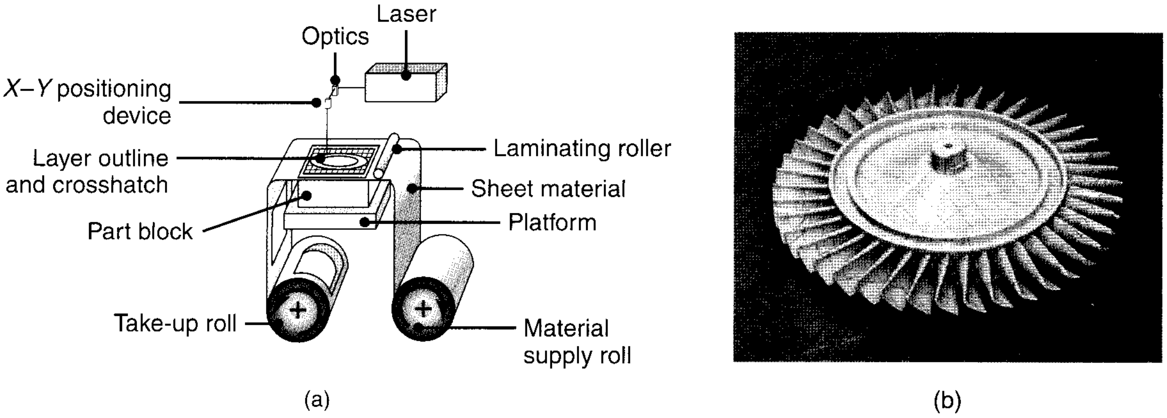


FIGURE 20.14 (a) Schematic illustration of the laminated-object-manufacturing process. (b) Turbine prototype made by LOM. *Source*: Courtesy of M. Feygin, Cubic Technologies, Inc.

20.3.8 Solid-ground Curing

This process is unique in that entire slices of a part are manufactured at one time. As a result, a large throughput is achieved, compared with that from other rapid-prototyping processes. However, *solid-ground curing* (SGC) is among the most expensive processes; hence, its adoption has been much less common than that of other types of rapid prototyping, and new machines are not available.

Basically, the method consists of the following steps:

1. Once a slice is created by the computer software, a mask of the slice is printed on a glass sheet by an electrostatic printing process similar to that used in laser printers. A mask is required because the area of the slice where the solid material is desired remains transparent.
2. While the mask is being prepared, a thin layer of photoreactive polymer is deposited on the work surface and is spread evenly.
3. The photomask is placed over the work surface, and an ultraviolet floodlight is projected through the mask. Wherever the mask is clear, the light shines through to cure the polymer and causes the desired slice to be hardened.
4. The unaffected resin (still liquid) is vacuumed off the surface.
5. Water-soluble liquid wax is spread across the work area, filling the cavities previously occupied by the unexposed liquid polymer. Since the workpiece is on a chilling plate and the workspace remains cool, the wax hardens quickly.
6. The layer is then milled to achieve the correct thickness and flatness.
7. This process is repeated—layer by layer—until the part is completed.

Solid-ground curing has the advantage of a high production rate, because entire slices are produced at once and two glass screens are used concurrently. That is, while one mask is being used to expose the polymer, the next mask already is being prepared, and it is ready as soon as the milling operation is completed.

20.3.9 Laser-engineered Net Shaping

More recent developments in additive manufacturing processes involve the principle of using a laser beam to melt and deposit metal powder or wire—again, layer by layer—over a previously molten layer. The patterns of deposited layers are controlled by a CAD file. This near-net-shaping process is called *laser-engineered net shaping* (LENS, a trade name) and is based on the technologies of laser welding and cladding. The heat input and cooling are controlled precisely to obtain a favorable microstructure.

The deposition process is carried out inside a closed area 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. Also, it can be used for repairing thin and delicate components. There are other, similar processing methods using lasers, including *controlled-metal buildup* (CMB) and *precision-metal deposition* (PMD, a trade name).

20.4 Virtual Prototyping

Virtual prototyping is a purely software form of prototyping that uses advanced graphics and virtual-reality environments to allow designers to examine a part. In a way, this technology is used by common, conventional CAD packages to render a part so that the designer can observe and evaluate it as it is drawn. However, virtual-prototyping systems should be recognized as extreme cases of rendering detail.

The simplest forms of such systems use complex software and three-dimensional graphics routines to allow viewers to change the view of the parts on a computer screen. More complicated versions will use virtual-reality headgear and gloves with appropriate sensors to let the user observe a computer-generated prototype of the desired part in a completely virtual environment.

Virtual prototyping has the advantage of affording an instantaneous rendering of parts for evaluation, but the more advanced systems are costly. Because familiarity with software interfaces is a prerequisite to their application, these systems have very steep learning curves. Furthermore, many manufacturing and design practitioners prefer a physical prototype to evaluate, rather than a video-screen rendering. They often perceive virtual-reality prototypes to be inferior to mechanical prototypes, even though designers debug as many or more errors in the virtual environment.

There have been some important examples of complicated products produced without any physical prototype whatsoever (**paperless design**). Perhaps the best known example is the Boeing 777 aircraft, for which mechanical fits and interferences were evaluated on a CAD system and difficulties were corrected before the first production model was manufactured (see Section 38.5).

20.5 Direct Manufacturing and Rapid Tooling

While extremely beneficial as a demonstration and visualization tool, rapid-prototyping processes also have been used as a manufacturing step in production. There are two basic methodologies used:

1. Direct production of engineering metal, ceramic, and polymer components or parts by rapid prototyping.
2. Production of tooling or patterns by rapid prototyping for use in various manufacturing operations.

Not only are the polymer parts that can be obtained from various rapid-prototyping operations useful for design evaluation and troubleshooting, but occasionally these processes can be used to manufacture parts directly—referred to as *direct manufacturing*. 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 rapid prototyping for direct manufacturing, or *rapid manufacturing*, are as follows:

- 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 small enough to justify direct manufacturing through rapid-prototyping technologies.
- The long-term and consistent performance of rapidly manufactured parts (compared with the more traditional methods of manufacturing them) may be suspect, especially with respect to fatigue, wear, and life cycle.

Much progress is being made to address these concerns to make rapid manufacturing a more competitive and viable option in manufacturing. The future of these processes remains challenging and promising, especially in view of the fact that rapid manufacturing is now being regarded as a method of producing a product on demand. Customers will be able to order a particular part, which will be produced within a relatively short waiting time.

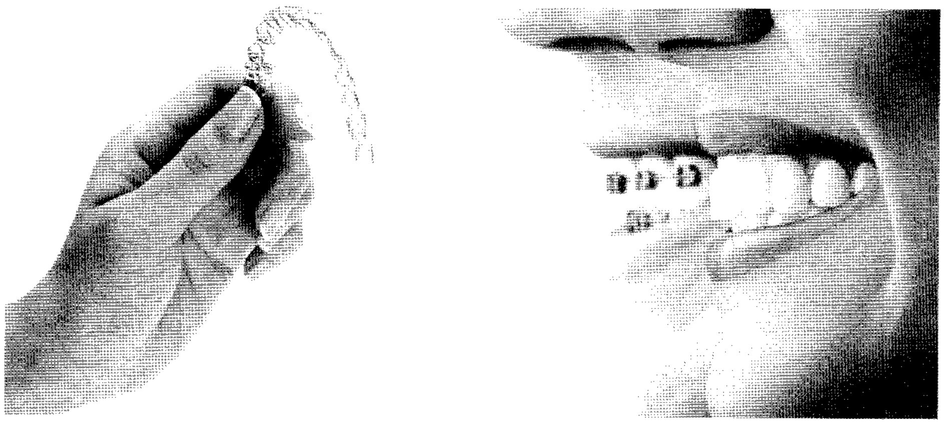
CASE STUDY 20.1 Invisalign® Orthodontic Aligners

Orthodontic braces have been available to straighten teeth for more than 50 years. The braces involve metal, ceramic, or plastic brackets that are bonded adhesively to teeth with fixtures for attachment to a wire, which then forces compliance on the teeth and straightens them to the desired shape within a few years. Conventional orthodontic braces are a well-known and successful technique for ensuring long-term dental health. However, there are several drawbacks to conventional braces, including the facts that (a) they are aesthetically unappealing; (b) the sharp wires and brackets can be painful; (c) they trap food leading to premature tooth decay; (d) brushing and flossing teeth are far more difficult and less effective with braces in place; and (e) certain foods must be avoided because they will damage the braces.

One solution is the Invisalign system, made by Align Technology, Inc. It consists of a series of aligners, each of which the person wears for approximately two weeks. Each aligner (see Fig. 20.15) consists of a precise geometry that incrementally moves the teeth to the desired positions. Because the aligners can be removed for eating, brushing, and flossing, most of the drawbacks of conventional braces are eliminated. Furthermore, since they are produced from a transparent plastic, the aligners do not seriously affect the person's appearance.

The Invisalign product uses a combination of advanced technologies in the production process, shown in Fig. 20.16. The treatment begins with an orthodontist or a general dentist creating a polymer impression of the patient's teeth (Fig 20.16a). These impressions then are used to create a three-dimensional CAD representation of the patient's teeth, as shown in Fig. 20.16b. Proprietary computer-aided design software then assists in the development of a treatment strategy for moving the teeth in an optimal manner.

Once the treating doctor has approved the treatment plan and it has been developed, the computer-based information is used to produce the aligners. This is done through a novel application of stereolithography. Although a number of materials are available for stereolithography, they have a characteristic yellow-brown shade to them and therefore are unsuitable for direct application as an orthodontic product. Instead, the Align process uses a stereolithography machine that produces patterns of the desired incremental positions of the teeth (Fig. 20.16c). A sheet of clear polymer is then thermoformed (see Section 19.6) over these patterns to produce the aligners, which are sent to the treating orthodontist. With the doctor's supervision, patients are instructed to change the next set of aligners every two weeks.



(a) (b)

FIGURE 20.15 (a) An aligner for orthodontic use, manufactured by a combination of rapid tooling and thermoforming. (b) Comparison of conventional orthodontic braces with the use of transparent aligners. *Source*: Courtesy of Align Technology, Inc.

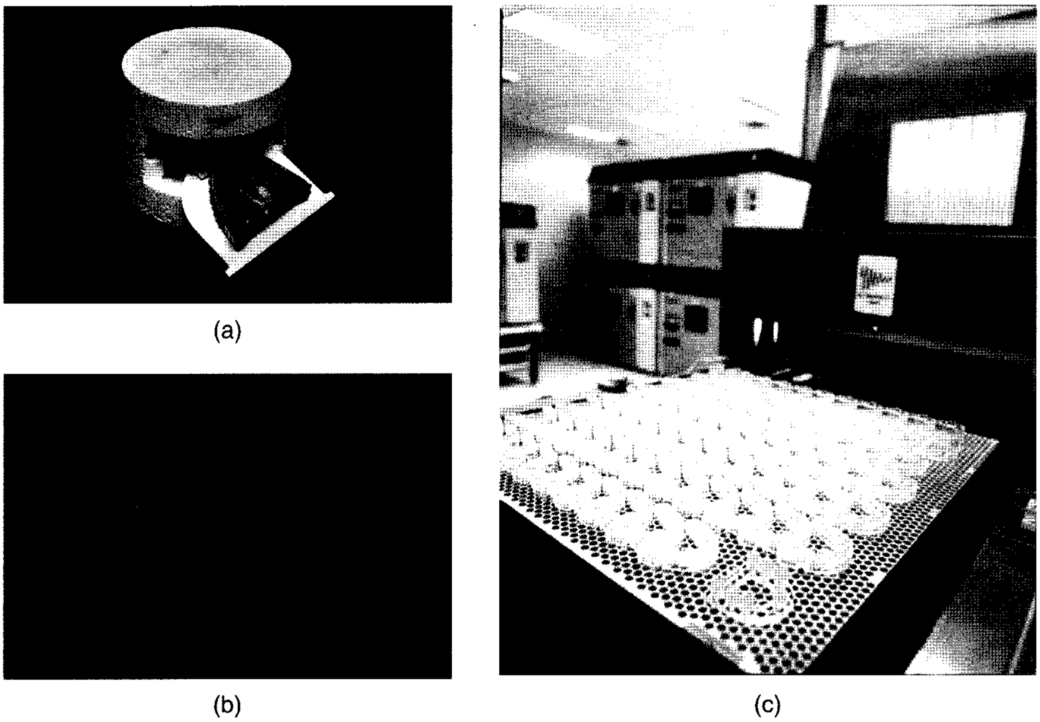


FIGURE 20.16 The manufacturing sequence for Invisalign orthodontic aligners. (a) Creation of a polymer impression of the patient's teeth. (b) Computer modeling to produce CAD representations of desired tooth profiles. (c) Production of incremental models of desired tooth movement. An aligner is produced by thermoforming a transparent plastic sheet against this model.
Source: Courtesy of Align Technology, Inc.

The Invisalign product has proven to be very popular for patients who wish to promote dental health and to preserve their teeth long into their lives. The use of stereolithography to produce accurate tools quickly and inexpensively allows this orthodontic treatment to be economically viable.

Source: Courtesy of Align Technology, Inc.

20.5.1 Rapid Tooling

Several methods have been devised for the rapid production of tooling (RT) by means of rapid-prototyping processes. The advantages to rapid tooling include the following:

1. The high cost of labor and short supply of skilled patternmakers can be overcome.
2. There is a major reduction in lead time.
3. Hollow designs can be adopted easily so that lightweight castings can be produced more easily.
4. The integral use of CAD technologies allows the use of modular dies with base-mold tooling (match plates) and specially fabricated inserts. This modular technique can further reduce tooling costs.
5. Chill- and cooling-channel placement in molds can be optimized more easily, leading to reduced cycle times.

6. Shrinkage due to solidification or thermal contraction can be compensated for automatically through software to produce tooling of the proper size and, in turn, to produce the desired parts.

The main shortcoming of rapid tooling is the potentially reduced tool or pattern life (compared to those obtained from machined tool and die materials, such as tool steels or tungsten carbides).

The simplest method of applying rapid-prototyping operations to other manufacturing processes is in the direct production of patterns or molds. As an example, Fig. 20.17 shows an approach for investment casting. Here, the individual patterns are made in a rapid-prototyping operation (in this case, stereolithography) and then used as patterns in assembling a tree for investment casting. Note that this approach requires a polymer that will completely melt and burn from the ceramic mold; such polymers are available for all forms of polymer rapid-prototyping operations. Furthermore, as drawn in CAD programs, the parts are usually software modified to account for shrinkage, and it is then that the modified part is produced in the rapid-prototyping machinery.

As another example, 3DP 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

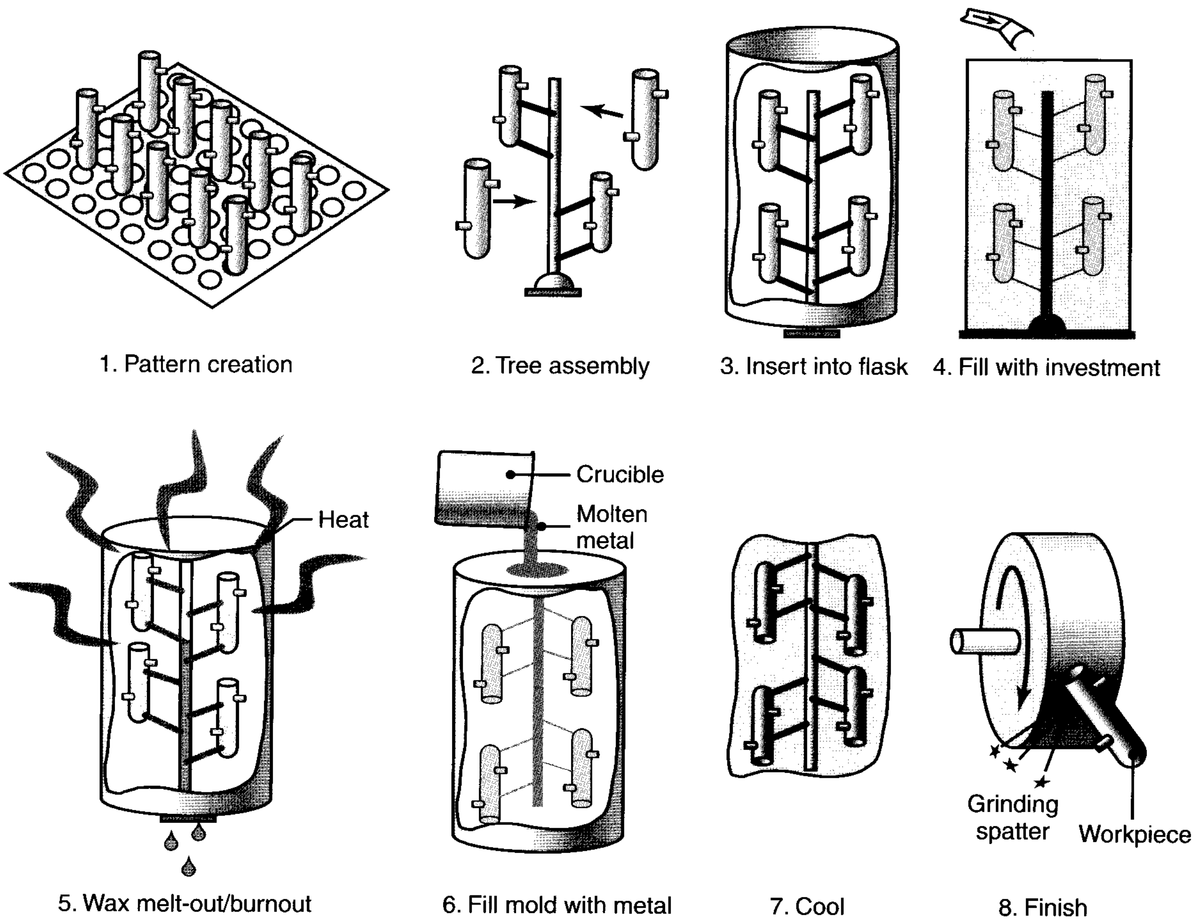


FIGURE 20.17 Manufacturing steps for investment casting with rapid-prototyped wax parts as blanks. This method uses a flask for the investment, but a shell method also can be used. *Source*: Courtesy of 3D Systems, Inc.

aluminum-silica powder is fused with a silica binder. The molds have to be post-processed in two steps: curing at around 150°C and then firing at 1000°–1500°C.

Another common application of rapid tooling is injection molding (see Section 19.3), in which the mold or, more typically, a *mold insert* is manufactured by rapid prototyping. Molds for slip casting of ceramics (see Section 18.2.1) also can be produced in this manner. To produce individual molds, rapid-prototyping processes are used directly, but 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. Furthermore, the design is simplified, because the designer need only analyze 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 rapid-prototyping technology to tool or pattern production, other rapid-tooling approaches based on rapid-prototyping technologies have been developed.

Room-temperature vulcanizing (RTV) **molding/urethane casting** can be performed by preparing a pattern of a part by any rapid-prototyping operation. The pattern is coated with a parting agent and may or may not be modified to define mold parting lines. Liquid RTV rubber is 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 (see Section 19.3.1). One main limitation of this approach is a lesser mold life, because the polyurethane in the mold causes progressive damage and the mold may be suitable for as few as 25 parts.

Epoxy or *aluminum-filled epoxy molds* also can be produced, but mold design then requires special care. With RTV rubber, the mold flexibility allows it to be peeled off the cured part. With epoxy molds, the high stiffness precludes this method of part removal, and mold design is more complicated. Thus, drafts are needed, 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 rapid prototyping (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 in injection molding, mold life may be as few as 10 parts, although a few hundred parts per mold are possible.

Sprayed-metal tooling. In this process, shown in Fig. 20.18, a pattern is created through rapid prototyping. A metal spray operation (see 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; two such mold halves are then suitable for use in injection-molding operations. Mold life is highly dependent on the material and temperatures used, and can vary from a few to thousands of parts.

Keltool process. In the *Keltool process*, an RTV mold is 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

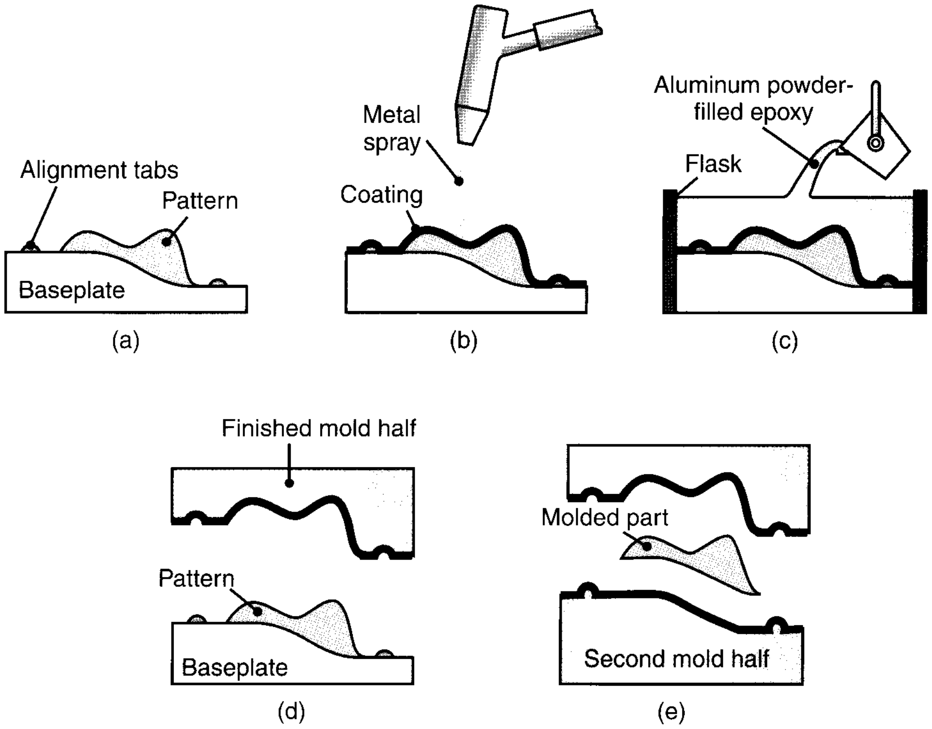


FIGURE 20.18 Production of tooling for injection molding by the sprayed-metal tooling process. (a) A pattern and baseplate are prepared through a rapid-prototyping operation; (b) a zinc–aluminum alloy is sprayed onto the pattern (see Section 34.5); (c) the coated baseplate and pattern assembly are placed 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.

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 attain a superior surface finish and good dimensional tolerances. Keltool molds are limited in size to around 150 × 150 × 150 mm, so, typically, a mold insert suitable for high-volume molding operations is produced. Depending on the material and processing conditions, mold life can range from 100,000 to 10 million parts.

EXAMPLE 20.5 Casting of Plumbing Fixtures

A global manufacturer of plumbing fixtures and accessories for baths and kitchens used rapid tooling to transform its development practice. One of the company's major product lines is decorative water faucets produced from brass castings that are subsequently polished to achieve the desired surface finish. The ability to produce prototypes from brass is essential to quickly evaluate designs and identify processing complications that may result.

A new faucet design was prepared in a CAD program; the finished product is shown in Fig. 20.19. As part of the product development cycle, it was desired 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 desired to validate the design through a sand-casting process followed by polishing. This approach allowed evaluation of the cast parts in terms of porosity and other casting defects, and also would identify processing difficulties that might arise in the finishing stages.

A sand mold was produced first, as shown in Fig. 20.20. The mold material was a blend of foundry

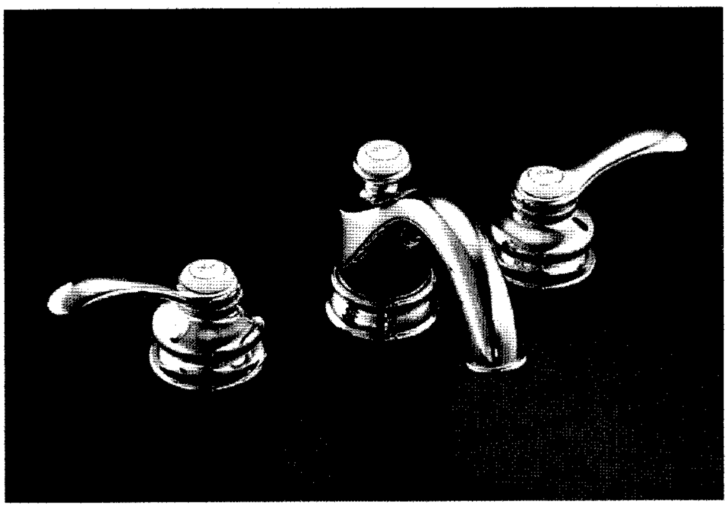


FIGURE 20.19 A new faucet design, produced by casting from rapid-prototyped sand molds.

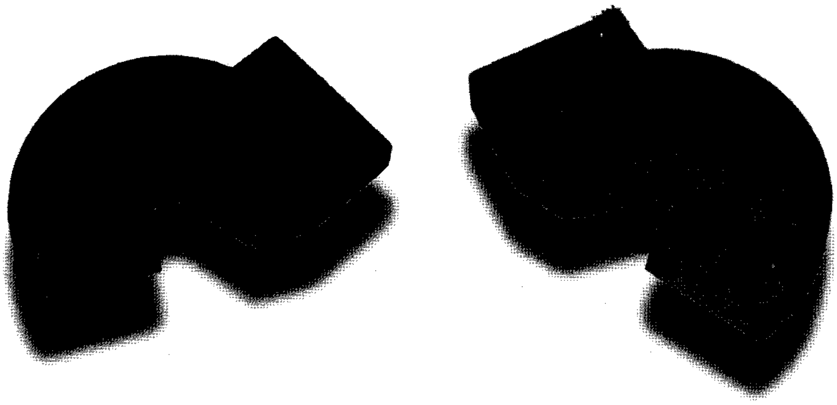


FIGURE 20.20 Sand molds produced through three-dimensional printing.

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 removed from the mold, especially for complicated casting designs. Therefore, the core is produced separately and 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, but would cost over $10,000 and add several months to the lead time. The incorporation of 3D printing into the design process provided new capabilities that confirmed the design aesthetics and function, as well as manufacturing robustness and reliability.

Source: Courtesy of Z Corporation.

SUMMARY

- Rapid prototyping has grown into a unique manufacturing discipline within the past two decades. As a physical-model-producing technology, it is a useful technique for identifying and correcting design errors. Several techniques have been developed for producing parts through rapid prototyping.
- 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 it 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 ballistic particle manufacturing 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.
- Rapid-prototyping 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, rapid-prototyping 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 even forging), thereby significantly reducing the lead time between design and manufacture.

KEY TERMS

ACES
Additive processes
Ballistic-particle manufacturing
Desktop machines
Direct AIM
Direct manufacturing
Electron-beam melting
Free-form fabrication
Fused-deposition modeling
Keltool
Laminated-object manufacturing
Multijet modeling
Photopolymer
Polyjet
Prototype
Rapid tooling
RTV molding/urethane casting
Selective laser sintering
Solid-ground curing
Sprayed metal tooling
Stereolithography
Subtractive processes
Three-dimensional printing
Virtual prototyping

BIBLIOGRAPHY

Beaman, J.J., Barlow, J.W., Bourell, D.L., and Crawford, R., **Solid Freeform Fabrication**, Kluwer, 1997.

Bennett, G. (ed.), **Developments in Rapid Prototyping and Tooling**, Institution of Mechanical Engineers, 1997.

Chua, C.K., and Fua, L.K., **Rapid Prototyping: Principles and Applications in Manufacturing**, Wiley, 1997.

Gebhardt, A., **Rapid Prototyping**, Hanser Gardner, 2004.

Hilton, P.D., and Jacobs, P.F., **Rapid Tooling: Technologies and Industrial Applications**, CRC Press, 2000.

Kamrani, A., and Nasr, E.A. (eds.), **Rapid Prototyping: Theory and Practice**, Springer, 2007.

Noorani, R.I., **Rapid Prototyping: Principles and Applications**, Wiley, 2006.

Pham, D.T., and Dimov, S.S., **Rapid Tooling: The Technologies and Applications of Rapid Prototyping and Rapid Tooling**, Springer-Verlag, 2001.

Wood, L., **Rapid Automated Prototyping: An Introduction**, Industrial Press, 1993.

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) SGC, (e) FDM, and (f) LENS.

20.8. What starting materials can be used in fused-deposition modeling? In three-dimensional printing?

20.9. What are the cleaning and finishing operations in rapid-prototyping processes? Why are they necessary?

QUALITATIVE PROBLEMS

20.10. 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.

20.11. How would you rapidly manufacture tooling for injection molding? Explain any difficulties that may be encountered.

20.12. Explain the significance of rapid tooling in manufacturing.

20.13. List the processes described in this chapter that are best suited for the production of ceramic parts. Explain.

20.14. Few parts in commercial products today are directly manufactured through rapid-prototyping operations. Explain.

20.15. Can rapid-prototyped parts be made of paper? Explain.

20.16. 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.

20.17. Why are the metal parts in three-dimensional printing often infiltrated by another metal?

20.18. Make a list of the advantages and limitations of each of the rapid-prototyping operations described in this chapter.

20.19. In making a prototype of a toy automobile, list the post-rapid-prototyping finishing operations that you think would be necessary. Explain.

QUANTITATIVE PROBLEMS

▶**20.20.** Using an approximate cost of \$160 per litre for the liquid polymer, estimate the material cost of a rapid-prototyped rendering of a typical computer mouse.

▶**20.21.** The extruder head in a fused-deposition modeling setup has a diameter of 1.25 mm and produces layers that are 0.25 mm thick. If the extruder head and polymer extrudate velocities are both 50 mm/s, estimate the production time for the generation of a 38-mm solid cube. Assume that there is a 10-second delay between layers as the extruder head is moved over a wire brush for cleaning.

▶**20.22.** Using the data for Problem 20.21 and assuming that the porosity for the support material is 50%, calculate the production rate for making a 100-mm high cup with an outside diameter of 90 mm and a wall thickness of 4 mm. Consider the cases (a) with the closed end up and (b) with the closed end down.

20.23. 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

20.24. Rapid-prototyping machines represent a large capital investment; consequently, few companies can 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 rapid-prototyping machines that they use, and determine the percentage use of each class.

20.25. One of the major advantages of stereolithography is that it can use semitransparent polymers, so that internal details of parts can readily be discerned. List and describe several parts in which this feature is valuable.

20.26. 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?

20.27. Identify the rapid-prototyping 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.

20.28. Design a machine that uses rapid-prototyping technologies to produce ice sculptures. Describe its basic features, commenting on the effect of size and shape complexity on your design.

20.29. 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.

PART IV

Machining Processes and Machine Tools

Parts manufactured by the casting, forming, and shaping processes described in Parts II and III, including many parts made by near-net or net-shape methods, often require further operations before the product is ready for use. Consider, for example, the following features on parts and whether they could be produced by the processes discussed thus far:

- Smooth and shiny surfaces, such as the bearing surfaces of the crankshaft shown in Fig. IV.1.
- Small-diameter deep holes in a part such as the injector nozzle shown in Fig. IV.2.
- Parts with sharp features, a threaded section, and 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 mechanical assembly with other components.
- Special surface finishes and textures for functional purposes or for appearance.

A brief review will indicate that none of the forming and shaping processes described thus far is capable of producing parts with such specific characteristics and that the parts will require further manufacturing operations. **Machining** is a general term describing a group of processes that consist of the *removal* of material and *modification* of the surfaces of a workpiece after it has been produced by various methods. Thus, machining involves *secondary* and *finishing* operations.

The very wide variety of shapes produced by machining can be seen clearly in an automobile, as shown in Fig. IV.4. It also should be recognized that some parts may be produced to final shape (net shape) and at high quantities by forming and shaping processes, such as die casting and powder metallurgy. However, machining may be more economical, provided that the number of parts required is relatively small or the material and shape allow the parts to be machined at high rates and quantities and with high dimensional accuracy. A good example is the production of brass screw-machine parts on multiple-spindle automatic screw machines.

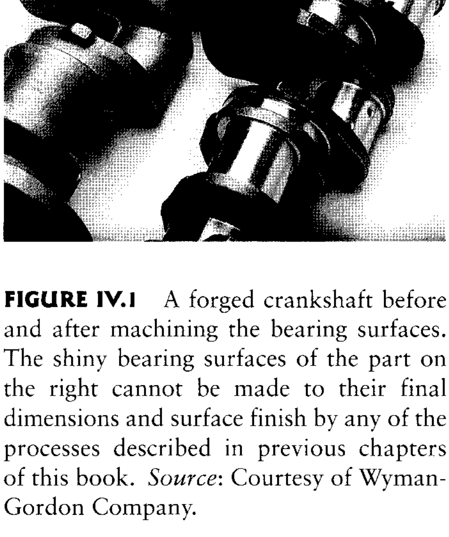


FIGURE IV.1 A forged crankshaft before and after machining the bearing surfaces. The shiny bearing surfaces of the part on the right cannot be made to their final dimensions and surface finish by any of the processes described in previous chapters of this book. *Source*: Courtesy of Wyman-Gordon Company.

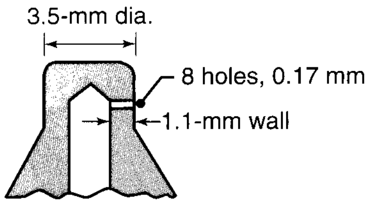


FIGURE IV.2 Cross section of a fuel-injection nozzle, showing a small hole made by the electrical-discharge machining process, as described in Section 27.5. The material is heat-treated steel.

In general, however, resorting to machining suggests that a part could not have been produced to the final desired specifications by the primary processes used in making them and that additional operations are necessary. We again emphasize the importance of *net-shape manufacturing*, as described in Section I.5, to avoid these additional steps and reduce production costs.

Furthermore, in spite of their advantages, material-removal processes have the following disadvantages:

- They *waste material* (although the amount may be relatively small).
- The processes generally takes *longer* than other processes.
- 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.

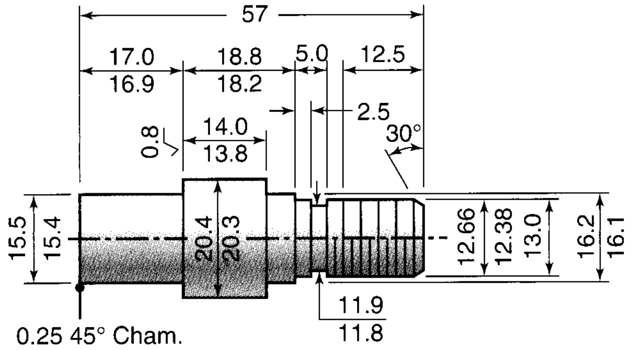


FIGURE IV.3 A machined and threaded part, showing various dimensions and tolerances; all dimensions are in mm. Note that some tolerances are only a few tenths of a mm.

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 related processes (Chapter 26).
- **Advanced machining processes** utilizing electrical, chemical, laser, thermal, and hydrodynamic methods to accomplish this task (Chapter 27).

The machines on which these operations are performed are called **machine tools**. As described throughout Part IV, their construction and characteristics greatly influence these operations, as well as product quality, surface finish, and dimensional accuracy.

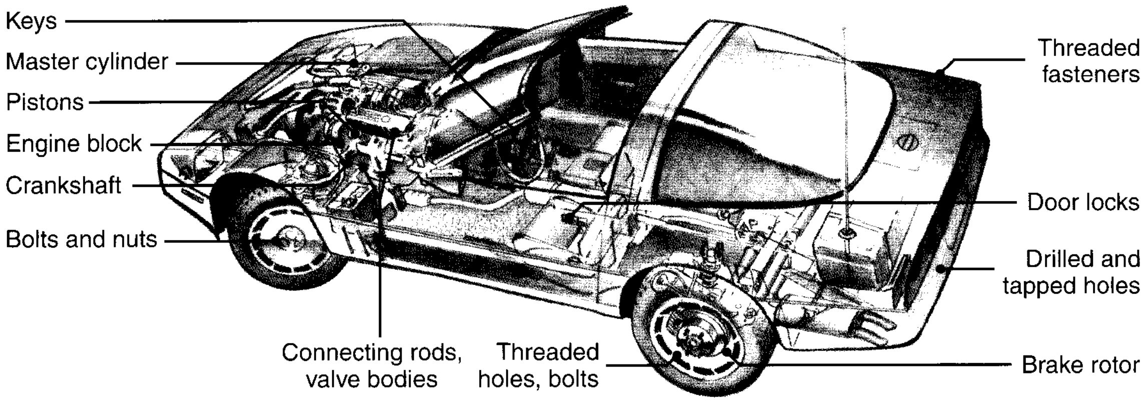


FIGURE IV.4 Typical parts on an automobile that require machining operations to impart desirable surface characteristics, dimensions, and tolerances.

As can be seen in Table I.2 in the General Introduction, the first primitive tools (dating back many millennia) were made for the main purpose of chipping away and cutting all types of natural materials (such as wood, stone, vegetation, and livestock) for the purpose of food and shelter. Note also that it was in the 1500s that progress began on manufacturing products by machining operations, particularly with the introduction of lathes. We now have available a wide variety of computer-controlled machine tools and modern techniques (using various materials and energy sources) and are capable of making functional parts as small as tiny insects and with cross sections smaller than a human hair.

As in other manufacturing operations, it is important to view machining operations as a *system,* consisting of the

- *Workpiece*
- *Cutting tool*
- *Machine tool,* and
- *Production personnel.*

Machining cannot be carried out efficiently or economically and also meet stringent part specifications without a thorough knowledge of the interactions among these four elements.

In the next seven chapters, the basic mechanics of chip formation in machining are described: tool forces, power, temperature, tool wear, surface finish and integrity, cutting tools, and cutting fluids. We then discuss specific machining processes—their capabilities, limitations, and typical applications—and identify important machine-tool characteristics for operations such as turning, milling, boring, drilling, and tapping. The features of *machining centers*—versatile machine tools controlled by computers and capable of efficiently performing a variety of operations—also are described.

The next processes described are those in which the removal of material (to a very high dimensional accuracy and surface finish) is carried out by **abrasive processes** and related operations. Common examples are grinding, wheels, coated abrasives, honing, lapping, buffing, polishing, shot-blasting, and ultrasonic machining.

For technical and economic reasons, some parts cannot be machined satisfactorily by cutting or abrasive processes. Since the 1940s, important developments have taken place in **advanced machining processes**, such as chemical, electrochemical, electrical-discharge, laser-beam, electron-beam, abrasive-jet, and hydrodynamic machining.

The knowledge gained in Part IV will enable us to assess the capabilities and limitations of material-removal processes; machine tools and related equipment; their proper selection for maximum efficiency, productivity, and low cost; and how these processes fit into the broader scheme of manufacturing operations.

CHAPTER

21

Fundamentals of Machining

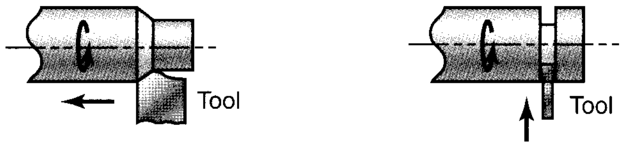
- This chapter is an introduction to the fundamentals of the machining processes to be covered in subsequent chapters and, as such, presents the basic concepts relevant to all machining operations.
- The chapter opens with a discussion of the mechanics of chip formation in machining and the model typically used for orthogonal cutting operations; the model and its oblique-cutting extensions allow the calculation of force and power in machining.
- Temperature rise in chip and cutting tool are then discussed.
- Mechanisms of tool wear and failure follow, with flank wear characterized by the Taylor tool-life equation.
- Crater wear, nose wear, and other forms of wear are also described.
- The chapter ends with a discussion of surface finish, the integrity of parts produced by machining, and the factors involved in the machinability of various metallic and nonmetallic materials.

21.1 Introduction

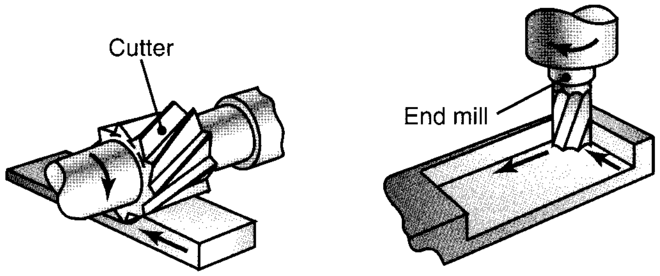
Cutting processes remove material from the surface of a workpiece by producing **chips**. Some of the more common cutting processes, illustrated in Fig. 21.1 (see also Fig. I.6e), are as follows:

- **Turning**, in which the workpiece is rotated and a cutting tool removes a layer of material as the tool moves to the left, as in Fig. 21.1a.
- **Cutting off**, in which the cutting tool moves radially inward and separates the right piece from the bulk of the blank.
- **Slab milling**, in which a rotating cutting tool removes a layer of material from the surface of the workpiece.
- **End milling**, in which a rotating cutter travels along a certain depth in the workpiece and produces a cavity.

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 velocity as the workpiece rotates. The *feed*, or *feed rate*, is the distance the tool travels



(a) Straight turning (b) Cutting off



(c) Slab milling (d) End milling

FIGURE 21.1 Some examples of common machining operations.

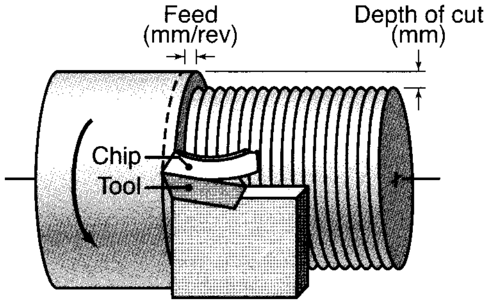


FIGURE 21.2 Schematic illustration of the turning operation, showing various features.

horizontally per unit revolution of the workpiece (mm/rev). This movement of the tool produces a chip, which moves up the face of the tool.

In order to analyze this process in 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. A chip is produced ahead of the tool by plastically deforming and shearing the material continuously along the shear plane. This phenomenon can be demonstrated by slowly scraping the surface of a stick of butter lengthwise with a sharp knife and observing the formation of a chip. 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 (dimension perpendicular to the page) in the idealized model. These 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.

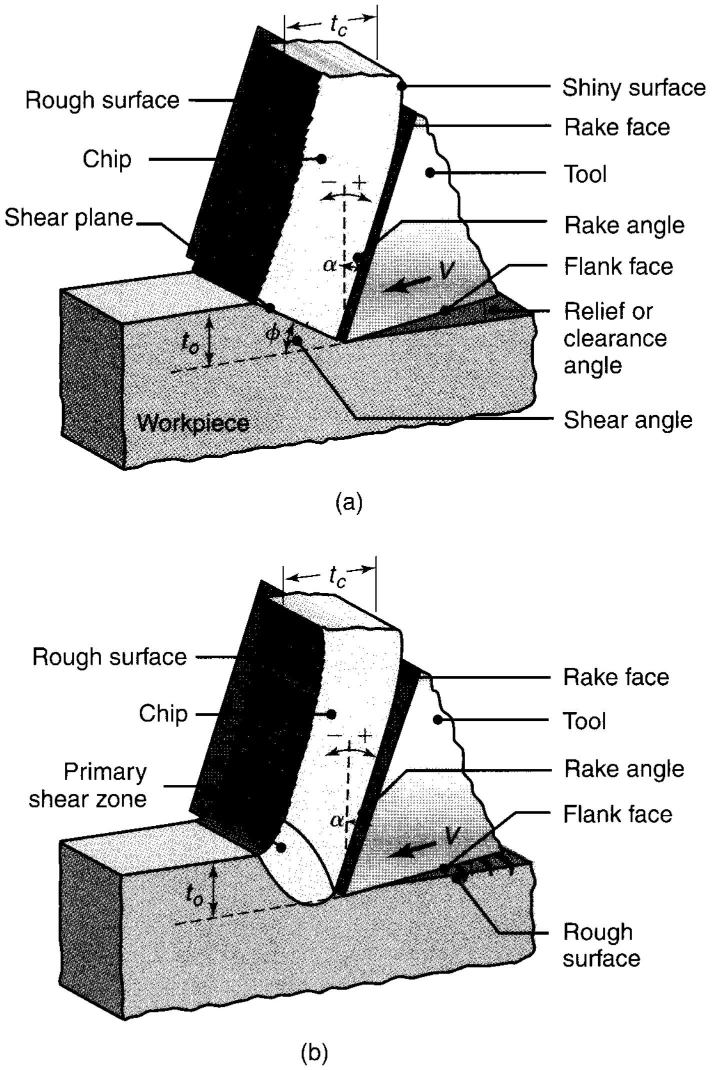


FIGURE 21.3 Schematic illustration of a two-dimensional cutting process, also called orthogonal cutting: (a) Orthogonal cutting with a well-defined shear plane, also known as the M.E. Merchant model. Note that the tool shape, the depth of cut, t_o, and the cutting speed, V, are all independent variables. (b) Orthogonal cutting without a well-defined shear plane.

21.2 Mechanics of Cutting

The factors that influence the cutting process are outlined in Table 21.1. In order to appreciate the contents of this table, let's now identify the major *independent variables* in the cutting process: (a) tool material and coatings; (b) tool shape, surface finish, and sharpness; (c) workpiece material and condition; (d) cutting speed, feed, and depth of cut; (e) cutting fluids; (f) characteristics of the machine tool; and (g) work holding and fixturing.

Dependent variables in cutting are those that are influenced by changes in the independent variables listed above, and 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.

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 automated machinery
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

When machining operations yield unacceptable results, normal troubleshooting requires a systematic investigation. A typical question posed is which of the independent variables should be changed first, and to what extent, if (a) the surface finish of the workpiece being cut is unacceptable, (b) the cutting tool wears rapidly and becomes dull, (c) the workpiece becomes very hot, and (d) the tool begins to vibrate and chatter.

In order to understand these phenomena and respond to the question posed, let's first study the mechanics of chip formation—a subject that has been studied extensively since the early 1940s. Several models (with varying degrees of complexity) have been proposed. As is being done in other manufacturing processes (such as casting, molding, shaping, and forming), advanced machining models are being continuously developed. The methods used include *computer simulation* of the machining process, with the purpose of studying the complex interactions of the many variables involved while developing capabilities to *optimize* machining operations.

The simple model shown in Fig. 21.3a (referred to as the M.E. Merchant model, developed in the early 1940s) is sufficient for our purposes. This model is known as **orthogonal cutting**, because it is two dimensional and the forces involved (as we later show) are perpendicular to each other. The cutting tool has a **rake angle** of α (positive, as shown in the figure) and a **relief** or **clearance angle**.

Microscopic examination of chips obtained in actual machining operations have revealed that they are produced by *shearing* (as modeled in Fig. 21.4a)—similar to the movement in a deck of cards sliding against each other. Shearing takes place in a **shear zone** (usually along a well-defined plane referred to as the **shear plane**) at an angle ϕ (called the **shear angle**). Below the shear plane, the workpiece remains undeformed; above it, the chip that is already formed moves up the rake face of the tool. The dimension d in the figure is highly exaggerated to show the mechanism involved. In reality, this dimension is only on the order of 10^{-2} to 10^{-3} mm.

Some materials (notably cast irons at low speeds) do not shear along a well-defined plane but instead shear in a zone, as shown in Fig. 21.3b. Shearing in such a

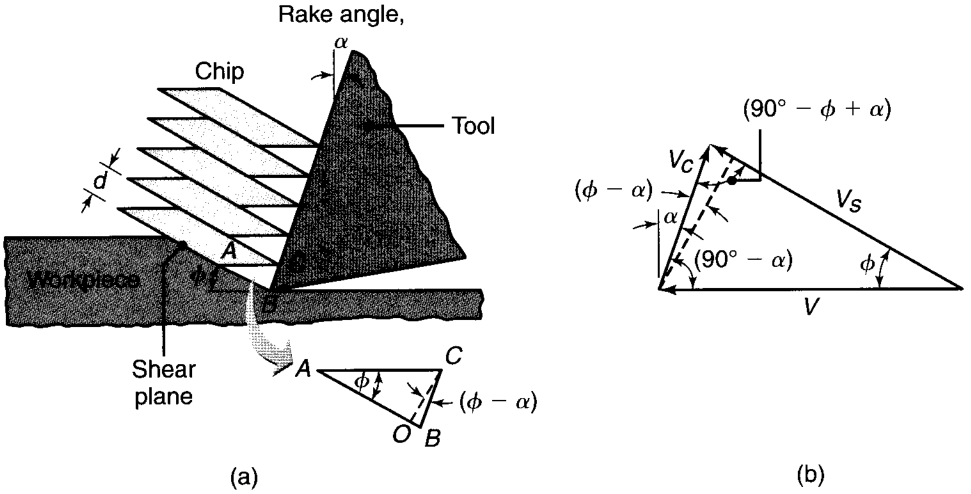


FIGURE 21.4 (a) Schematic illustration of the basic mechanism of chip formation by shearing. (b) Velocity diagram showing angular relationships among the three speeds in the cutting zone.

volume is not in itself objectionable, but it can lead to surface defects in the workpiece (as will be discussed later).

Cutting Ratio. It can be seen 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, and is related to the two angles by the following relationships:

$$\tan\phi = \frac{r\cos\alpha}{1 - r\sin\alpha} \tag{21.1a}$$

and

$$r = \frac{t_o}{t_c} = \frac{\sin\phi}{\cos(\phi - \alpha)}. \tag{21.1b}$$

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* and is a measure of how thick the chip has become compared with the depth of cut; hence, the chip-compression ratio always is greater than unity. The depth of cut also is referred to as the *undeformed chip thickness*, as may be visualized by reviewing Fig. 21.3.

The cutting ratio is an important and useful parameter for evaluating cutting conditions. Since the undeformed chip thickness, t_o, is a machine setting and is therefore known, the cutting ratio can be calculated easily by measuring the chip thickness with a micrometer. With the rake angle also known for a particular cutting operation (it is a function of the tool and workpiece geometry in use), Eq. (21.1) allows calculation of the shear angle.

Although we have referred to t_o as the *depth of cut*, note that in a machining process such as turning, as shown in Fig. 21.2, this quantity is the *feed*. To visualize the situation, assume, for instance, that the workpiece in Fig. 21.2 is a thin-walled tube and the width of the cut is the same as the thickness of the tube. Then, by rotating Fig. 21.3 clockwise by 90°, note that it is now similar to the view in Fig. 21.2.

Shear Strain. Referring now to Fig. 21.4a, we can see that the **shear strain,** γ, that the material undergoes can be expressed as

$$\gamma = \frac{AB}{OC} = \frac{AO}{OC} + \frac{OB}{OC},$$

or

$$\gamma = \cot\phi + \tan(\phi - \alpha). \qquad (21.2)$$

Note that large shear strains are associated with low shear angles or with low or negative rake angles. Shear strains of 5 or higher have been observed in actual cutting operations. Compared to forming and shaping processes, the workpiece material undergoes greater deformation during cutting, as is also seen in Table 2.4. Furthermore, deformation in cutting 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 is high. (We discuss the nature and size of the deformation zone further in Section 21.3.)

The shear angle has great significance in the mechanics of cutting operations. It influences force and power requirements, chip thickness, and temperature. Consequently, much attention has been focused on determining the relationships among shear angle, cutting process variables, and workpiece material properties. One of the earliest analyses was based on the assumption that the shear angle adjusts itself to minimize the cutting force or that the shear plane is a plane of maximum shear stress. This analysis yielded the expression

$$\phi = 45^\circ + \frac{\alpha}{2} - \frac{\beta}{2}, \qquad (21.3)$$

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 the many shear-angle relationships developed, another useful formula that generally is applicable is

$$\phi = 45^\circ + \alpha - \beta. \qquad (21.4)$$

The coefficient of friction in metal cutting generally ranges from about 0.5 to 2, indicating that the chip encounters considerable frictional resistance while moving up the tool's rake face. 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 (21.3) indicates that (a) as the rake angle decreases or the friction at the tool–chip interface (rake face) increases, the shear angle decreases and the chip becomes thicker; (b) thicker chips mean more energy dissipation because the shear strain is higher [see Eq. (21.2)]; and (c) because work done during cutting is converted into heat, the temperature rise is also higher. The effects of these phenomena are described throughout the rest of this chapter.

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. Since mass continuity has to be maintained,

$$Vt_o = V_c t_c, \quad \text{or} \quad V_c = Vr.$$

Hence,

$$V_c = \frac{V \sin\phi}{\cos(\phi - \alpha)}. \qquad (21.5)$$

A *velocity diagram* also can be constructed, as shown in Fig. 21.4b, in which, from trigonometric relationships, we obtain the equation

$$\frac{V}{\cos(\phi - \alpha)} = \frac{V_s}{\cos\alpha} = \frac{V_c}{\sin\phi} \tag{21.6a}$$

where V_s is the velocity at which shearing takes place in the shear plane. Note also that

$$r = \frac{t_o}{t_c} = \frac{V_c}{V}. \tag{21.6b}$$

These velocity relationships will be utilized further in Section 21.3 in describing power requirements in cutting 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 as follows:

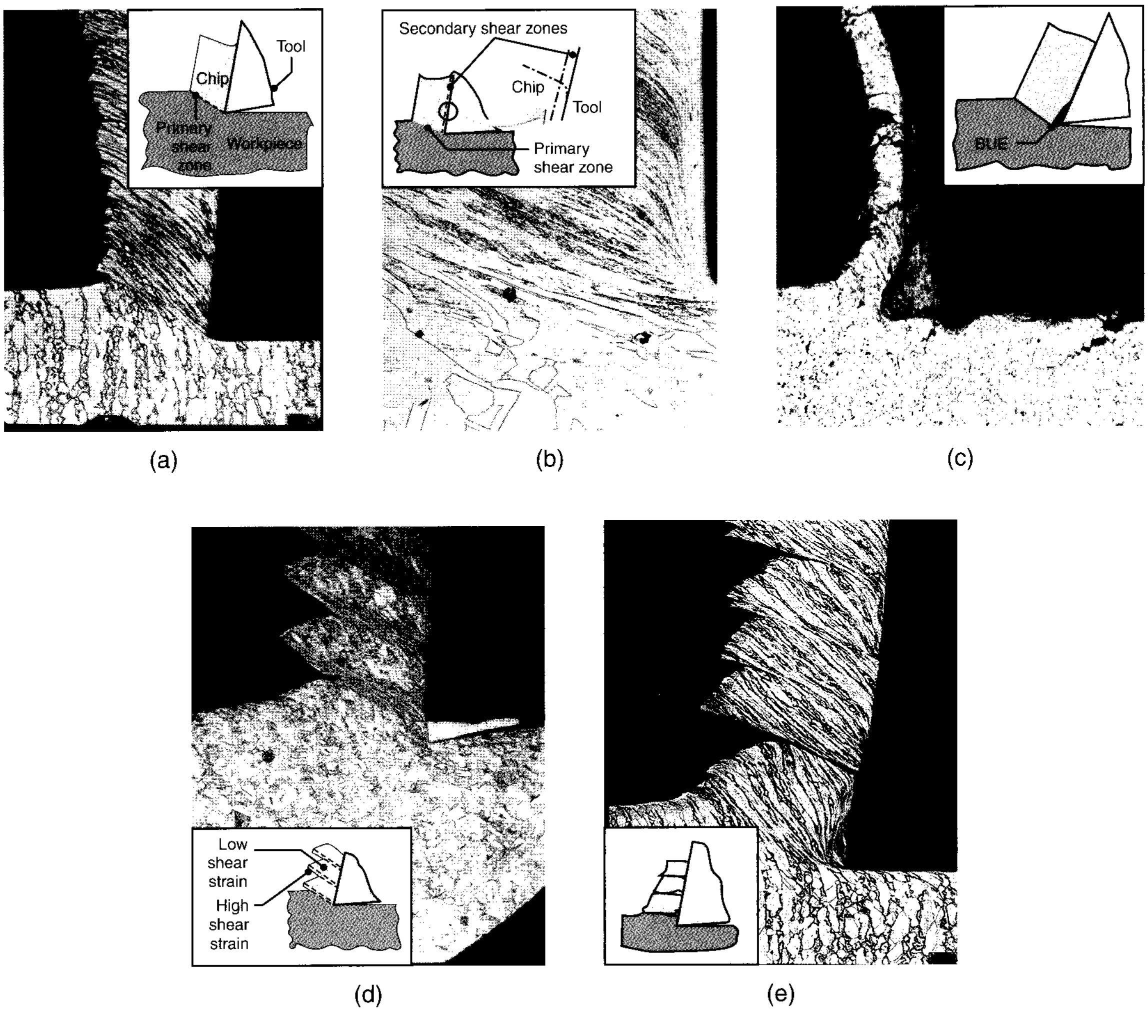


FIGURE 21.5 Basic types of chips produced in orthogonal metal cutting, their schematic representation, and photomicrographs of the cutting zone: (a) continuous chip with narrow, straight, and primary shear zone; (b) continuous chip with secondary shear zone at the chip–tool interface; (c) built-up edge; (d) segmented or nonhomogeneous chip; and (e) discontinuous chip. *Source*: After M.C. Shaw, P.K. Wright, and S. Kalpakjian.

- Continuous
- Built-up edge
- Serrated or segmented
- Discontinuous.

Let's first note that a chip has two surfaces. One surface has been in contact with the rake face of the tool and has a shiny and burnished appearance caused by rubbing as the chip moves up the tool face. The other surface is from the original surface of the workpiece; it has a jagged, rough 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. This surface is exposed to the environment and has not come into any contact with any solid body.

Continuous Chips. *Continuous chips* usually are formed with ductile materials that are machined at high cutting speeds and/or high rake angles (Fig. 21.5a). The 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 thicker 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 in the machined subsurface. This situation occurs generally in machining soft metals at low speeds and low rake angles. It usually results in a poor surface finish and induces surface residual stresses, which may be detrimental to the properties of the machined part in their service life.

Although they generally produce a good surface finish, continuous chips are not necessarily desirable, particularly with computer-controlled machine tools in wide use, as they tend to become tangled around the toolholder, the fixturing, and the workpiece, as well as around the chip-disposal systems (see Section 23.3.7). The operation may have to be stopped to clear away the chips. This problem can be alleviated with **chip breakers** (discussed shortly), by changing parameters such as cutting speed, feed, and depth of cut, or by using 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, the BUE becomes unstable and eventually breaks apart. Part of the BUE material is carried away by the tool side of the chip; the rest is deposited randomly on the workpiece surface. The cycle of BUE formation and destruction is repeated continuously during the cutting operation until corrective measures are taken. In effect, a built-up edge changes the geometry of the cutting edge and dulls it (Fig. 21.6a).

Built-up edge commonly is observed in practice. It is a major factor that adversely affects surface finish, as can be seen in Figs. 21.5c and 21.6b and c. However, a thin, stable BUE usually is regarded as desirable because it reduces tool wear by protecting its rake face. Cold-worked metals generally have less of a tendency to form BUE than when in their annealed condition. Because of work hardening and deposition of successive layers of material, the BUE hardness increases significantly (Fig. 21.6a). As the cutting speed increases, the size of the BUE decreases; in fact it may not form at all.

The tendency for BUE formation can be reduced by one or more of the following means:

- Increase the cutting speeds
- Decrease the depth of cut
- Increase the rake angle
- Use a sharp tool
- Use an effective cutting fluid
- Use a cutting tool that has lower chemical affinity for the workpiece material.

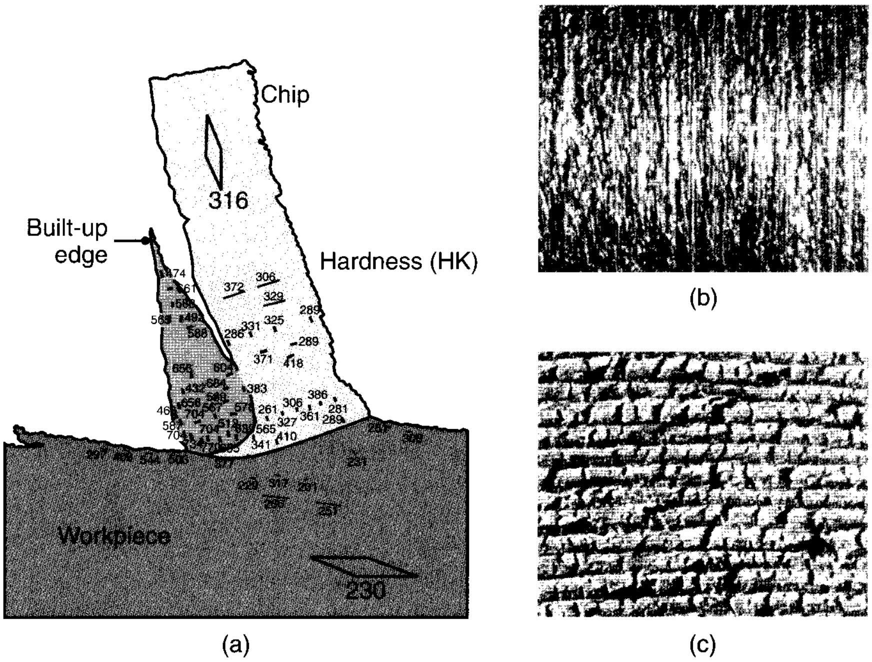


FIGURE 21.6 (a) Hardness distribution with a built-up edge in the cutting zone (material: 3115 steel). Note that some regions in the built-up edge are as much as three times harder than the bulk metal of the workpiece. (b) Surface finish produced in turning 5130 steel with a built-up edge. (c) Surface finish on 1018 steel in face milling. Magnifications: 15×. *Source*: Courtesy of Metcut Research Associates, Inc.

Serrated Chips. *Serrated chips* (also called *segmented* or *nonhomogeneous* chips, see Fig. 21.5d) are semicontinuous chips with large zones of low shear strain and small zones of high shear strain, hence the latter zone is called *shear localization*. Metals with low thermal conductivity and strength that decreases sharply with temperature (thermal softening) exhibit this behavior, most notably titanium. The chips have a sawtooth-like appearance. (This type of chip should not be confused with the illustration in Fig. 21.4a, in which the dimension *d* is highly exaggerated.)

Discontinuous Chips. *Discontinuous chips* consist of segments that may be attached firmly or loosely to each other (Fig. 21.5e). Discontinuous chips usually form under the following conditions:

- Brittle workpiece materials, because they do not have the capacity to undergo the high shear strains involved in cutting.
- Workpiece materials that contain hard inclusions and impurities or have structures such as the graphite flakes in gray cast iron.
- Very low or very high cutting speeds.
- Large depths of cut.
- Low rake angles.
- Lack of an effective cutting fluid.
- Low stiffness of the toolholder or the machine tool, thus allowing vibration and chatter to occur.

Because of the discontinuous nature of chip formation, forces continually vary during cutting. Consequently, the stiffness or rigidity of the cutting-tool holder, the work-holding devices, and the machine tool are important in cutting with serrated chips as well as with discontinuous chips. If it is not sufficiently stiff, the machine tool may begin to vibrate and chatter, as discussed in detail in Section 25.4. This, in turn, adversely affects the surface finish and dimensional accuracy of the machined part and may cause premature wear or damage to the cutting tool—even to the components of the machine tool if the vibration is excessive.

Chip Curl. In all cutting operations performed on metals, as well as nonmetallic materials such as plastics and wood, chips develop a curvature (*chip curl*) as they leave the workpiece surface (Fig. 21.5). Among the factors affecting the chip curl are the following:

- The distribution of stresses in the primary and secondary shear zones.
- Thermal effects.
- Work-hardening characteristics of the workpiece material.
- The geometry of the cutting tool.
- Cutting fluids.

Process variables also affect chip curl. Generally, as the depth of cut decreases, the radius of curvature decreases; that is, the chip becomes curlier. Also, cutting fluids can make chips become more curly (the radius of curvature decreases), thus reducing the tool–chip contact area and concentrating the heat closer to the tip of the tool. As a result, tool wear increases.

Chip Breakers. As stated previously, continuous and long chips are undesirable, as they tend to become entangled, severely interfere with machining operations, and also become a potential safety hazard. If all of the process variables are under control, the usual procedure employed to avoid such a situation is to break the chip intermittently with cutting tools that have *chip-breaker* features, as shown in Fig. 21.7.

Chip breakers, traditionally pieces of metal clamped to the tool's rake face, bend and break the chip. However, most modern 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.

With soft workpiece materials (such as pure aluminum or copper), chip breaking by the means just described may not be effective, in which case machining may be done in small increments (pausing so that a chip is not generated) or by reversing the feed by small increments. In interrupted-cutting operations (such as milling), chip breakers generally are not necessary, since the chips already have finite lengths.

Controlled Contact on Tools. 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. This reduction in contact length affects chip-formation mechanics. Primarily, it reduces the cutting forces and thus the energy and temperature. Determining an optimum length is important, as too small a contact length would concentrate the heat at the tool tip, thus increasing wear.

Cutting Nonmetallic Materials. A variety of chips are encountered in cutting thermoplastics, depending on the type of polymer and process parameters, such as depth

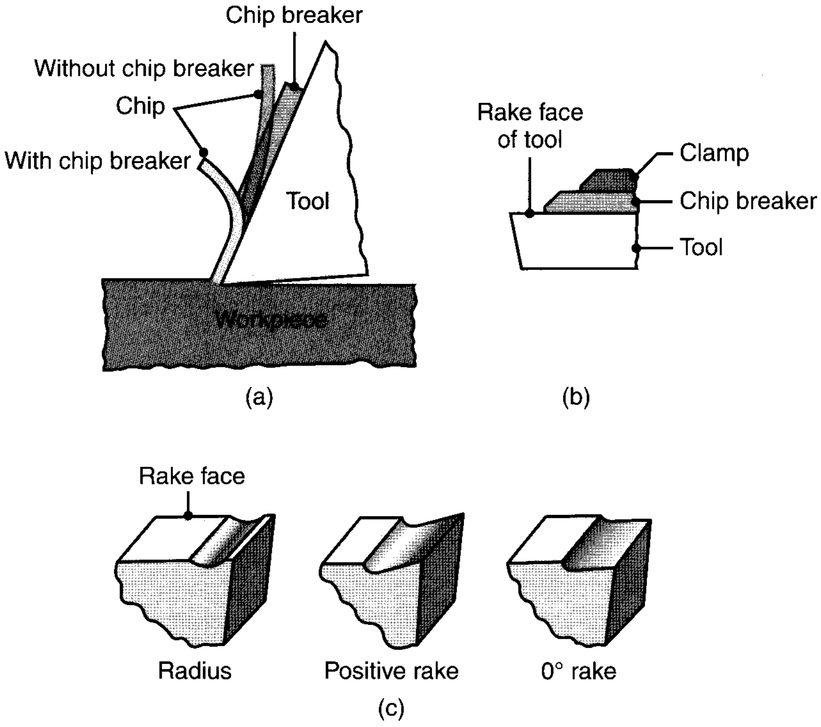


FIGURE 21.7 (a) Schematic illustration of the action of a chip breaker. Note that the chip breaker decreases the radius of curvature of the chip and eventually breaks it. (b) Chip breaker clamped on the rake face of a cutting tool. (c) Grooves in cutting tools acting as chip breakers. Most cutting tools used now are *inserts* with built-in chip-breaker features.

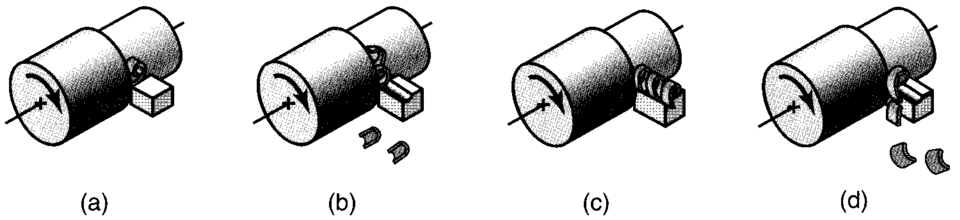


FIGURE 21.8 Chips produced in turning: (a) tightly curled chip; (b) chip hits workpiece and breaks; (c) continuous chip moving radially away from workpiece; and (d) chip hits tool shank and breaks off. *Source*: After G. Boothroyd.

of cut, tool geometry, and cutting speed. Many of the discussions concerning metals also are applicable to polymers. Because they are brittle, thermosetting plastics and ceramics generally produce discontinuous chips. [For characteristics of other machined materials (such as wood, ceramics, and composite materials) see Section 21.7.3].

21.2.2 Oblique Cutting

The majority of machining operations involve tool shapes that are three dimensional; thus, the cutting is *oblique*. The basic difference between oblique and orthogonal cutting can be seen in Fig. 21.9a. Whereas in orthogonal cutting the chip slides directly up the face of the tool, in oblique cutting the chip is *helical* and at an angle *i*, called

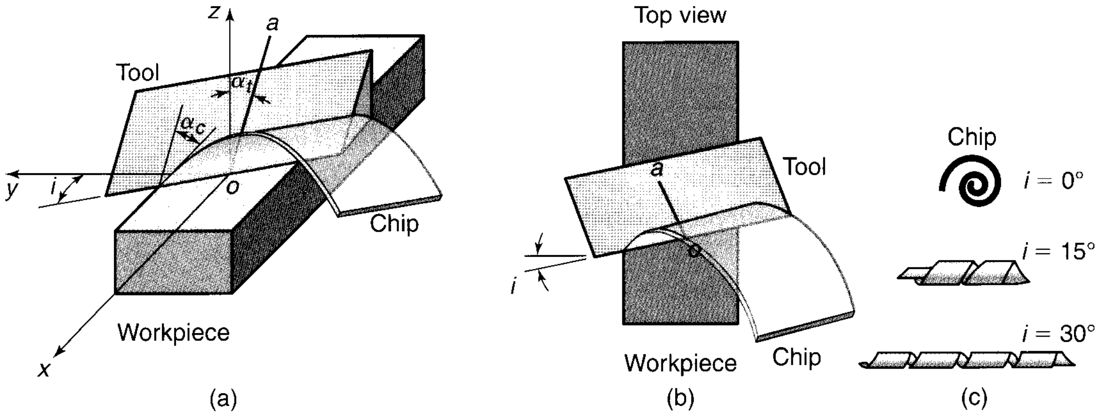


FIGURE 21.9 (a) Schematic illustration of cutting with an oblique tool. Note the direction of chip movement. (b) Top view, showing the inclination angle, *i*. (c) Types of chips produced with tools at increasing inclination angles.

the **inclination angle** (Fig. 21.9b). Note the lateral direction of chip movement in oblique cutting—a situation that is similar to a snowplow blade, that throws the snow sideways. It can be seen that such a helical chip moves sideways and away from the cutting zone and doesn't obstruct it as it would in orthogonal cutting.

Note that the chip in Fig. 21.9a flows up the rake face of the tool at angle α_c (the **chip flow angle**), which is measured in the plane of the tool face. Angle α_i is the **normal rake angle** and is a basic geometric property of the tool. This is the angle between line *oz* normal to the workpiece surface and line *oa* on the tool face.

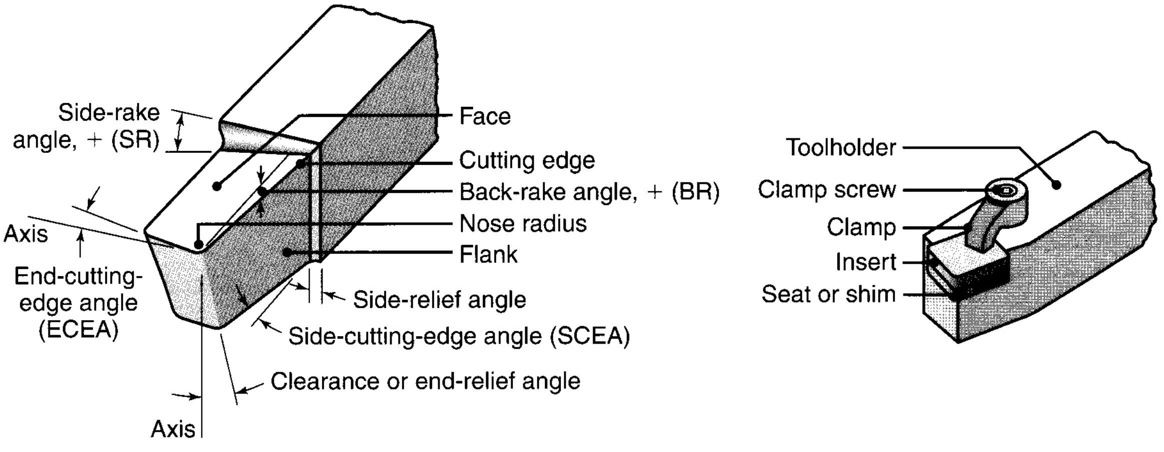
The workpiece material approaches the cutting tool at a velocity V and leaves the surface (as a chip) with a velocity V_c. The effective rake angle, α_e, is calculated in the plane of these two velocities. Assuming that the chip flow angle, α_c, is equal to the inclination angle, i (and this assumption has been verified experimentally), the effective rake angle, α_e, is

$$\alpha_e = \sin^{-1}(\sin^2 i + \cos^2 i \, \sin\alpha_n). \tag{21.7}$$

Since both i and α_n can be measured directly, the effective rake angle can 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 usually can be produced by grinding, the majority of cutting tools are now 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 to shave wood. *Shaving* is useful particularly in improving the surface finish and dimensional accuracy of sheared parts and punched slugs (Fig. 16.9). Another application of shaving is in finishing gears with a cutter that has the shape of the gear tooth (see Section 24.7). Parts that are long or have a combination of shapes are shaved by *skiving* with a specially shaped cutting tool that moves tangentially across the length of the workpiece.



(a) (b)

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 made of carbides and other materials of various shapes and sizes.

21.3 Cutting Forces and Power

Knowledge of the *cutting forces* and *power* involved in machining operations is important for the following reasons:

- Data on cutting forces is essential so that
  - **a.** Machine tools can be properly designed to minimize distortion of the machine components, maintain the desired dimensional accuracy of the machined part, and help select appropriate toolholders and work-holding devices.
  - **b.** The workpiece is capable of withstanding these forces without excessive distortion.
- Power requirements must be known in order to enable the selection of a machine tool with adequate electric power.

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 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. It can also be shown that

$$F = R \sin\beta \tag{21.8a}$$

and

$$N = R \cos\beta. \tag{21.8b}$$

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. It can

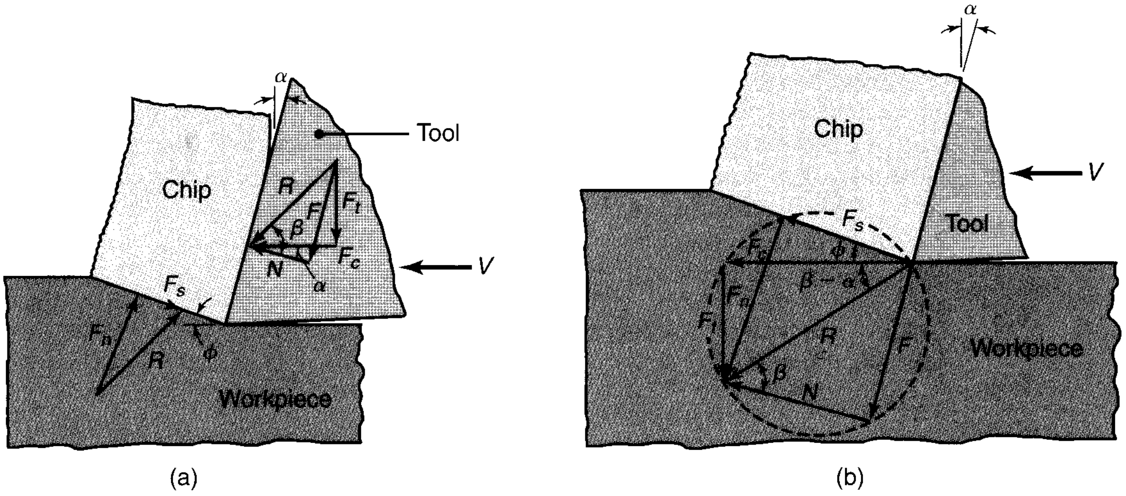


FIGURE 21.11 (a) Forces acting in the cutting zone during two-dimensional cutting. Note that the resultant force, R, must be colinear to balance the forces. (b) Force circle to determine various forces acting in the cutting zone.

be shown that these forces can be expressed, respectively, as

$$F_s = F_c \cos\phi - F_t \sin\phi \tag{21.9}$$

and

$$F_n = F_c \sin\phi + F_t \cos\phi. \tag{21.10}$$

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 be determined.

The ratio of F to N is the **coefficient of friction**, μ, at the tool–chip interface, and the angle β is the **friction angle** (as in Fig. 21.11). The magnitude of μ can be determined as

$$\mu = \frac{F}{N} = \frac{F_t + F_c \tan\alpha}{F_c - F_t \tan\alpha}. \tag{21.11}$$

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 *pressures* on the 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 and, sometimes, chipping and fracture of the tool.

Thrust Force. A knowledge of the *thrust force* in cutting is important because the toolholder, the work-holding devices, and the machine tool 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 be pushed away from the workpiece surface being machined. This movement will, in turn, reduce the depth of cut, resulting in less dimensional accuracy in the machined part.

We also can show the effect of rake angle and friction angle on the direction of thrust force by noting from Fig. 21.11b that

$$F_t = R \sin(\beta - \alpha), \tag{21.12a}$$

or

$$F_t = F_c \tan(\beta - \alpha). \tag{21.12b}$$

Note that 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 relative magnitudes of β and α. When $\beta > \alpha$, the sign of F_t is positive (that is, *downward*), and when $\beta < \alpha$, the sign is negative (that is, *upward*). Therefore, it is 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 work holders and in the stability of the cutting process.

Power. *Power* is the product of force and velocity. Thus, from Fig. 21.11, the power input in cutting is

$$\text{Power} = F_c V. \tag{21.13}$$

This 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, it can be seen that the power dissipated in the shear plane is

$$\text{Power for shearing} = F_s V_s. \tag{21.14}$$

Denoting the width of cut as w, the **specific energy for shearing,** u_s, is given by

$$u_s = \frac{F_s V_s}{w t_o V}. \tag{21.15}$$

Similarly, the power dissipated in friction is

$$\text{Power for friction} = F V_c, \tag{21.16}$$

and the **specific energy for friction,** u_f, is

$$u_f = \frac{F V_c}{w t_o V} = \frac{F r}{w t_o}. \tag{21.17}$$

The **total specific energy,** u_t, is thus

$$u_t = u_s + u_f. \tag{21.18}$$

Because of the many factors involved, reliable prediction of cutting forces and power still is based largely on experimental data, such as given in Table 21.2. The wide range of values shown can be attributed to differences in strength within each material group and to various other factors, such as friction, use of cutting fluids, and processing variables.

The sharpness of the tool tip also influences forces and power. Because the tip rubs against the machined surface and makes the deformation zone ahead of the tool larger, duller tools require higher forces and power.

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. Transducers have a much higher natural frequency and stiffness than dynamometers, which are prone to excessive deflection and vibration. Also, it is possible to calculate the cutting force from the **power consumption** during cutting, provided that the mechanical efficiency of the machine tool is known or can be determined. The *specific energy* in cutting (such as that shown in Table 21.2) also can be used to calculate cutting forces.

TABLE 21.2

Approximate Range of Energy Requirements in Cutting Operations at the Drive Motor of the Machine Tool (for Dull Tools, Multiply by 1.25)

Material	Specific energy $W \cdot s/mm^3$
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

EXAMPLE 21.1 Relative Energies in Cutting

In an orthogonal cutting operation, $t_o = 0.13$ mm, $V = 120$ m/min, $\alpha = 10°$ and the width of cut = 6 mm. It is observed that $t_c = 0.23$ mm, $F_c = 500$ N, and $F_t = 200$ N. 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

$$r = \frac{t_o}{t_c} = \frac{0.13}{0.23} = 0.565,$$

$$F = R\sin\beta,$$

$$F_c = R\cos(\beta - \alpha),$$

and

$$R = \sqrt{(F_t^2 + F_c^2)} = \sqrt{200^2 + 500^2} = 539 \text{ N}.$$

Thus,

$$500 = 539 \cos(\beta - 10),$$

so

$$\beta = 32°$$

and

$$F = 539 \sin 32° = 286 \text{ N}.$$

Hence,

$$\text{Percentage} = \frac{(286)(0.565)}{500} = 0.32, \text{ or } 32\%.$$

21.4 Temperatures in Cutting

As in all metalworking processes where plastic deformation is involved, the energy dissipated in cutting is converted into heat that, in turn, raises the temperature in the cutting zone. *Temperature rise* is a very important factor in machining because of its major adverse effects, such as the following:

- Excessive temperature lowers the strength, hardness, stiffness, and wear resistance of the cutting tool; tools also may soften and undergo plastic deformation; thus, tool shape is altered.

- Increased heat causes uneven dimensional changes in the part being machined, making it difficult to control its dimensional accuracy and tolerances.
- An excessive temperature rise can induce thermal damage and metallurgical changes in the machined surface, adversely affecting its properties.

From the preceding sections, it can be seen that the main sources of heat in machining are: (a) the 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 for dull or worn tools. Much effort has been expended in establishing relationships among temperature and various material and process variables in cutting. A comprehensive expression for the *mean temperature*, T_{mean}, in *orthogonal cutting* is

$$T = \frac{0.000665 Y_f}{\rho c} \sqrt[3]{\frac{V t_o}{K}}, \tag{21.19a}$$

where the mean temperature is in K, Y_f is the flow stress 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 depend on temperature, it is important to use appropriate values that are compatible with the predicted temperature range. It can be seen from Eq. (21.19a) 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.

An expression for the *mean temperature in turning* on a lathe is given by

$$T_{\text{mean}} \propto V^a f^b, \tag{21.19b}$$

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.

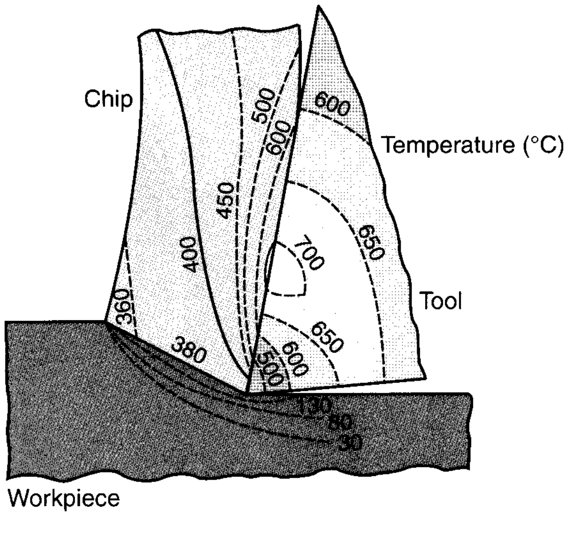


FIGURE 21.12 Typical temperature distribution in the cutting zone. Note the severe temperature gradients within the tool and the chip, and that the workpiece is relatively cool. *Source*: After G. Vieregge.

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 in 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. From the preceding discussions, it will be apparent that the particular temperature pattern depends on several factors pertaining to material properties and cutting conditions, including the type of cutting fluid (if any) used during machining.

The temperatures 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, respectively, as a function of the distance from the tip of the tool. The temperature 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.

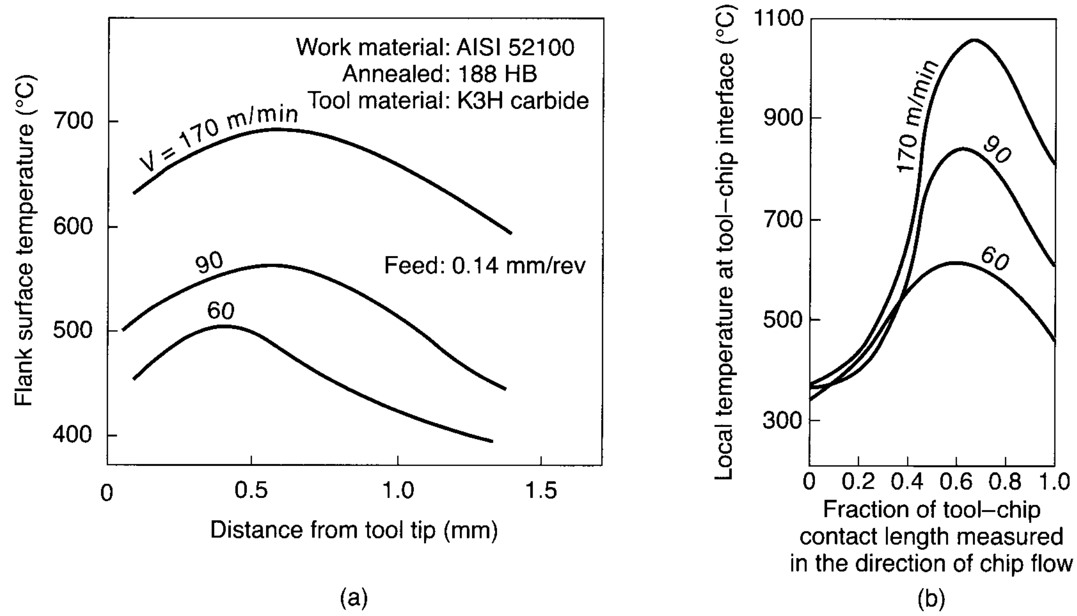


FIGURE 21.13 Temperatures developed in turning 52100 steel: (a) flank temperature distribution and (b) tool–chip interface temperature distribution. *Source*: After B.T. Chao and K.J. Trigger.

Note 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) produced at high cutting speeds. Chips can become red hot and create a safety hazard for the operator.

From Eq. (21.19b) 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, and hence the temperature rises (eventually becoming almost an adiabatic process). This effect of speed can be demonstrated easily by rubbing your hands together faster and faster.

As can be seen from Fig. 21.14, the chip carries away most of the heat generated. It has been estimated that 90% of the energy is removed by the chip during a typical machining operation, with the rest by the tool and the workpiece. Note in this figure that, as the cutting speed increases, a larger proportion of the total heat generated is carried away by the chip, and less heat goes into the workpiece or the tool. This is one reason that machining speeds have been increasing significantly over the years (see *high-speed machining*, Section 25.5). The other main benefit is associated with the favorable economics in reducing machining time, as described in Section 25.8.

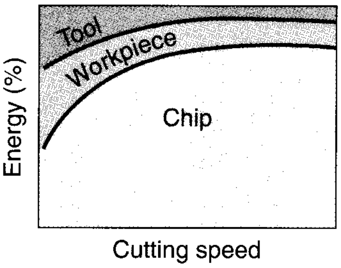


FIGURE 21.14 Proportion of the heat generated in cutting transferred to the tool, workpiece, and chip as a function of the cutting speed. Note that the chip removes most of the heat.

Techniques for Measuring Temperature. Temperatures and their distribution in the cutting zone may be determined from **thermocouples** embedded in the tool or the workpiece. This technique has been used successfully, although it involves considerable effort. It is easier to determine the *mean* temperature with the **thermal emf** (electromotive force) at the tool–chip interface, which acts as a *hot junction* between two different materials (i.e., tool and chip). **Infrared radiation**

from the cutting zone may be monitored with a *radiation pyrometer*. However, this technique indicates only surface temperatures; the accuracy of the results depends on the emissivity of the surfaces, which is difficult to determine accurately.

21.5 Tool Life: Wear and Failure

We have seen 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 along the rake face, and (d) sliding of the tool along the newly cut workpiece surface. These conditions induce **tool wear**, which is a major consideration in all machining operations, as are mold and die wear in casting and metalworking. Tool wear adversely affects tool life, the quality of the machined surface and its dimensional accuracy, and, consequently, the economics of cutting operations.

Wear is a gradual process, much like the wear of the tip of an ordinary pencil. The rate of tool wear depends on tool and workpiece materials, tool geometry, process parameters, cutting fluids, and the characteristics of the machine tool. Tool wear and the changes in tool geometry during cutting manifest themselves in different ways, generally classified as *flank wear*, *crater wear*, *nose wear*, *notching*, *plastic deformation of the tool tip*, *chipping*, and *gross fracture* (Fig. 21.15).

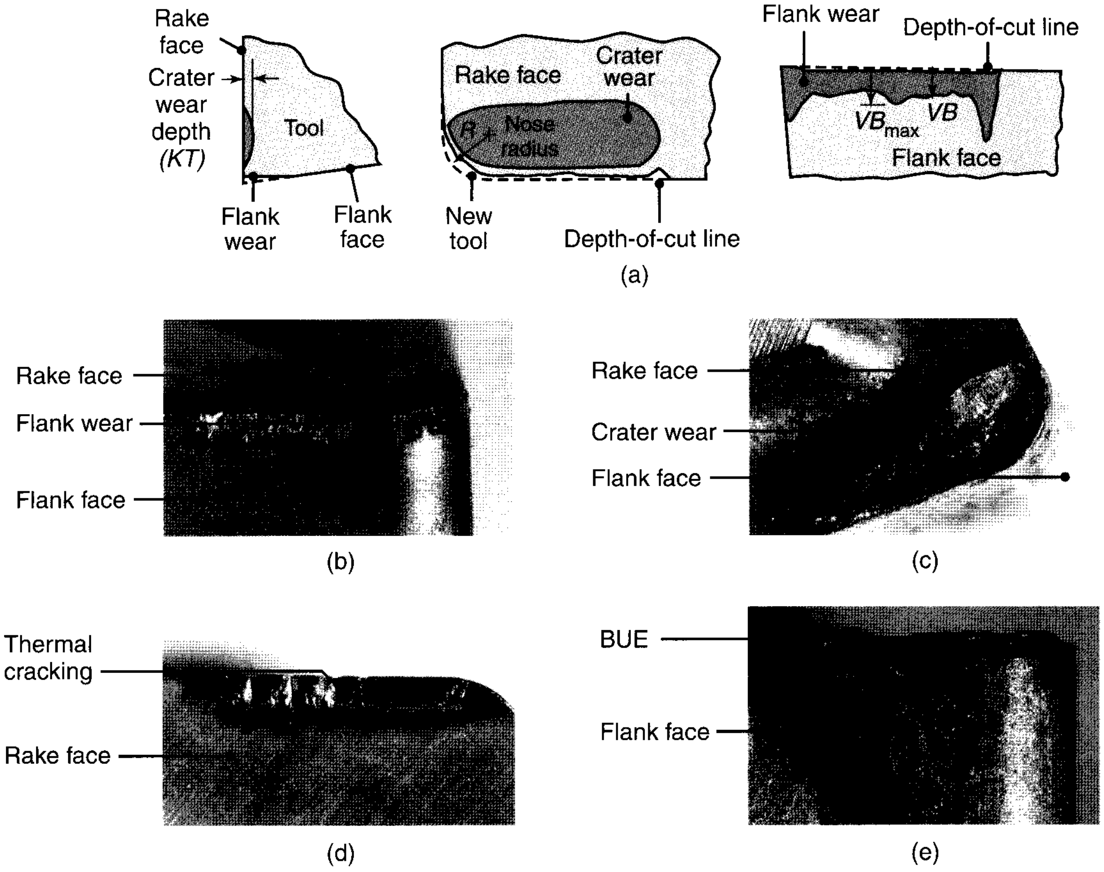


FIGURE 21.15 (a) Features of tool wear in a turning operation. The VB indicates average flank wear. (b)–(e) Examples of wear in cutting tools: (b) flank wear, (c) crater wear, (d) thermal cracking, and (e) flank wear and built-up edge. *Source*: (a) Terms and definitions reproduced with the permission of the International Organization for Standardization, ISO, copyright remains with ISO. (b)–(e) Courtesy of Kennametal Inc.

21.5.1 Flank Wear

Flank wear occurs on the relief (flank) face of the tool (Figs. 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 (see Section 33.5) and (b) high temperatures, which adversely affect 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:

$$VT^n = C. \tag{21.20a}$$

Here, 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 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. Generally, however, 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. (21.20) can be rewritten as

$$T = \left(\frac{C}{V}\right)^{\frac{1}{n}}, \tag{21.20b}$$

where it can be seen that for constant values of C, the smaller the value of n, the lower is the tool life.

Cutting speed is the most important process variable associated with tool life, followed by depth of cut and feed, f. For turning, Eq. (21.20) can be modified to

$$VT^n d^x f^y = C, \tag{21.21}$$

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.

We can rewrite Eq. (21.21) as

$$T = C^{1/n} V^{-1/n} d^{-x/n} f^{-y/n}, \tag{21.22}$$

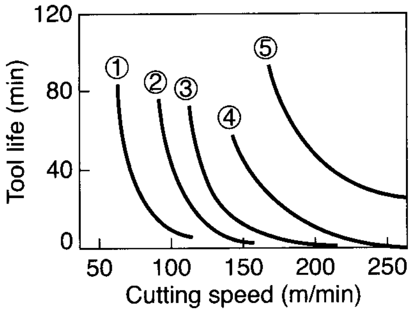
or, using typical values, as

$$T \approx C^7 V^{-7} d^{-1} f^{-4}. \tag{21.23}$$

TABLE 21.3

Ranges of n Values for the Taylor Equation (21.20a) 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



	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.

To obtain a constant tool life, the following observations can be made from Eq. (21.23): (a) If the feed or the depth of cut is increased, the cutting speed must be decreased (and vice versa), and (b) depending on 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 by performing cutting tests on various materials 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 workpiece-material microstructures.

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 and hard constituents in the material or on the surface of the workpiece (such as rust, scale, and slag) also are important factors, because their abrasive action reduces tool life.

The exponent *n* can be determined from tool-life curves (Fig. 21.17). Note that the smaller the *n* value, 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, the exponent *n* can indeed become *negative* at low cutting speeds, meaning that tool-life curves actually can reach a maximum and then curve downward. Because of this possibility, caution should 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, it is understandable that it also strongly influences wear. Thus, as temperature increases, flank wear rapidly increases.

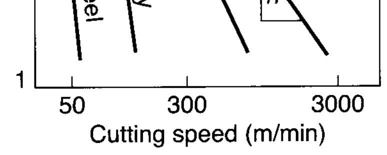


FIGURE 21.17 Tool-life curves for a variety of cutting-tool materials. The negative reciprocal of the slope of these curves is the exponent *n* in the Taylor tool-life Equation (21.20a), and *C* is the cutting speed at $T = 1$ min, ranging from about 60 to 3,000 m/min in this figure.

EXAMPLE 21.2 Increasing Tool Life by Reducing the Cutting Speed

Using the Taylor Equation (21.20a) for tool life and letting $n = 0.5$ and $C = 120$, calculate the percentage increase in tool life when the cutting speed is reduced by 50%.

Solution Since $n = 0.5$, the Taylor equation can be rewritten as $VT^{0.5} = 120$. Let's denote V_1 as the initial speed and V_2 the reduced speed; thus, $V_2 = 0.5V_1$. Because C is the constant 120, we have the relationship

$$0.5V_1\sqrt{T_2} = V_1\sqrt{T_1}.$$

Simplifying this equation, $T_2/T_1 = 1/0.25 = 4$. This indicates that 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%. Thus, 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.

Allowable Wear Land. We realize that we have to sharpen a knife or a pair of scissors when the quality of the cut deteriorates or the forces required are too high. Similarly, cutting tools need to be replaced (or resharpened) 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* (*VB* in Fig. 21.15a) for various machining conditions is given in Table 21.4. For improved dimensional accuracy, tolerances, 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 is generally the one that yields a tool life of 60 to 120 min, and for a carbide tool, it is 30 to 60 min. However, depending on the particular workpiece, the operation, and the high-productivity considerations due to the use of modern, computer-controlled machine tools, the cutting speeds selected can vary significantly from these values.

Optimum Cutting Speed. We have noted that as cutting speed increases, tool life is reduced rapidly. On the other hand, if the cutting speed is low, tool life is long, but the *rate* at which material is removed is also low. Thus, there is an *optimum cutting speed*. Because it involves several other parameters, we will describe this topic further in Section 25.8.

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, VB_{max}, is about twice that for *VB*.

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 "one" condition (that is, as cast with a hardness of 265 HB). We note that when the cutting speed is 60 m/min, tool life is about 40 min. Thus, the tool travels a distance of 60 m/min × 40 min = 2400 m before it has to be replaced. However, when the cutting speed is increased to 120 m/min, the tool life is reduced to about 5 min and the tool travels 120 m/min × 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. It is important to 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 Figs. 21.15a, and c, and Fig. 21.18, which illustrates various types of tool wear and failures. It readily can be seen that crater wear changes the tool–chip interface contact geometry. The most significant factors influencing crater wear are (a) the temperature at the tool–chip interface and (b) the chemical affinity between the tool and workpiece materials. Additionally, the factors influencing flank wear may affect crater wear.

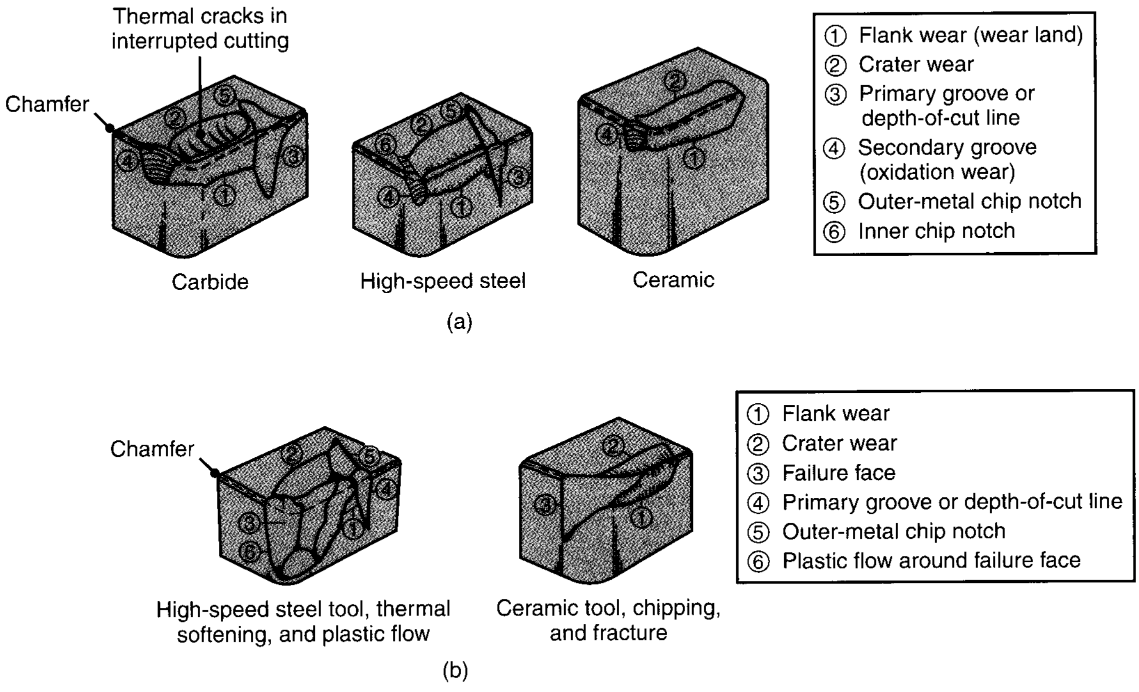


FIGURE 21.18 (a) Schematic illustrations of types of wear observed on various cutting tools. (b) Schematic illustrations of catastrophic tool failures. A wide range of parameters influence these wear and failure patterns. *Source*: Courtesy of V.C. Venkatesh.

Crater wear generally is attributed to a **diffusion mechanism**—that is, the movement of atoms across the tool–chip interface. Since diffusion rate increases with increasing temperature, crater wear increases as temperature increases. Note in Fig. 21.19, for example, how rapidly crater wear increases within a narrow range of temperatures. Applying protective *coatings* to tools is an effective means of slowing the diffusion process and thus reducing crater wear. Typical coatings are titanium nitride, titanium carbide, 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 cut at high speeds, is shown in Fig. 21.20. Note that the crater-wear pattern on the tool coincides with its discoloration pattern, which is an indication of the presence of high temperatures.

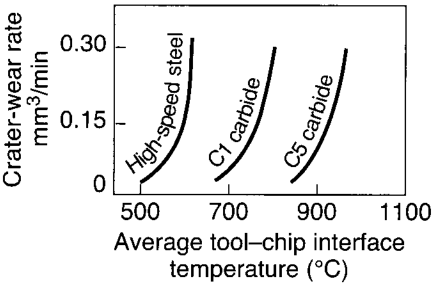


FIGURE 21.19 Relationship between crater-wear rate and average tool–chip interface temperature: (1) high-speed steel, (2) C1 carbide, and (3) C5 carbide (see Table 22.4). Note how rapidly crater-wear rate increases with an incremental increase in temperature. *Source*: After B.T. Chao and K.J. Trigger.

21.5.3 Other Types of Wear, Chipping, and Fracture

We now describe the factors involved in other types of cutting-tool wear 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 chip formation, and causes rubbing of the tool over the workpiece, raising its temperature and possibly 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 also may undergo *plastic deformation* because of temperature rises in the cutting zone, where temperatures can easily reach 1000°C in machining steels and can be higher in stronger materials.

Notches or *grooves* observed on cutting tools, as shown in Figs. 21.15a and 21.18, have been attributed to the fact that the region they occupy is the boundary where the chip is no longer in contact with the tool. Known as the **depth-of-cut line** (DOC) with a depth *VN*, this boundary oscillates because of inherent variations in the cutting operation. Furthermore, the region is in contact with the machined surface generated during the previous cut; the thin work-hardened layer that can develop will contribute to the formation of the wear groove. If sufficiently deep, the notch can lead to gross chipping of the tool tip because of its reduced cross section, as well as the notch sensitivity of the tool material.

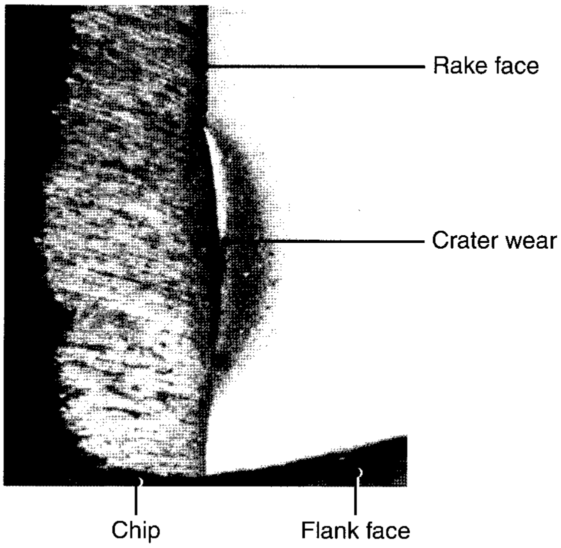


FIGURE 21.20 Interface of a cutting tool (right) and chip (left) in machining plain-carbon steel. The discoloration of the tool indicates the presence of high temperatures. Compare this figure with the temperature profiles shown in Fig. 21.12. *Source*: Courtesy of P.K. Wright.

Scale and oxide layers on a workpiece surface also contribute to notch wear, because these layers are hard and abrasive. Thus, light cuts should not be taken on rusted workpieces, and the depth of cut should be greater than the thickness of the oxide film or the work-hardened layer. 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 subject to wear, tools may undergo *chipping*, in which 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 the tip of a pencil if it is too sharp. The chipped fragments from the cutting tool may be very small (**microchipping** or **macrochipping**), 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 a small crack or defect already exists. Unlike wear, which is a gradual process, chipping is a sudden loss of tool material and a corresponding change in its shape. As can be expected, chipping has a major detrimental effect on surface finish, surface integrity, and the dimensional accuracy of the workpiece.

Two main causes of chipping are the following:

- **Mechanical shock** (i.e., impact due to interrupted cutting, as in turning a splined shaft on a lathe).
- **Thermal fatigue** (i.e., cyclic variations in the temperature of the tool in interrupted cutting).

Thermal cracks usually are 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.

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 toward the tool tip, thus weakening the tip because of reduced material volume and causing chipping.

21.5.4 Tool-condition Monitoring

With computer-controlled machine tools and automated manufacturing, the reliable and repeatable performance of cutting tools is a critical consideration. As described in Chapters 23 through 25, modern machine tools operate with little direct supervision by a machine operator and generally are enclosed, making it impossible or difficult to monitor the machining operation and the condition of the tool. It is therefore essential to continuously and indirectly monitor the condition of the cutting tool so as to note, for example, wear, chipping, or gross tool failure.

In modern machine tools, tool-condition monitoring systems are 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 the periodic observation of changes in the tool profile. This is a common and reliable technique and is done with a microscope (**toolmakers' microscope**). However, this requires that the cutting operation be stopped for tool observation. Another direct method involves programming the tool to contact a sensor after every machining cycle; this approach allows the detection of broken tools. Usually, the sensor has the appearance of a pin that must be depressed by the tool tip.

Indirect methods of observing tool conditions involve the correlation of the tool condition with parameters such as cutting forces, power, temperature rise, workpiece surface finish, vibration, and chatter. A powerful 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 tool-condition monitoring system consists of **transducers** that are installed in original machine tools or are retrofitted on existing machines. They continually monitor torque and forces during cutting. The signals are preamplified, and a microprocessor analyzes and interprets their content. The system is capable of differentiating the signals that come from different sources, such as tool breakage, tool wear, a missing tool, overloading of the machine tool, or colliding with machine components. The system also can compensate automatically for tool wear and thus improve the dimensional accuracy of the part being machined.

The design of transducers must be such that they are (a) nonintrusive to the machining operation, (b) accurate and repeatable in signal detection, (c) resistant to abuse and the shop-floor environment, and (d) cost effective. Continued progress is being made in the development of **sensors**, including the use of *infrared* and *fiber-optic techniques* for temperature measurement during machining.

In lower cost computer numerical-control machine tools, monitoring is done by **tool-cycle time**. In a production environment, once the life expectancy of a cutting tool or insert has been determined, it can be entered into the machine control unit, so that the operator is prompted to make a tool or cutter change when that time is reached. This process is inexpensive and fairly reliable, although not totally so, because of the inherent statistical variation in tool life.

21.6 Surface Finish and Integrity

Surface finish influences not only the dimensional accuracy of machined parts but also their properties and their performance in service. The term *surface finish* describes the geometric features of a surface (see Chapter 33), and *surface integrity* pertains to material properties, such as fatigue life and corrosion resistance, that are strongly influenced by the nature of the surface produced.

With its significant effect on the tool-tip profile, the *built-up edge* has the greatest influence on surface finish. Figure 21.21 shows the surfaces obtained in two

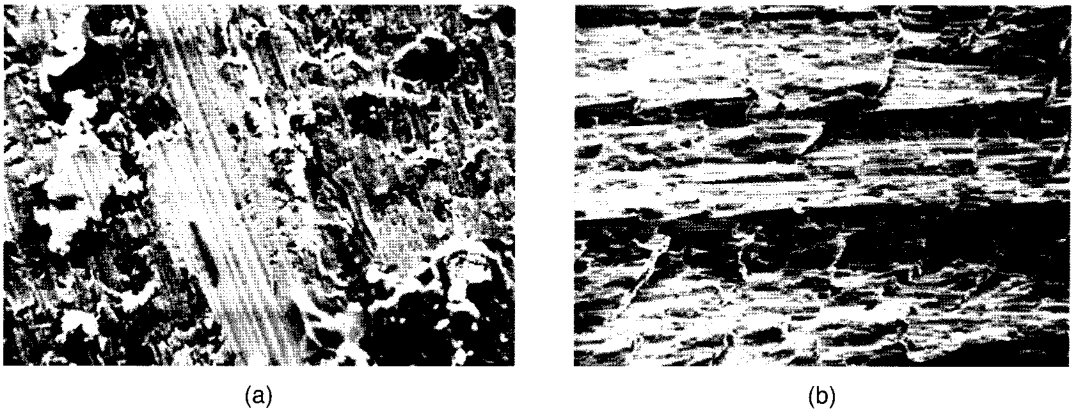

(a) (b)

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*: Courtesy of J T. Black and S. Ramalingam.

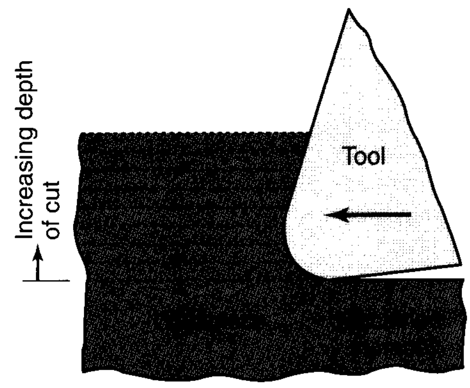


FIGURE 21.22 Schematic illustration of a dull tool with respect to the depth of cut in orthogonal machining (exaggerated). Note that the tool has a positive rake angle, but as the depth of cut decreases, the rake angle effectively can become negative. The tool then simply rides over the workpiece (without cutting) and burnishes its surface; this action raises the workpiece temperature and causes surface residual stresses.

different cutting operations. Note the considerable damage to the surfaces from BUE; its damage is manifested in the scuffing marks, which deviate from the straight grooves that would result from normal machining, as seen in Fig. 21.2. Ceramic and diamond 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 edges, just as the tip of a dull pencil or the cutting 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.

If the *tip radius* of the tool (not to be confused with the 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 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, the tool leaves a spiral profile (**feed marks**) on the machined surface as it moves across the workpiece, as shown in Figs. 21.2 and 21.23. We can see that the higher the feed, f, and the smaller the tool-nose radius, R, the more prominent these marks will be. It can be shown that the surface roughness for such a case is given by

$$R_t = \frac{f^2}{8R}, \qquad (21.24)$$

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. (Further details on surface roughness are given for individual machining processes as they are discussed.)

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 cutting, it will affect the workpiece surface finish adversely. The reason is that a vibrating tool periodically changes the dimensions of the cut. Excessive chatter also can cause chipping and premature failure of the more brittle cutting tools, such as ceramics and diamond.

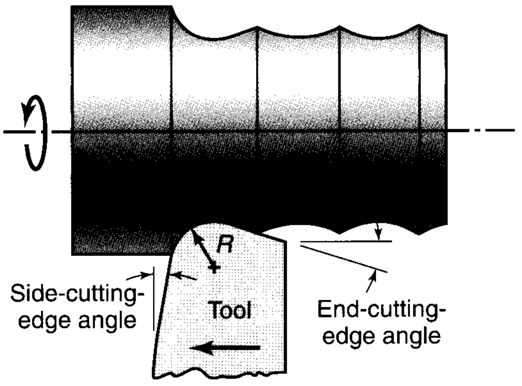


FIGURE 21.23 Schematic illustration of feed marks on a surface being turned (exaggerated).

Factors influencing *surface integrity* are as follows:

- Temperatures generated during processing and possible metallurgical transformations.
- Surface residual stresses.
- Severe plastic deformation and strain hardening of the machined surfaces, tearing, and cracking.

Each of these factors can have major adverse effects on the machined part but can be taken care of by careful selection and maintenance of cutting tools and control of process variables.

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. Surface finish is not a primary consideration, since it will be improved during finish machining. Of course, it is important that there be no subsurface-damage results from rough machining that cannot 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 required.
4. The level of difficulty in chip control.

Thus, good machinability indicates good surface finish and surface integrity, a long tool life, and low force and power requirements. As for chip control, and as stated earlier regarding continuous chips, long, thin, stringy, and curled chips can interfere severely with the cutting operation by becoming entangled in the cutting zone.

Because of the complex nature of cutting operations, it is difficult to establish relationships that quantitatively define the machinability of a particular material. In machining practice, tool life and surface roughness generally are considered to be the most important factors in machinability. Although not used much anymore due to their qualitative and misleading nature, approximate **machinability ratings (indexes)** have been available for many years for each type of material and its condition.

In these ratings, the standard material is AISI 1112 steel (resulfurized), with a rating of 100. This means that, for a tool life of 60 min, this steel should be machined at a cutting speed of 30 m/min. Examples of typical ratings are 3140 steel at 55; free-cutting brass at 300; 2011 wrought aluminum at 200; pearlitic gray iron at 70; and precipitation-hardening 17-7 steel at 20.

These qualitative aspects of machinability are not sufficient (of course) to guide a machine operator in determining the machining parameters in order to produce an acceptable part economically. Hence, in subsequent chapters, we present several tables in which, for various groups of materials, *specific recommendations* are given regarding such parameters as cutting speed, feed, depth of cut, cutting tools and their 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. Because steels are among the most important engineering materials (as also noted in Chapter 5) their machinability has been studied extensively. Carbon steels have a wide range of machinability, depending on their ductility and hardness. If a carbon steel is too ductile, chip formation can produce 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. Cold-worked carbon steels are desirable from a machinability standpoint.

An important group of steels is **free-machining steels**, containing sulfur and phosphorus. *Sulfur* forms manganese-sulfide inclusions (second-phase particles), 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. Note that soft steels can be difficult to machine because of their tendency for built-up edge formation and the resulting poor surface finish.

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. Because of its low shear strength, lead acts as a *solid* lubricant and is smeared over the tool–chip interface during cutting.

When the temperature developed 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. In addition to having this effect, lead lowers the shear stress in the primary shear zone, thus reducing cutting forces and power consumption. Lead can be used with every grade of steel and is identified by the letter "L" between the second and third numerals in steel identification (e.g., 10L45). (Note that in stainless steels, a similar use of the letter *L* means "low carbon," a condition that improves their corrosion resistance.)

However, because lead is a well-known *toxin* and a pollutant, there are serious environmental concerns about its use in steels (estimated at 4000 metric tons of lead consumption every year in the production of steels). Consequently, 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, although their performance is not as good.

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. Increases in temperature are reduced correspondingly. Consequently, these steels produce less crater wear, especially at high cutting speeds.

Alloy steels can have a wide variety of compositions and hardnesses. Consequently, their machinability cannot be generalized, although they have higher levels of hardness and other mechanical properties. 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 cutting tools, producing good surface finish, integrity, and dimensional accuracy.

Effects of Various Elements in Steels. The presence of *aluminum* and *silicon* in steels is always harmful, because these elements combine with oxygen to form 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 are more 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. The machinability of most steels is improved by cold working, which hardens the material and reduces the tendency for built-up edge formation.

Other alloying elements (such as *nickel*, *chromium*, *molybdenum*, and *vanadium*) that improve the properties of steels 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: The higher the oxygen content, the lower the aspect ratio, and the higher the machinability.

In improving the machinability of steels, however, it is important to consider the possible detrimental effects of the alloying elements on the properties and strength of the machined part in service. At elevated temperatures, for example, lead causes *embrittlement* of steels (liquid-metal embrittlement and hot shortness; see Section 1.5.2), although at room temperature it has no effect on mechanical properties.

Sulfur can reduce the hot workability of steels 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 (anisotropy). Rephosphorized steels are significantly less ductile and are produced solely to improve machinability.

Stainless Steels. Austenitic (300 series) steels generally are difficult to machine. Chatter can be a problem, necessitating machine tools with high stiffness. Ferritic stainless steels (also 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, but they can be abrasive depending on composition, especially pearlite. Free carbides in castings reduce their machinability and cause tool chipping or fracture. *Nodular* and *malleable irons* are machinable with hard tool materials.

21.7.2 Machinability of Nonferrous Metals

Following is a summary of the machinability of nonferrous metals and alloys, 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. Thus, high cutting 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** generally is machinable, but because the fine particles produced during machining are toxic, it requires machining in 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, although 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. Bronzes are more difficult to machine than brass.
- **Magnesium** is very easy to machine, with good surface finish and prolonged tool life. However, care should be exercised because of its high rate of oxidation (*pyrophoric*) and the danger of fire.
- **Molybdenum** is ductile and work hardening. It can produce poor surface finish; thus, sharp tools are essential.
- **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), 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; hence, its machinability is low, although it improves greatly at elevated temperatures.
- **Zirconium** has good machinability, but it requires a coolant-type cutting fluid because of the danger of explosion and fire.

21.7.3 Machinability of Miscellaneous Materials

Thermoplastics generally have low thermal conductivity and a 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 support of the workpiece. External cooling of the cutting zone may be necessary to keep the chips from becoming gummy and sticking to the tools. Cooling usually can be achieved with a jet of air, a vapor mist, or 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 that are present; hence, they are difficult to machine. Fiber tearing, pulling, and edge delamination are significant problems and can lead to severe reduction in the load-carrying capacity of the machined component. Machining of these materials requires careful handling and removal of debris 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 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 8.2.5) and with the selection of appropriate processing parameters, such as *ductile-regime cutting* (described in Section 25.7).

Wood is an orthotropic 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., remains largely an art. Two 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 1.7, and Tables 1.2 and 3.1). Thus, steels are hot machined above a 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) a reduced tendency for vibration and chatter.

SUMMARY

- The Merchant model of orthogonal cutting, although a simple model, nonetheless takes into account all of the major process parameters in machining. Central to the model is that machining takes place through localized deformation occurring on the shear plane.
- 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 with other shaping techniques. On the other hand, these processes generally take more time, waste some material in the form of chips, and 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 shape and tool material; cutting conditions such as speed, feed, and depth of cut; the use of cutting fluids; and the 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 is an important consideration, since it can have adverse effects on tool life, as well as on the 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; and the characteristics of the machine tool. Tool failure also may occur by notching, chipping, and gross fracture.
- The surface finish of machined components can affect product integrity adversely. Important variables are the geometry and condition of the cutting tool, the type of chip produced, and process variables.
- Machinability generally is defined in terms of surface finish, tool life, force and power requirements, and chip control. The machinability of materials depends on their composition, properties, and microstructure. Thus, proper selection and control of process variables are important.

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
Machinability ratings
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

Astakhov, V.P., **Metal Cutting Mechanics**, CRC Press, 1998.

ASM Handbook, Vol. 16: *Machining*, ASM International, 1989.

Boothroyd, G., and Knight, W.A., **Fundamentals of Machining and Machine *Tools***, 3rd ed., Marcel Dekker, 2006.

Brown, J., **Advanced Machining Technology Handbook**, McGraw-Hill, 1998.

Cheng, K., **Machining Dynamics: Fundamentals, Applications and Practice**, Springer, 2008.

Cormier, D., **McGraw-Hill Machining and Metalworking Handbook**, 2005.

Davim, J.P. (ed.), **Machining: Fundamentals and Recent Advances**, Springer, 2008.

Davis, J.R. (ed.), **Tool Materials**, ASM International, 1995.

DeVries, W.R., **Analysis of Material Removal Processes**, Springer, 1992.

Machinery's Handbook, revised periodically, Industrial Press.

Metal Cutting Tool Handbook, 7th ed., Industrial Press, 1989.

Oxley, P.L.B., **Mechanics of Machining: An Analytical Approach to Assessing Machinability**, Wiley, 1989.

Shaw, M.C., **Metal Cutting Principles**, 2nd ed., Oxford, 2005.

Smith, G.T., **Cutting Tool Technology: Industrial Handbook**, Spinger, 2008.

Stephenson, D.A., and Agapiou, J.S., **Metal Cutting: Theory and Practice**, 2nd ed., CRC Press, 2006.

Stout, K.J., Davis, E.J., and Sullivan, P.J., **Atlas of Machined Surfaces**, Chapman and Hall, 1990.

Trent, E.M., and Wright, P.K., **Metal Cutting**, 4th ed., Butterworth Heinemann, 2000.

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. Explain the difference between positive and negative rake angles. What is the importance of the rake angle?

21.4. Comment on the role and importance of the relief angle.

21.5. Explain the difference between discontinuous chips and segmented chips.

21.6. Why should we be interested in the magnitude of the thrust force in cutting?

21.7. What are the differences between orthogonal and oblique cutting?

21.8. Is there any advantage to having a built-up edge on a tool? Explain.

21.9. What is the function of chip breakers? How do they function?

21.10. Identify the forces involved in a cutting operation. Which of these forces contributes to the power required?

21.11. Explain the characteristics of different types of tool wear.

21.12. List the factors that contribute to poor surface finish in cutting.

21.13. Explain what is meant by the term *machinability* and what it involves. Why does titanium have poor machinability?

QUALITATIVE PROBLEMS

21.14. Are the locations of maximum temperature and crater wear related? If so, explain why.

21.15. Is material ductility important for machinability? Explain.

21.16. Explain why studying the types of chips produced is important in understanding cutting operations.

21.17. 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.)

21.18. Tool life can be almost infinite at low cutting speeds. Would you then recommend that all machining be done at low speeds? Explain.

21.19. Explain the consequences of allowing temperatures to rise to high levels in cutting.

21.20. The cutting force increases with the depth of cut and decreasing rake angle. Explain why.

21.21. Why is it not always advisable to increase the cutting speed in order to increase the production rate?

21.22. What are the consequences if a cutting tool chips?

21.23. What are the effects of performing a cutting operation with a dull tool? A very sharp tool?

21.24. 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?

21.25. Explain why it is possible to remove more material between tool resharpenings by lowering the cutting speed.

21.26. Noting that the dimension d in Fig. 21.4a is very small, explain why the shear strain rate in metal cutting is so high.

21.27. Explain the significance of Eq. (21.7).

21.28. Comment on your observations regarding Figs. 21.12 and 21.13.

21.29. Describe the consequences of exceeding the allowable wear land (Table 21.4) for various cutting-tool materials.

21.30. Comment on your observations regarding the hardness variations shown in Fig. 21.6a.

21.31. Why does the temperature in cutting depend on the cutting speed, feed, and depth of cut? Explain in terms of the relevant process variables.

21.32. You will note that the values of a and b in Eq. (21.19b) are higher for high-speed steels than for carbides. Why is this so?

21.33. As shown in Fig. 21.14, the percentage of the total cutting energy carried away by the chip increases with increasing cutting speed. Why?

21.34. Describe the effects that a dull tool can have on cutting operations.

21.35. Explain whether it is desirable to have a high or low (a) n value and (b) C value in the Taylor tool-life equation.

21.36. The tool-life curve for ceramic tools in Fig. 21.17 is to the right of those for other tool materials. Why?

21.37. Why are tool temperatures low at low cutting speeds and high at high cutting speeds?

21.38. Can high-speed machining be performed without the use of a cutting fluid?

21.39. Given your understanding of the basic metal-cutting process, what are the important physical and chemical properties of a cutting tool?

QUANTITATIVE PROBLEMS

▶**21.40.** Let $n = 0.5$ and $C = 90$ in the Taylor equation for tool wear. What is the percent increase in tool life if the cutting speed is reduced by (a) 50% and (b) 75%?

▶**21.41.** Assume that, in orthogonal cutting, the rake angle is 25° and the coefficient of friction is 0.2. Using Eq. (21.3), determine the percentage increase in chip thickness when the friction is doubled.

▶**21.42.** Derive Eq. (21.11).

▶**21.43.** Taking carbide as an example and using Eq. (21.19b), determine how much the feed should be reduced in order to keep the mean temperature constant when the cutting speed is doubled.

▶**21.44.** 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.

▶**21.45.** An orthogonal cutting operation is being carried out under the following conditions: $t_o = 0.1$ mm, $t_c = 0.2$ mm, width of cut = 5 mm, $V = 2$ m/s, rake angle = 10°, $F_c = 500$ N, and $F_t = 200$ N. Calculate the percentage of the total energy that is dissipated in the shear plane.

▶**21.46.** Explain how you would go about estimating the C and n values for the four tool materials shown in Fig. 21.17.

21.47. Derive Eqs. (21.1).

▶**21.48.** 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 50°. (*Note:* do not use Eq. (21.3) or Eq. (21.4)).

▶**21.49.** Show that, for the same shear angle, there are two rake angles that give the same cutting ratio.

▶**21.50.** 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.

▶**21.51.** 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 $y = 0.6$.

▶**21.52.** In Example 21.3, if the cutting speed V is doubled, will the answer be different? Explain.

21.53. Using Eq. (21.24), select an appropriate feed for $R = 1$ mm and a desired roughness of 1 μm. How would you adjust this feed to allow for nose wear of the tool during extended cuts? Explain your reasoning.

▶**21.54.** With a carbide tool, the temperature in a cutting operation is measured as 650K 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 480K?

21.55. 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

▶**21.56.** 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.

21.57. 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.

21.58. Design an experimental setup whereby orthogonal cutting can be simulated in a turning operation on a lathe.

21.59. 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?

21.60. 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.

▶**21.61.** 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.

21.62. 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.

21.63. 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.

Cutting-Tool Materials and Cutting Fluids

CHAPTER 22

- 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.
- The properties and applications of high-speed steels, carbides, ceramics, cubic boron nitride, diamond, and coated tools are described in detail.
- The types of cutting fluids in common use are then described, including their functions and how they affect the machining operation.
- Trends in near-dry and dry machining are also discussed, and their importance with respect to environmentally friendly machining operations are explained.

EXAMPLES:

22.1 Introduction

The selection of cutting-tool materials for a particular application is among the most important factors in machining operations, just as the selection of mold and die materials was critical for forming and shaping processes (Chapters 14–19). We will discuss throughout this chapter the relevant properties and performance characteristics of all major types of cutting-tool materials, which will help us in tool selection. However, as it will become apparent, the complex nature of this subject does not always render itself to the determination of appropriate tool materials; hence, we also must rely on general guidelines and *recommendations* that have been accumulated in industry over many years. Beginning with Chapter 23, more detailed information on tool material recommendations for specific workpiece materials and machining operations will be presented.

As noted in the preceding chapter, the cutting tool is subjected to (a) high temperatures, (b) high contact stresses, and (c) rubbing along the tool–chip interface and along the machined surface. Consequently, the cutting-tool material must possess the following characteristics:

- **Hot hardness,** so that the hardness, strength, and wear resistance of the tool are maintained at the temperatures encountered in machining operations. This property ensures that the tool does not undergo any plastic deformation and

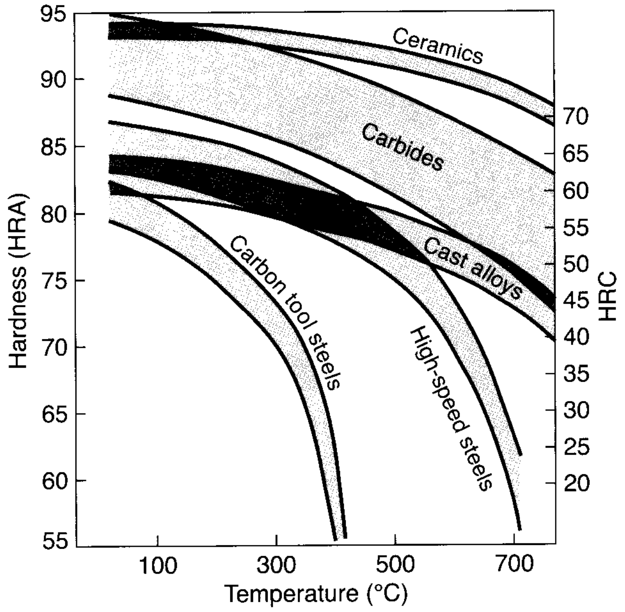


FIGURE 22.1 The hardness of various cutting-tool materials as a function of temperature (hot hardness). The wide range in each group of materials is due to the variety of tool compositions and treatments available for that group.

thus retains its shape and sharpness. Tool-material hardness as a function of temperature is shown in Fig. 22.1. Note the wide response of these materials and (not surprisingly) how well ceramics maintain their hardness at high temperatures. Carbon tool steels commonly were used as tool materials until the development of high-speed steels in the early 1900s. Note that carbon tool steels rapidly begin to lose their hardness even at moderate temperatures, meaning that they cannot be used for machining at high speeds (and thus at high temperatures). Consequently, the production rate will be low and costs will be high.

- **Toughness** and **impact strength** (or mechanical shock resistance), so that impact forces on the tool that are encountered repeatedly in interrupted cutting operations (such as milling and turning a splined shaft on a lathe) 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 encountered in interrupted cutting.
- **Wear resistance,** so that an acceptable tool life is obtained before replacement is necessary.
- **Chemical stability** and **inertness** with respect to the material being machined, to avoid or minimize any adverse reactions, adhesion, and tool–chip diffusion that would contribute to tool wear.

To respond to these demanding requirements, various cutting-tool materials with a wide range of mechanical, physical, and chemical properties have been developed over the years, as shown in Table 22.1. The properties listed in the first column 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 to be machined.
- Impact strength is important in making interrupted cuts in machining, such as in milling.
- Melting temperature of the tool material is important as compared to the temperatures developed in the cutting zone.
- The physical properties of thermal conductivity and coefficient of thermal expansion are important in determining the resistance of the tool materials to thermal fatigue and shock.

It must be recognized, however, that the tool materials may not have all of the desired properties for a particular machining operation. This situation can readily be seen from Table 22.2 by noting the opposite directions of the long horizontal arrows. For example, (a) high-speed steels are tough, but they have limited hot hardness, and (b) ceramics have high resistance to temperature and wear, but they are brittle and can chip. Note also that the cost of tools increases as we move to the right in Table 22.2—from HSS to diamond (which is the most expensive).

Table 22.3 shows the operating characteristics of tool materials in machining. Tool materials generally are divided into the following categories, listed in the order in which they were developed and implemented in industry (note that many of these

TABLE 22.1

General Characteristics of Tool Materials

Property	High-speed steels	Cast-cobalt alloys	Carbides		Ceramics	Cubic boron nitride	Single-crystal diamond*
			WC	TiC			
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.

TABLE 22.2

General Characteristics of Cutting-tool Materials (These Tool 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 to heavy	Light to heavy	Light to heavy	Light to heavy	Light to heavy	Light to heavy	Very light for single-crystal diamond
Processing method	Wrought, cast, HIP* sintering	Cast and HIP sintering	Cold pressing and sintering	CVD or PVD†	Cold pressing and sintering or HIP sintering	High-pressure, high-temperature sintering	High-pressure, high-temperature 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 use 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 use 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 low chemical stability at higher temperature
Diamond	High hardness and toughness, abrasive wear resistance	Chipping, oxidation, graphitization	Low strength, and low chemical stability at higher temperatures

Source: After R. Komanduri and other sources.

materials also are used for dies and molds in casting, forming, and shaping metallic and nonmetallic materials):

1. High-speed steels
2. Cast-cobalt alloys
3. Carbides
4. Coated tools
5. Alumina-based ceramics
6. Cubic boron nitride
7. Silicon-nitride-based ceramics
8. Diamond
9. Whisker-reinforced materials and nanomaterials.

Carbon steels are the oldest tool materials and have been used widely for drills, taps, broaches, and reamers since the 1880s. Low-alloy and medium-alloy steels were developed later 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 cutting at high speeds when the temperature rises significantly. Their use is limited to very low speed cutting operations, particularly in woodworking; hence, they are not of any particular significance in modern machining operations.

In this chapter, the following are described:

- The characteristics, applications, and limitations of cutting-tool materials in machining operations, including the required characteristics and costs.
- The applicable range of process variables for optimal performance.
- The types and characteristics of cutting fluids and their specific applications in a wide variety of machining operations.

22.2 High-speed Steels

High-speed steel (HSS) tools are so named because they were developed to machine at higher speeds than was previously possible. First produced in the early 1900s, high-speed steels are the most highly alloyed of the tool steels. They can be hardened to various depths, have good wear resistance, and are relatively inexpensive. Because of their toughness (hence high resistance to fracture), high-speed steels are suitable especially for (a) high positive rake-angle tools (i.e., those with small included angles), (b) interrupted cuts, (c) machine tools with low stiffness that are subject to vibration and chatter, and (d) complex and single-piece tools, such as drills, reamers, taps, and gear cutters. Their most important limitation (due to their lower hot hardness) is that their cutting speeds are low compared with those of carbide tools, as can be seen in 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 formed in the steel 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, and is less expensive. Consequently, 95% of all high-speed steel tools are made of the M-series steels. Table 5.6 lists three of these steels and their characteristics.

High-speed steel tools are available in wrought (rolled or forged), cast, and powder-metallurgy (sintered) forms. They can be **coated** for improved performance, as described in Section 22.5. High-speed steel tools also may be subjected to *surface treatments* (such as case hardening for improved hardness and wear resistance; see Section 4.10) or steam treatment at elevated temperatures to develop a hard, black oxide layer (*bluing*) for improved performance, including a reduced tendency for built-up edge formation.

EXAMPLE 22.1 Alloying Elements in High-speed Steel Cutting Tools

List the major alloying elements in high-speed steels and describe their effects in cutting tools.

Solution The major alloying elements in HSS are chromium, vanadium, tungsten, cobalt, and molybdenum. To appreciate their role in cutting tools, refer to Section 5.5.1 on the effects of various elements in steels and note the following:

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 following composition ranges: 38 to 53% Co, 30 to 33% Cr, and 10 to 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, these alloys are cast and ground into relatively simple shapes. They are now used only for special applications that involve deep, continuous **roughing cuts** at relatively high feeds and speeds—as much as twice the rates possible with high-speed steels. As described in Section 23.2, roughing cuts usually involve high feed rates and large depths of cut, with the primary purpose of removing large amounts of material with little regard for surface finish. Conversely, **finishing cuts** are performed at lower feeds and depths of cut, and the surface finish produced is a priority.

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. However, such speeds often are necessary to improve plant 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 used for machining are *tungsten* *carbide* and *titanium carbide*. In order to differentiate them from the coated tools described in Section 22.5, plain-carbide tools usually 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. These tools are manufactured using powder-metallurgy techniques (hence the term *sintered carbides* or *cemented carbides*), as described in Chapter 17. Tungsten-carbide particles are first combined with *cobalt* in a mixer, 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 the desired *insert* shapes. Tungsten carbides frequently are 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 of WC decrease, while its toughness increases because of the higher toughness of cobalt. Tungsten-carbide tools generally are used for cutting steels, cast irons, and abrasive nonferrous materials and largely have replaced HSS tools because of their better performance.

Micrograin Carbides. Cutting tools also are made of submicron and ultra-fine-grained (*micrograin*) carbides, including tungsten carbide, titanium carbide, and tantalum carbide. The grain size is in the range from 0.2 to 0.8 μm. Compared with the traditional carbides described previously, these tool materials are stronger, harder, and more wear resistant, thus improving productivity. In one application, microdrills with diameters on the order of 100 μm are being made from micrograin carbides and used in the fabrication of microelectronic circuit boards (Chapter 28).

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 promote tool life and performance. However, they are more expensive and cannot be justified for all applications.

22.4.2 Titanium Carbide

Titanium carbide (TiC) consists of 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 cutting at speeds higher than those appropriate for tungsten carbide.

22.4.3 Inserts

High-speed steel tools are shaped in one piece and ground to various geometric features (Fig. 21.10a); such tools include drill bits and milling and gear cutters. After the cutting edge wears, the tool has to be removed from its holder and reground.

Although a supply of sharp or resharpened tools usually is maintained, tool-changing operations are time consuming and inefficient. The need for a more effective method has led to the development of *inserts,* which are individual cutting tools with several cutting points (Fig. 22.2). A square insert has eight cutting points, and a triangular insert has six. Inserts usually are clamped on the *toolholder* with various locking mechanisms (Fig. 22.3). Although not used as commonly, inserts also may be *brazed* to the tool shank, but this practice has largely been abandoned.

Clamping is the preferred method of securing an insert because each insert has a number of cutting points and, after one edge is worn, it is **indexed** (rotated in its holder) to make another cutting point available. In addition to the examples in this figure, a wide variety of other toolholders is available for specific applications, including those with quick insertion and removal features.

Carbide inserts are available in a variety of shapes, such as square, triangle, diamond, and round. The strength of the cutting edge of an insert depends on its shape. The smaller the included angle (Fig. 22.4), the lower the strength of the edge. In order to further improve edge strength and prevent chipping, insert edges usually are honed, chamfered, or produced with a

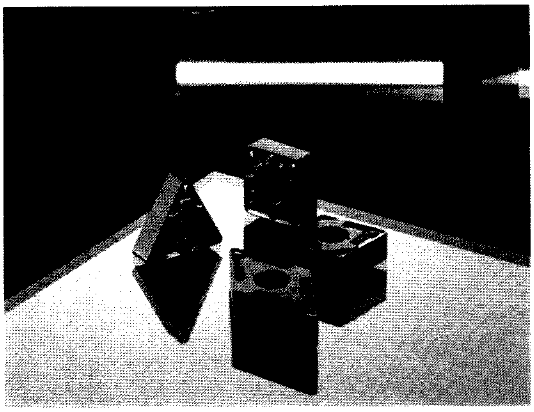


FIGURE 22.2 Typical carbide inserts with various shapes and chip-breaker features: Round inserts also are available, as can be seen in Figs. 22.3c and 22.4. The holes in the inserts are standardized for interchangeability in toolholders. *Source*: Courtesy of Kyocera Engineered Ceramics, Inc.

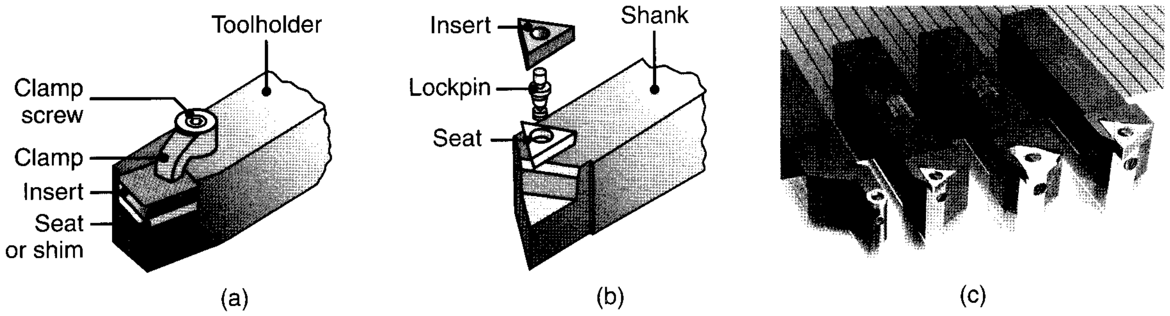


FIGURE 22.3 Methods of mounting inserts on toolholders: (a) clamping and (b) wing lockpins. (c) Examples of inserts mounted with threadless lockpins, which are secured with side screws. *Source*: Courtesy of Valenite.

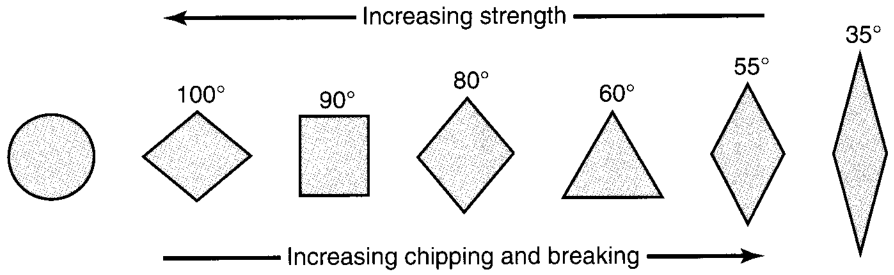


FIGURE 22.4 Relative edge strength and tendency for chipping of inserts with various shapes. Strength refers to the cutting edge indicated by the included angles. *Source*: Courtesy of Kennametal Inc.

negative land (Fig. 22.5). Most inserts are honed to a radius of about 0.025 mm.

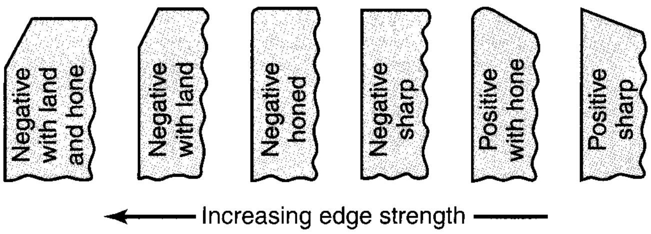


FIGURE 22.5 Edge preparation for inserts to improve edge strength. *Source*: Courtesy of Kennametal Inc.

Chip-breaker features on inserts (described in Section 21.2.1) are for the purposes of (a) controlling chip flow during machining, (b) eliminating long chips, and (c) reducing vibration and heat generated. Carbide inserts are available commercially with a wide variety of complex chip-breaker features, typical examples of which are shown in 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 cutting, and whether it is a roughing or finishing cut. Optimum chip-breaker geometries are being developed by computer-aided design and finite-element analysis techniques.

Stiffness of the machine tool (Section 25.3) is of major importance with 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 edges of the tool, increasing the tendency for the edges to chip off.

Low cutting speeds tend to encourage cold welding of the chip to the tool. Cutting fluids should be applied continuously and in large quantities if used to minimize the heating and cooling of the tool in interrupted cutting operations.

22.4.4 Classification of Carbides

With rapidly increasing global manufacturing and wider use of ISO (International Organization for Standardization) standards, carbide grades are classified using the letters P, M, and K (as shown in Tables 22.4 and 22.5) for a range of applications, including the traditional C grades used in the United States. Because of the wide variety of carbide compositions available and the broad range of machining applications and workpiece materials involved, efforts at ISO classification continue to be a difficult task. This is true especially when comparing ISO grades with the traditional grades classified by the American National Standards Institute (ANSI), which range from C1 to C8.

TABLE 22.4

ISO Classification of Carbide Cutting Tools According to Use

Symbol	Workpiece material	Color code	Designation in increasing order of wear resistance and decreasing order of toughness in each category (in increments of 5)
P	Ferrous metals with long chips	Blue	P01, P05–P50
M	Ferrous metals with long or short chips, nonferrous metals	Yellow	M10–M40
K	Ferrous metals with short chips, nonferrous metals, nonmetallic materials	Red	K01, K10–K40

TABLE 22.5

Classification of Tungsten Carbides According to Machining Applications

ISO standard	ANSI classification number (grade)	Materials to be machined	Machining operation	Type of carbide	Characteristics of	
					Cut	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

As described in Part I, new alloys and engineered materials are being developed continuously, particularly since the 1960s. These materials have high strength and toughness, but generally are abrasive and chemically reactive with tool materials. The difficulty of machining these materials efficiently and the need for improving their performance in machining the more common engineering materials have led to important developments in *coated tools*. Coatings have unique properties, such as

- Lower friction
- Higher adhesion
- Higher resistance to wear and cracking
- Acting as a diffusion barrier
- Higher hot hardness and impact resistance.

Coated tools can have lives 10 times longer than those of uncoated tools, allowing for high cutting speeds and thus reducing both the time required for machining operations and production costs. As can be seen in Fig. 22.6, machining time has been reduced steadily by a factor of more than 100 since 1900. This improvement has had a major impact on the economics of machining operations in conjuction with continued improvements in the design and construction of modern machine tools and their computer controls. As a result, coated tools now are used in 40 to 80% of all machining operations, particularly turning, milling, and drilling.

22.5.1 Coating Materials and Coating Methods

Commonly used coating materials are titanium nitride (TiN), titanium carbide (TiC), titanium carbonitride (TiCN), and aluminum oxide (Al_2O_3). These coatings,

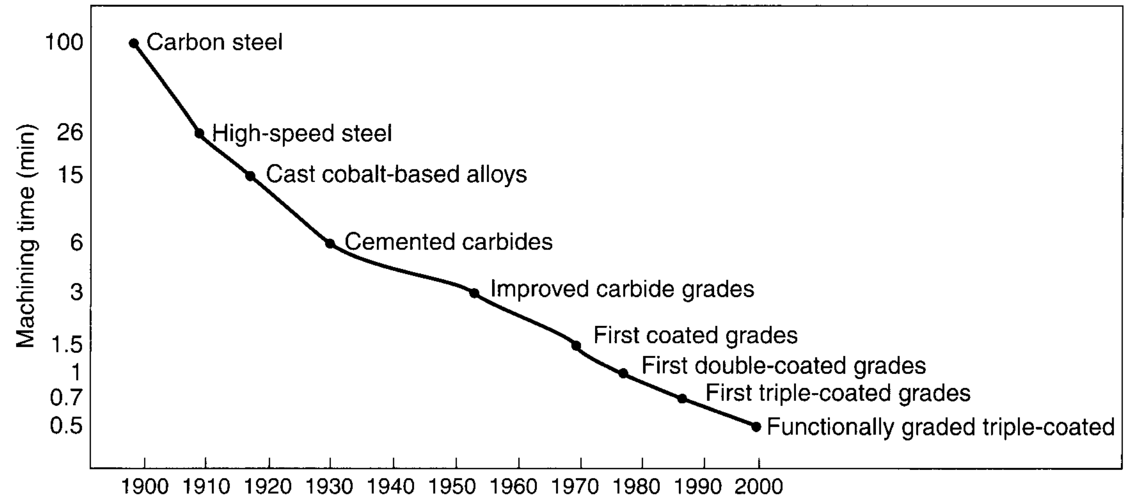


FIGURE 22.6 Relative time required to machine with various cutting-tool materials, indicating the year the tool materials were first introduced. Note that machining time has been reduced by two orders of magnitude within a hundred years. *Source*: Courtesy of Sandvik.

generally in the thickness range from 2 to 15 μm, are applied on cutting tools and inserts 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).

The CVD process is the most commonly used method for carbide tools with multiphase and ceramic coatings, both of which are described later in this section. Nonetheless, 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 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 for cutting tools and dies should have the following general characteristics:

- **High hardness** at elevated temperatures, to resist wear.
- **Chemical stability** and **inertness** to the workpiece material, to reduce wear.
- **Low thermal conductivity**, to prevent temperature rise in the substrate.
- **Compatibility and good bonding** to the substrate, to prevent flaking or spalling.
- **Little or no porosity** in the coating, to maintain its integrity and strength.

The effectiveness of coatings is enhanced by the hardness, toughness, and high thermal conductivity of the substrate (which may be carbide or high-speed steel). Honing of the cutting edges is an important procedure for the maintenance of coating strength; otherwise, the coating may peel or chip off at sharp edges and corners.

Titanium-nitride Coatings. Titanium-nitride coatings have low friction coefficients, high hardness, resistance to high temperature, and good adhesion to the substrate. Consequently, they greatly improve the life of high-speed steel tools, as well as the lives of carbide tools, drill bits, and cutters. Titanium-nitride-coated tools (gold in color) perform well at higher cutting speeds and feeds. Flank wear is significantly lower than that of uncoated tools (Fig. 22.7), and flank surfaces can be reground after use, since regrinding the tool does not remove the coating on the rake face of the tool. However, coated tools do not perform as well at low cutting speeds, because the coating can be worn off by chip adhesion. Therefore, the use of appropriate cutting fluids to discourage adhesion is important.

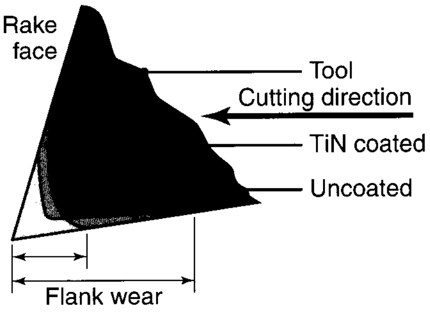


FIGURE 22.7 Schematic illustration of typical wear patterns on uncoated high-speed steel tools and titanium-nitride-coated tools. Note that flank wear is significantly lower for the coated tool.

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 suitable coating materials for tools. The most commonly used ceramic coating is aluminum oxide (Al_2O_3). However, because they are very stable (not chemically reactive), oxide coatings generally bond weakly to the substrate.

Multiphase Coatings. The desirable properties of the coatings just described can be combined and optimized with the use of *multiphase coatings*. Carbide tools are now available with two or three layers of such coatings and are particularly effective in machining cast irons and steels.

For example, one could first deposit TiC over the substrate, followed by Al_2O_3, and then TiN. The first layer should 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 as follows:

1. High-speed, continuous cutting: TiC/Al_2O_3
2. Heavy-duty, continuous cutting: TiC/Al_2O_3/TiN
3. Light, interrupted cutting: TiC/TiC + TiN/TiN.

Coatings also are available in **alternating multiphase layers**. The thickness of these layers is on the order of 2 to 10 μm—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 that is similar to the increase in the strength of metals with decreasing grain size. Therefore, thinner layers are harder than thicker layers.

A typical multiphase-coated carbide tool may consist of the following layers, starting from the top, along with their primary functions:

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, 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.

Diamond Coatings. The properties and applications of diamond, diamond coatings, and *diamond-like 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

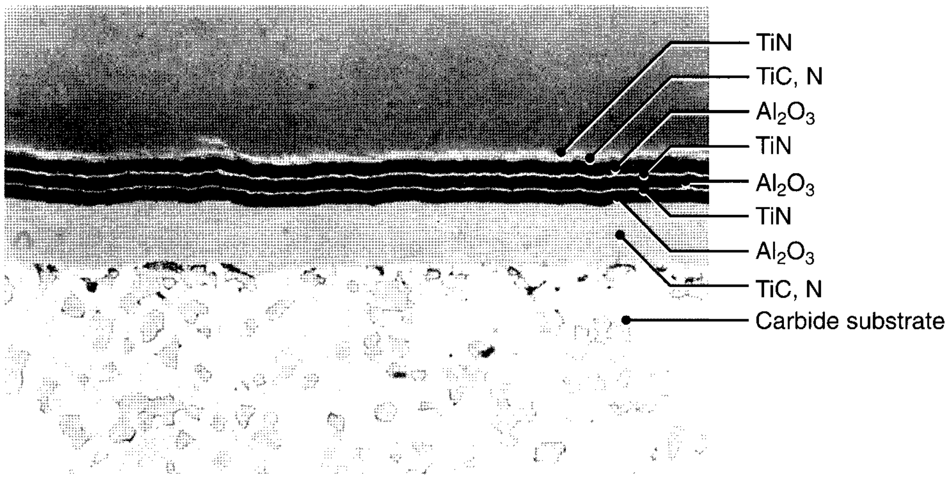


FIGURE 22.8 Multiphase coatings on a tungsten-carbide substrate. Three alternating layers of aluminum oxide are separated by very thin layers of titanium nitride. Inserts with as many as 13 layers of coatings have been made. Coating thicknesses are typically in the range from 2 to 10 μm. *Source*: Courtesy of Kennametal Inc.

is being used widely as a coating for cutting tools, particularly on tungsten-carbide and silicon-nitride inserts. Diamond-coated tools are particularly effective in machining nonferrous metals, abrasive materials such as aluminum alloys containing silicon, fiber-reinforced and metal-matrix composite materials, and graphite. As many as tenfold improvements in tool life have been obtained over the lives of other coated tools.

Diamond-coated inserts are available commercially. Thin films are deposited on substrates through PVD or CVD techniques. Thick films are obtained 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* also are being developed, with interlocking layers of diamond that give strength to the coating. As with all coatings, it is important to develop good adherence of the diamond film to the substrate and minimize the difference in thermal expansion between the diamond and substrate materials.

22.5.2 Miscellaneous Coating Materials

Major advances have been taking place in further improving the performance of coated tools. **Titanium carbonitride** (TiCN) and **titanium-aluminum nitride** (TiAlN) are effective in cutting stainless steels. TiCN (which is deposited through 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. Chromium-based coatings, such as **chromium carbide** (CrC), have been found to be effective in machining softer metals that tend to adhere to the cutting tool, such as aluminum, copper, and titanium. Other new materials include **zirconium nitride** (ZrN) and **hafnium nitride** (HfN). Further experimental data are required before these coatings and their behavior can be fully assessed for proper applications in machining.

More recent developments are (a) **nanolayer coatings,** including carbide, boride, nitride, oxide, or some combination thereof (see also Section 8.8), and (b) **composite coatings,** using a variety of materials. The hardness of some of these coatings approaches that of cubic boron nitride. Although they are still in experimental stages, it is expected that they will have the combined benefits of various types of coatings, as well as wider applications in machining operations.

22.5.3 Ion Implantation

In this process, ions are introduced into the surface of the cutting tool, improving its surface properties (Section 34.7). The process does not change the dimensions of tools. **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** (see also Section 8.2). They are cold pressed into insert shapes under high pressure and 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, so they have less tendency to adhere to metals during cutting and a correspondingly lower tendency to form a built-up edge. Consequently, in cutting cast irons and steels, good surface finish is obtained with ceramic tools. However, ceramics lack toughness, and their use may result in premature tool failure by chipping or in catastrophic failure.

Ceramic inserts are available in shapes similar to those of carbide inserts. They are effective in high-speed, uninterrupted cutting operations, such as finishing or semifinishing by turning. To reduce thermal shock, cutting should be performed either dry or with a copious amount of cutting fluid applied in a steady stream. Improper or intermittent applications of the fluid can cause thermal shock and fracture of the ceramic tool.

Ceramic tool shape and setup are important. Negative rake angles (large included angles) generally are preferred in order 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 work-holding devices, thus reducing vibration and chatter.

Cermets. *Cermets* (from the words *cer*amic and *met*al) were first used in the early 1950s and consist of ceramic particles in a metallic matrix. They were introduced in the 1960s and are *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. Further refinements of these tools have resulted in improved

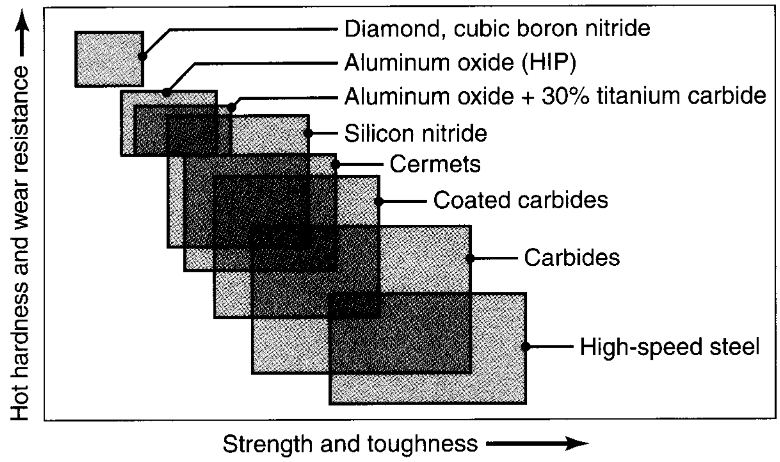


FIGURE 22.9 Ranges of mechanical properties for various groups of tool materials. HIP means hot isostatically pressed. (See also Tables 22.1 through 22.5.)

strength, toughness, and reliability. 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. Although cermets can be coated, the benefits of coated cermets are somewhat controversial, as the improvement in wear resistance appears to be marginal.

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). Cubic-boron-nitride tools also are made in small sizes without a substrate.

At elevated temperatures, cBN is chemically inert to iron and nickel. (Hence, there is no wear due to diffusion.) Its resistance to oxidation is high; thus, it is particularly suitable for cutting hardened ferrous and high-temperature alloys (see *hard machining*, Section 25.6) and for high-speed machining operations (see Section 25.5). It also is used as an abrasive. Because cBN tools are brittle, stiffness of the machine tool and the fixturing is important to avoid vibration and chatter. Furthermore, in order to avoid chipping and cracking due to thermal shock, machining generally should be performed dry (i.e., cutting fluids should be avoided), particularly in interrupted cutting operations (such as milling), which repeatedly subject the tool to thermal cycling.

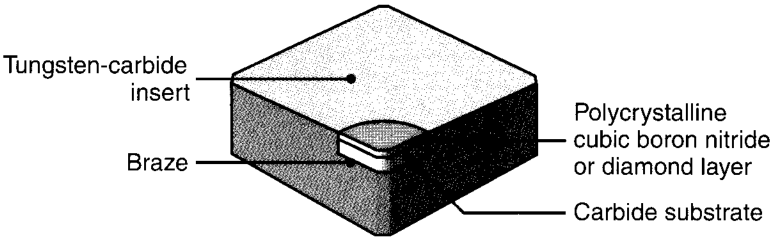


FIGURE 22.10 An insert of a polycrystalline cubic boron nitride or a diamond layer on tungsten carbide.

22.8 Silicon-nitride-based Ceramics

Developed in the 1970s, *silicon-nitride* (SiN) *based ceramic* tool 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 of which it is composed: *si*licon, *al*uminum, oxygen, and *n*itrogen. Sialon has higher thermal-shock resistance than silicon nitride and 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.

22.9 Diamond

Of all known materials, the hardest substance is diamond, described earlier in Section 8.7. As a cutting tool, it has highly desirable properties, such as low friction, high wear resistance, and the ability to maintain a sharp cutting edge. Diamond is used when a good surface finish and dimensional accuracy are required, particularly

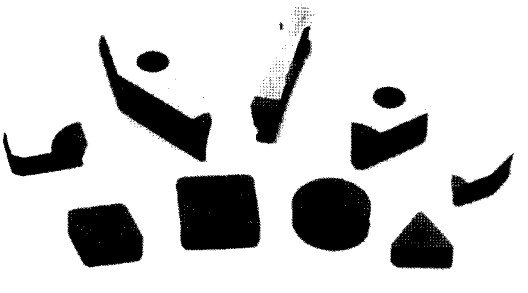


FIGURE 22.11 Inserts with polycrystalline cubic boron nitride tips (top row), and solid-polycrystalline cBN inserts (bottom row). *Source*: Courtesy of Valenite.

with soft nonferrous alloys and abrasive nonmetallic and metallic materials (especially some aluminum–silicon alloys). *Synthetic* or industrial *diamonds* are widely used because natural diamond has flaws and its performance can be unpredictable, as is the case with abrasives used in grinding wheels.

Single-crystal diamonds of various carats can be used for special applications. However, they have been replaced largely by **polycrystalline diamond** (PCD) tools, called *compacts*, which also are used as dies for fine wire drawing. These diamond tools consist of very small synthetic crystals fused by 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. 22.11). The random orientation of the diamond crystals prevents the propagation of cracks through the structure, significantly improving its toughness.

Because diamond is brittle, tool shape and sharpness are important. Low rake angles generally are used to provide a strong cutting edge (because of the larger included angles). Special attention should be given to proper mounting and crystal orientation in order to obtain 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 cutting). Diamond tools can be used satisfactorily at almost any speed, but are most suitable for light, uninterrupted finishing cuts. In order to minimize tool fracture, the single-crystal diamond must be resharpened as soon as it becomes dull. Because of its strong chemical affinity at elevated temperatures (resulting in diffusion), diamond is not recommended for machining plain-carbon steels or titanium, nickel, and cobalt-based alloys. Diamond also is used as an abrasive in grinding and polishing operations and as *coatings*.

22.10 Whisker-reinforced Materials and Nanomaterials

In order to further improve the performance and wear resistance of cutting tools (particularly in machining new materials and composites), continued progress is being made in developing new tool materials with enhanced properties such as the following:

- High fracture toughness
- Resistance to thermal shock
- Cutting-edge strength
- Creep resistance
- Hot hardness.

Advances include the use of *whiskers* as reinforcing fibers in composite cutting-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. However, the high reactivity of silicon carbide with ferrous metals 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, nanomaterials are applied as a thin coating, usually in an attempt to obtain a reasonable tool life without the use of a coolant (dry machining, Section 22.12.1) or to machine at high speeds (see Section 25.5). Note also that carbide cutting tools are nanomaterials, but since their widespread use predates the practice of classifying materials as nanomaterials, they are often overlooked.

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 cost for a typical 12.5-mm *insert* is approximately (a) $5 to $10 for uncoated carbides, (b) $6 to $10 for coated carbides, (c) $8 to $15 for ceramics, (d) $50 to $60 for diamond-coated carbides, (e) $60 to $100 for cubic boron nitride, and (f) $90 to $125 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 cutting tool, for example, can perform a large amount of material removal before it is indexed and eventually recycled. Recall from Section 21.5 that the expected tool life can be in the range of 30 to 60 minutes. Thus, considering that a square insert has eight cutting edges, it is indicated that the tool lasts a long time before it is removed from the machine and replaced.

Cutting tools can be **reconditioned** by resharpening them, using tool and cutter grinders with special fixtures. This operation may be carried out by hand or on computer-controlled tool and cutter grinders. Advanced methods of machining also are used, as described in Chapter 27. Reconditioning of coated tools also is done by recoating them, usually in special facilities for these purposes. It is important that reconditioned tools have the same geometric features as the original tools. Often, a decision has to be made whether further reconditioning of tools is economical, especially when the costs of typical small inserts are not a major contribution to total costs. Recycling 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 have been used extensively in machining operations to achieve the following results:

- Reduce friction and wear, thus improving tool life and the surface finish of the workpiece.
- Cool the cutting zone, thus improving tool life and reducing the temperature and thermal distortion of the workpiece.
- Reduce forces and energy consumption.
- Flush away the chips from the cutting zone, thus preventing the chips from interfering with the cutting process, particularly in operations such as drilling and tapping.
- Protect the machined surface from environmental corrosion.

Depending on the type of machining operation, the cutting fluid needed may be a **coolant**, a **lubricant**, or both. The effectiveness of cutting fluids depends on a number of 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, water is not an effective lubricant; hence, it does not reduce friction. Furthermore, it can cause oxidation (rusting) of workpieces and machine-tool components.

The need for a cutting fluid depends on the *severity* of the particular machining operation, which may be defined as the level of temperatures and forces encountered and the ability of the tool materials to withstand them, the tendency for built-up edge formation, the ease with which chips produced can be removed from the cutting zone, and how effectively the fluids can be applied to the proper region at the tool–chip interface. The relative severities of specific machining processes, in increasing order of severity, are as follows: 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. It has been shown that cutting fluids may cause the chip to become *more curly* and thus concentrate the heat closer to the tool tip, reducing tool life. More importantly, 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 thermal fatigue or thermal shock.

However, beginning with the mid-1990s, there has been a major trend toward **near-dry machining**, meaning a minimal use of cutting fluids, as well as toward *dry machining*; these trends are described in Section 22.12.1.

Cutting-fluid Action. The basic mechanisms of lubrication in metalworking operations are described in greater detail in Section 33.6. Here, the mechanisms by which cutting fluids influence machining operations are described. It is not immediately clear how a cutting fluid can penetrate the important rake face of the tool and influence a continuous cutting process. Studies have shown that the cutting fluid gains access to the tool–chip interface by seeping from the sides of the chip 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 proper wetting (surface tension) characteristics. For example, grease cannot be an effective lubricant in machining, but low-molecular-weight oils suspended in water (emulsions) are very effective. Note that discontinuous cutting operations have more straightforward mechanisms for lubricant application, but the tools are more susceptible to thermal shock.

EXAMPLE 22.2 Effects of Cutting Fluids on Machining

A machining operation is being carried out with a cutting fluid that is an effective lubricant. What will be the changes in the mechanics of the cutting operation if the fluid is shut off?

Solution Since the cutting fluid is a good lubricant, the following chain of events will take place after the fluid is shut off:

a. Friction at the tool–chip interface will increase.
b. The shear angle will decrease in accordance with Eq. (21.3).
c. The shear strain will increase, as seen from Eq. (21.2).
d. The chip will become thicker.
e. A built-up edge is likely to form.

As a result of these changes, the following events will occur:

a. The shear energy in the primary zone will increase.
b. The frictional energy in the secondary zone will increase.
c. The total energy will increase.
d. The temperature in the cutting zone will rise, causing greater tool wear.
e. Surface finish will begin to deteriorate and dimensional tolerances may be difficult to maintain because of the increased temperature and thermal expansion of the workpiece during machining.

Types of Cutting Fluids. The characteristics and applications of metalworking fluids are described in Section 33.7. Briefly, four general types of cutting fluids are commonly used in machining operations:

1. **Oils** (also called *straight oils*), including mineral, animal, vegetable, compounded, and synthetic oils, typically are used for low-speed operations where temperature rise is not significant.
2. **Emulsions** (also called *soluble oils*), a mixture of oil and water and additives, generally are used for high-speed operations because the temperature rise is significant. The presence of water makes emulsions highly effective coolants. The presence of oil reduces or eliminates the tendency of water to cause oxidation.
3. **Semisynthetics** are chemical emulsions containing little mineral oil, diluted in water, and with additives that reduce the size of oil particles, making them more effective.
4. **Synthetics** are chemicals with additives, diluted in water, and containing no oil.

Because of the complex interactions among the cutting fluid, the workpiece materials, temperature, time, and cutting-process variables, the application of fluids cannot be generalized. In Chapters 23 and 24, fluid recommendations for specific machining operations are given; these recommendations also are available from numerous suppliers of metalworking fluids.

Methods of Cutting-fluid Application. There are four basic methods of cutting-fluid applications in machining:

1. **Flooding.** This is the most common method (shown in Fig. 22.12, indicating good and poor flooding practices). 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 interfering with the operation.

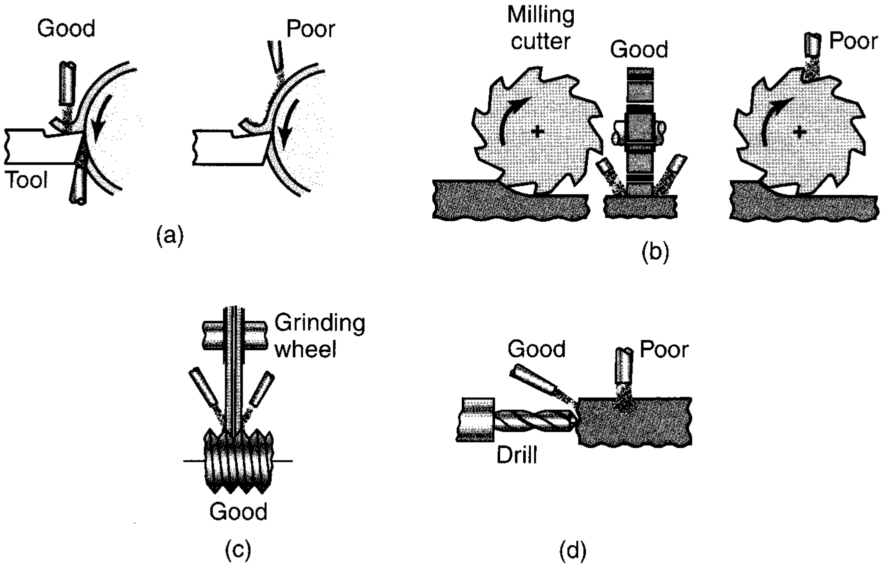


FIGURE 22.12 Schematic illustration of the proper methods of applying cutting fluids (flooding) in various machining operations: (a) turning, (b) milling, (c) thread grinding, and (d) drilling.

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 (compared with flood cooling). It is effective particularly with water-based fluids at air pressures ranging from 70 to 600 kPa. However, it has limited cooling capacity. Mist application requires venting to prevent the inhalation of airborne fluid particles by the machine operator and others nearby.
3. **High-pressure systems.** With the increasing speed and power of modern computer-controlled machine tools, heat generation in machining has become a significant factor. Particularly effective is the use of high-pressure refrigerated coolant systems to increase the rate of heat removal from the cutting zone. High pressures also are used in delivering the cutting fluid via specially designed nozzles that aim a powerful jet of fluid to the zone, particularly into the *clearance* or *relief face* of the tool (see Fig. 21.3). The pressures employed, which are usually in the range from 5.5 to 35 MPa, act as a chip breaker in situations where the chips produced would otherwise be long and continuous, interfering with the cutting operation. In order to avoid damage to the workpiece surface by impact from any particles present in the high-pressure jet, contaminant size in the coolant should not exceed 20 μm. Proper and continuous filtering of the fluid also is essential to maintain quality.
4. **Through the cutting tool system.** The severity of various machining operations have been described in terms of the difficulty of supplying fluids into the cutting zone and flushing away the chips. For a more effective application, narrow passages can be produced in cutting tools, as well as in toolholders, through which cutting fluids can be applied under high pressure. Two applications of this method are (a) gun drilling, shown in Fig. 23.22, with a long, small hole through the body of the drill itself, and (b) boring bars, shown in Fig. 23.17a, where there is a long hole through the shank (toolholder), to which an insert is clamped. Similar designs have been developed for cutting tools and inserts and for delivering cutting fluids 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 and machine tools
- Biological considerations
- The environment.

In selecting a cutting fluid, one should consider whether the machined component will be subjected to stress and adverse effects, possibly leading to stress-corrosion cracking. This consideration is particularly important for cutting fluids with sulfur and chlorine additives. For example, (a) cutting fluids containing sulfur should not be used with nickel-based alloys, and (b) fluids containing chlorine should not be used with titanium. Cutting fluids also may affect adversely the machine tool components; thus, their compatibility with various metallic and nonmetallic materials in the machine tool must be considered. Machined parts should be cleaned and washed in order to remove any cutting-fluid residue, as described in Section 34.16. This operation can be significant in time and cost. Consequently, the trend is to use water-based, low-viscosity fluids for ease of cleaning and filtering.

Because the machine-tool operator is usually in close proximity to cutting fluids, the health effects of operator contact with fluids 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 chemical constituents such as sulfur, chlorine, phosphorus, hydrocarbons, biocides, and various additives. Much progress has been made in ensuring the safe use of cutting fluids in manufacturing plants, including reducing or eliminating their use by adopting the more recent trends in dry or near-dry machining techniques (see Section 21.12.1).

Cutting fluids (as well as other metalworking fluids used in manufacturing operations) may undergo chemical changes as they are used repeatedly over time. These changes may be due to environmental effects or to contamination from various sources, including metal chips, fine particles produced during machining, and tramp oil (from leaks in hydraulic systems, oils on sliding members of machines, and lubricating systems for the machine tools). The changes involve the growth of microbes (bacteria, molds, and yeast), particularly in the presence of water, becoming an environmental hazard and also adversely affecting the characteristics and effectiveness of the cutting fluids.

Several techniques (such as settling, skimming, centrifuging, and filtering) are available for clarifying used cutting fluids. Recycling involves treatment of the fluids with various additives, agents, biocides, and deodorizers, as well as water treatment (for water-based fluids). Disposal practices for these fluids must comply with federal, state, and local laws and regulations.

22.12.1 Near-dry and Dry Machining

For economic and environmental reasons, there has been a continuing worldwide trend since the mid-1990s to minimize or eliminate the use of metalworking fluids. This trend has led to the practice of *near-dry machining* (NDM), with major benefits such as the following:

- 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.
- Further improving surface quality.

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/hr, 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).

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, but generally not for aluminum alloys.

One of the functions of a metal-cutting fluid is to flush chips from the cutting zone. This function seems to be problematic with dry machining. However, tool designs have been developed that allow the application of pressurized air, often through the tool shank. The compressed air doesn't serve a lubrication purpose and provides only limited cooling, but is very effective at clearing chips from the cutting interface.

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 is injected into the cutting zone. Because of the reduction in temperature, tool hardness is maintained and tool life is enhanced, thus allowing higher cutting speeds. Also, the chips are more brittle; hence, machinability is increased. Furthermore, the nitrogen simply evaporates and therefore has no adverse environmental impact.

SUMMARY

- A wide variety of cutting-tool materials have been developed over the past century, the most commonly used ones being high-speed steels, carbides, ceramics, cubic boron nitride, and diamond. Tool materials have a broad range of mechanical and physical properties, such as hot hardness, toughness, chemical stability and inertness, and resistance to chipping and wear.
- Various tool coatings have been developed with 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; diamond coatings also are gaining acceptance. 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 process parameters and the characteristics of the machine tool.
- Cutting fluids are important in machining operations and reduce friction, wear, cutting forces, and power requirements. Generally, slower cutting operations and those with high cutting-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
Carbides
Cast-cobalt alloys
Ceramics
Cermets
Chemical stability
Chip breaker
Coated tools
Coolants
Cryogenic machining
Cubic boron nitride
Cutting fluids
Diamond coatings
Diamond tools
Dry machining
Finishing cuts
Flooding
High-speed steels
Inserts
Lubricants
Micrograin carbides
Mist
Multiphase coatings
Nanocrystalline
Near-dry machining
Polycrystalline cubic boron nitride
Polycrystalline diamond
Reconditioning of tools
Roughing cuts
Sialon
Silicon-nitride-based ceramics
Stellite
Titanium carbide
Titanium nitride
Tool costs
Tool reconditioning
Toughness
Tungsten carbide
Uncoated carbides
Wear resistance
Whisker-reinforced tools

BIBLIOGRAPHY

ASM Handbook, Vol. 16: **Machining**, ASM International, 1989.

ASM Handbook, Vol. 3: **Properties and Selection: Stainless Steels, Tool Materials and Special Purpose Metals**, ASM International, 1980.

ASM Specialty Handbook: **Tool Materials**, ASM International, 1995.

Astakhov, V., **Tribology of Metal Cutting**, Elsevier, 2007.

Byers, J.P. (ed.), **Metalworking Fluids**, 2nd ed., CRC Press, 2006.

Davim, J.P. (ed.), **Machining: Fundamentals and Recent Advances**, Springer, 2008.

Davis, J.R. (ed.), **Tool Materials**, ASM International, 1995.

Jackson, M.J., and Morrell, J., **Machining with Nanomaterials**, Springer, 2009.

Komanduri, R., **Tool Materials**, in *Kirk–Othmer Encyclopedia of Chemical Technology*, 4th ed., Vol. 24, 1997.

Machinery's Handbook, Industrial Press, revised periodically.

Metal Cutting Tool Handbook, 7th ed., Industrial Press, 1989.

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.

Sluhan, C. (ed.), **Cutting and Grinding Fluids: Selection and Application**, Society of Manufacturing Engineers, 1992.

Smith, G.T., **Cutting Tool Technology: Industrial Handbook**, Spinger, 2008.

Stephenson, D.A., and Agapiou, J.S., **Metal Cutting Theory and Practice**, 2nd ed., CRC Press, 2006.

Trent, E.M., and Wright, P.K., **Metal Cutting**, 4th ed., Butterworth-Heinemann, 1999.

Tool and Manufacturing Engineers Handbook, 4th ed., Vol. 1: **Machining**, Society of Manufacturing Engineers, 1983.

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. Explain how cutting fluids penetrate the tool–chip interface.

22.8. List the methods by which cutting fluids are typically applied in machining operations.

22.9. Describe the advantages and limitations of (a) single-crystal and (b) polycrystalline diamond tools.

22.10. What is a cermet? What are its advantages?

22.11. Explain the difference between M-series and T-series high-speed steels.

QUALITATIVE PROBLEMS

22.12. Explain why so many different types of cutting-tool materials have been developed over the years. Why are they still being developed further?

22.13. Which tool-material properties are suitable for interrupted cutting operations? Why?

22.14. Describe the reasons for and advantages of coating cutting tools with multiple layers of different materials.

22.15. 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.

22.16. As stated in Section 22.1, tool materials can have conflicting properties when used for machining operations. Describe your observations regarding this matter.

22.17. Explain the economic impact of the trend shown in Fig. 22.6.

22.18. Why does temperature have such an important effect on tool life?

22.19. Ceramic and cermet cutting tools have certain advantages over carbide tools. Why, then, are they not completely replacing carbide tools?

22.20. Can cutting fluids have any adverse effects in machining? If so, what are they?

22.21. Describe the trends you observe in Table 22.2.

22.22. Why are chemical stability and inertness important in cutting tools?

22.23. Titanium-nitride coatings on tools reduce the coefficient of friction at the tool–chip interface. What is the significance of this property?

22.24. Describe the necessary conditions for optimal utilization of the capabilities of diamond and cubic-boron-nitride cutting tools.

22.25. Negative rake angles generally are preferred for ceramic, diamond, and cubic-boron-nitride tools. Why?

22.26. Do you think that there is a relationship between the cost of a cutting tool and its hot hardness? Explain.

22.27. 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.

22.28. In Table 22.1, the last two properties listed can be important to the life of a cutting tool. Why?

22.29. 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.

22.30. 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.

22.31. Referring to Table 22.1, state which tool materials would be suitable for interrupted cutting operations. Explain.

22.32. Which of the properties listed in Table 22.1 is, in your opinion, the least important in cutting tools? Explain.

22.33. 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.

22.34. What are the consequences of a coating on a tool having a different coefficient of thermal expansion than the substrate material?

22.35. Discuss the relative advantages and limitations of near-dry machining. Consider all relevant technical and economic aspects.

QUANTITATIVE PROBLEMS

▶**22.36.** 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.

22.37. 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).

▶**22.38.** The first column in Table 22.2 shows ten 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

22.39. Describe in detail your thoughts regarding the technical and economic factors involved in tool-material selection.

22.40. 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.

22.41. How would you go about measuring the effectiveness of cutting fluids? Describe your method and explain any difficulties that you might encounter.

22.42. 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.

22.43. 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.

22.44. 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.

Machining Processes: Turning and Hole Making

CHAPTER 23

- With the preceding two chapters as background, this chapter describes machining processes (turning) that are capable of generating round external or internal shapes.
- The most common machine tool used for such operations is the lathe, which is available in several types.
- Components and characteristics of lathes are examined.
- 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, powder metallurgy, powder injection molding, abrasive machining, thread rolling and rotary swaging.

EXAMPLES:

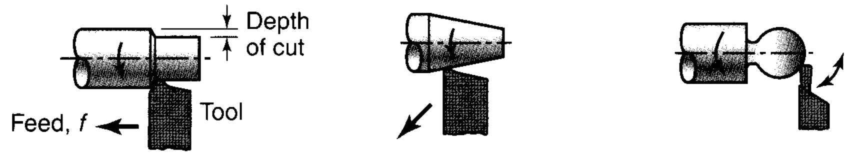
CASE STUDY:

23.1 Introduction

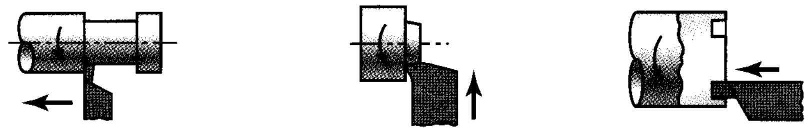
This chapter describes machining processes with the capability of producing parts that basically are round in shape. Typical products made are as small as miniature screws for the hinges of eyeglass frames and as large as turbine shafts for hydroelectric power plants, rolls for rolling mills, and gun barrels.

One of the most basic machining processes is *turning*, meaning that the part is rotated while it is being machined. The starting material is generally a workpiece that has been made by other processes, such as casting, forging, extrusion, drawing, or powder metallurgy, as described in Parts II and III. Turning processes, 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 a number of machining processes that produce a wide variety of shapes as the following list indicates:

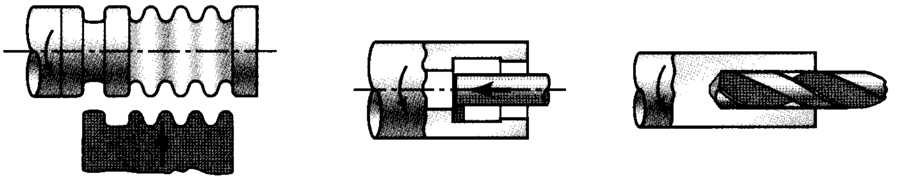
- **Turning:** to produce straight, conical, curved, or grooved workpieces (Figs. 23.1a through d), such as shafts, spindles, and pins.



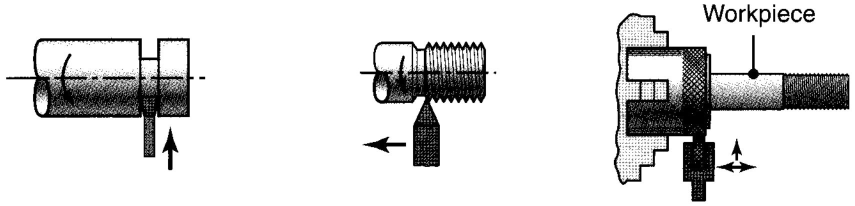
(a) Straight turning (b) Taper turning (c) Profiling



(d) Turning and external grooving (e) Facing (f) Face grooving



(g) Cutting with a form tool (h) Boring and internal grooving (i) Drilling



(j) Cutting off (k) Threading (l) Knurling

FIGURE 23.1 Miscellaneous cutting operations that can be performed on a lathe. Note that all parts are circular—a property known as axisymmetry. The tools used, their shape, and the processing parameters are described throughout this chapter.

- **Facing:** to produce a flat surface at the end of the part and perpendicular to its axis (Fig. 23.1e), useful for parts that are assembled with other components. *Face grooving* produces grooves for applications such as O-ring seats (Fig. 23.1f).
- **Cutting with form tools:** (Fig. 23.1g) to produce various axisymmetric shapes for functional or 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 may be followed by boring to improve its dimensional accuracy and surface finish.
- **Parting:** also called **cutting off**, to cut a piece from the end of a part, as is done in the production of slugs or blanks for additional processing into discrete products (Fig. 23.1j).
- **Threading:** to produce external or internal threads (Fig. 23.1k).
- **Knurling:** to produce a regularly shaped roughness on cylindrical surfaces, as in making knobs and handles (Fig. 23.1l).

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	0.013–0.025
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	0.05–0.08
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

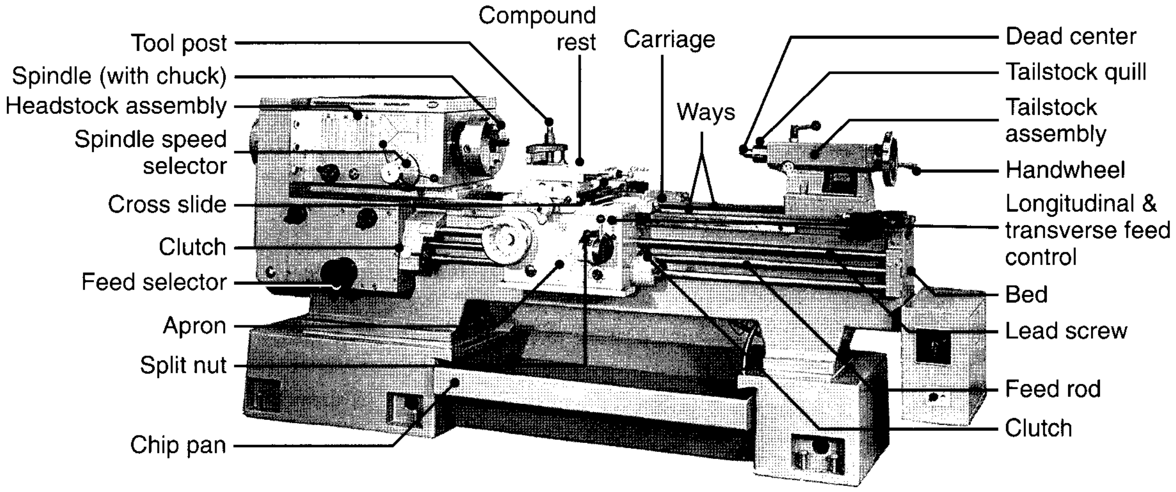
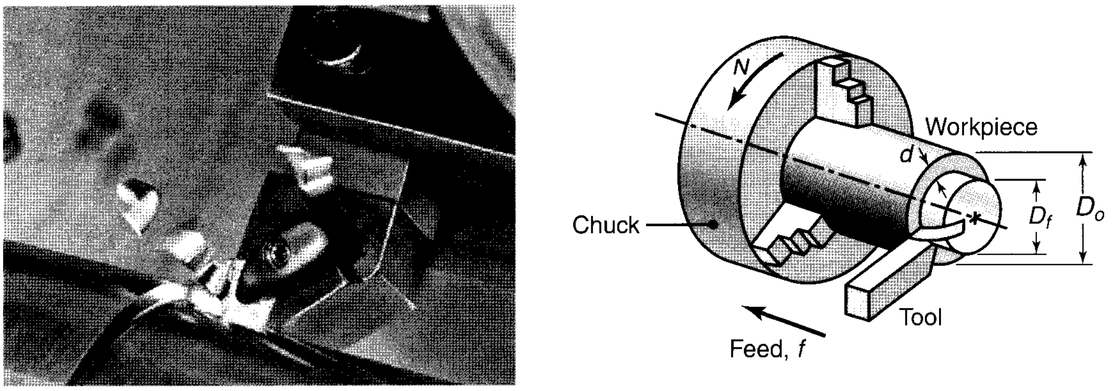


FIGURE 23.2 General view of a typical lathe, showing various components. *Source*: Courtesy of Heidenreich & Harbeck.

The cutting operations just summarized typically are performed on a **lathe** (Fig. 23.2), which is available in a variety of designs, sizes, capacities, and computer-controlled features (as discussed in Section 23.3 and Chapter 25). As shown in Figs. 21.2 and 23.3, turning is performed at various (1) rotational speeds, N, of the workpiece clamped in a spindle, (2) depths of cut, d, and (3) feeds, f, depending on the workpiece materials, cutting-tool materials, surface finish and dimensional accuracy required, and characteristics of the machine tool.



(a) (b)

FIGURE 23.3 (a) Photograph of a turning operation, showing insert and chip removal. The machine tool is traveling from left to right 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.

This chapter describes the process parameters, cutting tools, process capabilities, and characteristics of the machine tools that are used to produce a variety of parts with round shapes. Finally, design considerations to improve productivity for each group of processes are given.

23.2 The Turning Process

The majority of turning operations involve the use of simple single-point cutting tools, with the geometry of a typical right-hand cutting tool shown in Figs. 21.10 and 23.4. As can be seen, such tools are described by a standardized nomenclature. Each group of workpiece materials has an optimum set of tool angles, which have been developed largely through experience (Table 23.2).

The important process parameters that have a direct influence on machining processes and the importance of controlling these parameters for optimized productivity were described in Chapter 21. This section outlines turning-process parameters such as tool geometry and material-removal rate and gives data for recommended cutting practices, including cutting-tool materials, depth-of-cut, feed, cutting speed, and use of cutting fluids.

Tool Geometry. The various angles in a single-point cutting tool have important functions in machining operations. These angles are measured in a coordinate system consisting of the three major axes of the tool shank, as can be seen in Fig. 23.4. Note, however, that these angles may be different, with respect to the workpiece, after the tool is installed in the toolholder.

- **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. However, positive angles also result in a small included angle of the tool tip (as in Figs. 21.3 and 23.4), possibly leading to premature tool chipping and failure, depending on the toughness of the tool material.

TABLE 23.2

General Recommendations for Tool Angles in Turning

Material	High-speed steel Back rake	Side rake	End relief	Side relief	Side and end cutting edge	Carbide inserts 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

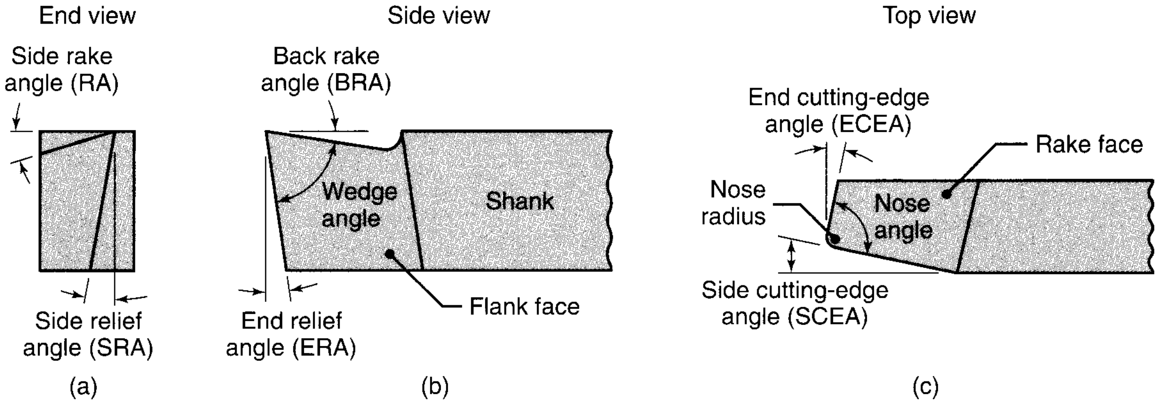


FIGURE 23.4 Designations for a right-hand cutting tool. *Right-hand* means that the tool travels from right to left, as shown in Fig. 23.3b.

- **Side rake angle** is more important than the **back rake angle**, although the latter usually controls the direction of chip flow. For machining metals and using carbide inserts, these angles typically are in the range from −5° to 5°.
- **Cutting-edge angle** affects chip formation, tool strength, and cutting forces to various degrees. Typically, the cutting-edge 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. Relief angles typically are 5°.
- **Nose radius** affects surface finish and tool-tip strength. The smaller the nose radius (sharp tool), the rougher the surface finish of the workpiece and the lower the strength of the tool. However, large nose radii can lead to tool *chatter*, as described in Section 25.4.

Material-removal Rate. The *material-removal rate* (MRR) in turning is the volume of material removed per unit time, with the units of mm^3/min. Referring to Figs. 21.2 and 23.3, note that, for each revolution of the workpiece, a ring-shaped layer of material is removed which has a cross-sectional area that equals the product of the distance the tool travels in one revolution (feed, f) and the depth of cut, d. The volume of this ring is the product of the cross-sectional area $(f)(d)$ and the average circumference of the ring, 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

$$\text{MRR} = \pi D_{\text{avg}} d f N. \tag{23.1a}$$

The dimensional accuracy of this equation can be checked by substituting dimensions into the right-hand side. For instance, (mm)(mm)(mm/rev)(rev/min) = mm^3/min, which indicates volume rate of removal. Note that Eq. (23.1a) also can be written as

$$\text{MRR} = d f V \tag{23.1b}$$

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 = (mm/rev)(rev/min) = mm/min. Since the distance traveled is l mm, the cutting time is

$$t = \frac{l}{fN}. \tag{23.2}$$

The cutting time in Eq. (23.2) does not include the time required for *tool approach* and *retraction*. Because the time spent in noncutting cycles of a machining operation is unproductive and adversely affects the overall economics, the time involved in approaching and retracting tools to and from the workpiece is an important consideration. Machine tools are designed and built to minimize this time. One method of accomplishing this aim is to rapidly traverse the tools during noncutting cycles, followed by a slower movement as the tool engages the workpiece.

The foregoing equations and the terminology used are summarized in Table 23.3.

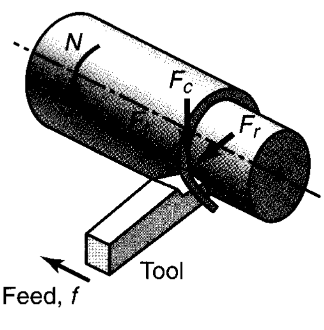


FIGURE 23.5 Forces acting on a cutting tool in turning. F_c is the cutting force, F_t is the thrust or feed force (in the direction of feed), and F_r is the radial force that tends to push the tool away from the workpiece being machined.

Forces in Turning. The three principal forces acting on a cutting tool 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 for precision-machining operations. The machine tool and its components must be able to withstand such forces without causing significant deflections, vibrations, and chatter in the overall operation.

The **cutting force**, F_c, acts downward on the tool tip and thus tends to deflect the tool downward and the workpiece upward. The cutting force supplies the energy required for the cutting operation, and it can be calculated with the data given in Table 21.2, from the energy per unit volume, described in Section 21.3. The product of the cutting force and its radius from the workpiece center determines the *torque* on the spindle. The product of the torque and the spindle speed determines the *power* required in the turning operation.

The **thrust force**, F_t, acts in the longitudinal direction. It also is called the **feed force**, because it is in the feed direction of the tool. This force tends to push the tool towards the right and away from the chuck. The **radial force**, F_r, acts in the radial direction and tends to push the tool away from the workpiece. Because of the many

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)
$= \pi D_{avg} N$ (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 or min
$= l/fN$
MRR = mm^3/min
$= \pi D_{avg} d f N$
Torque = N · m
$= F_c D_{avg}/2$
Power = kW or hp
= (Torque)(ω), where $\omega = 2\pi N$ rad/min

Note: The units given are those which are commonly used; however, appropriate units must be used and checked in the formulas.

factors involved in the cutting process, forces F_t and F_r are difficult to calculate directly; they usually are determined experimentally if desired.

Roughing and Finishing Cuts. In machining, the usual procedure is to first perform one or more *roughing cuts* at high feed rates and large depths of cut (and therefore high material-removal rates), but with little consideration for dimensional tolerance and surface roughness. These cuts are then followed by a *finishing cut*, at a lower feed and depth of cut in order to produce 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 these 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.

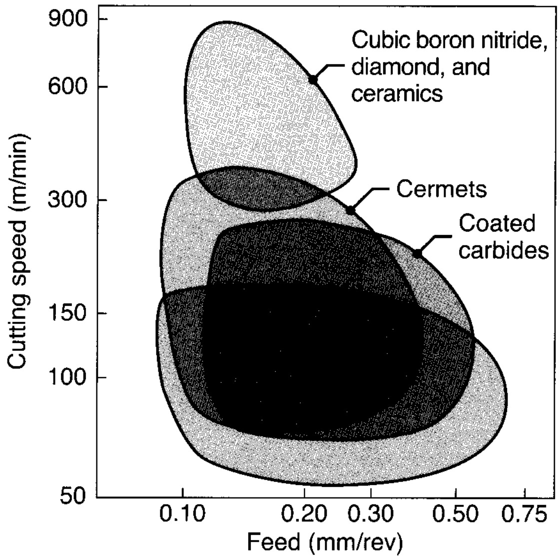


FIGURE 23.6 The range of applicable cutting speeds and feeds for a variety of tool materials.

TABLE 23.4

General Recommendations for Turning Operations

		General-purpose starting conditions			Range for roughing and finishing		
Workpiece material	Cutting tool	Depth of cut, mm	Feed, mm/rev	Cutting speed, m/min	Depth of cut, mm	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	"	"	120–425
	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	"	"	105–455
Medium and high-C steels	Uncoated carbide	1.2–4.0	0.30	75	2.5–7.6	0.15–0.75	45–120
	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

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	"	"	105–290
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 car bide	1.0–3.8	0.15	35–60	0.25–6.3	0.1–0.4	10–75
	TiN-coated carbide	"	"	30–60	"	"	10–100
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	"	"	760	"	"	305–3050
High silicon	Polycrystalline diamond	"	"	530	"	"	365–915

(continued)

TABLE 23.4 *(continued)*

General Recommendations for Turning Operations

		General-purpose starting conditions			Range for roughing and finishing		
Workpiece material	Cutting tool	Depth of cut, mm	Feed, mm/rev	Cutting speed, m/min	Depth of cut, mm	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	"	"	275–915
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	"	"	150–730
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.

TABLE 23.5

General Recommendations for Cutting Fluids for Machining (see also Section 33.7)

Material	Type of fluid
Aluminum	D, MO, E, MO + FO, CSN
Beryllium	MC, E, CSN
Copper	D, E, CSN, MO + FO
Magnesium	D, MO, MO + FO
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 = chemicals and synthetics; D = dry; E = emulsion; EP = extreme pressure; FO = fatty oil; and MO = mineral oil.

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 improve the operation significantly. General recommendations for cutting fluids appropriate to various workpiece materials are given in Table 23.5. However, there is a major current trend towards **near-dry** and **dry machining**, with important benefits, as described in Section 22.12.1.

EXAMPLE 23.1 Material-removal Rate and Cutting Force in Turning

A 150-mm-long, 12.5-mm-diameter 304 stainless-steel rod is being reduced in diameter to 12.0 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. 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,

$$V = \frac{(\pi)(12.5)(400)}{1000} = 15.7 \text{ m/min}$$

The cutting speed at the machined diameter is

$$V = \frac{(\pi)(12.0)(400)}{1000} = 15.1 \text{ m/min}$$

From the information given, note that the depth of cut is

$$d = \frac{12.5 - 12.0}{2} = 0.25 \text{ mm}$$

and the feed is

$$f = \frac{200}{400} = 0.5 \text{ mm/rev.}$$

According to Eq. (23.1a), the material-removal rate is then

$$\begin{aligned} \text{MRR} &= (\pi)(12.25)(0.25)(0.5)(400) \\ &= 1924 \text{ mm}^3\text{/min} = 2 \times 10^{-6} \text{ m}^3\text{/min} \end{aligned}$$

Equation (23.1b) also can be used, in which case we find that

$$\text{MRR} = (0.25)(0.5)(15.7)(1000) = 2 \times 10^{-6} \text{ m}^3\text{/min}$$

The actual time to cut, according to Eq. (23.2), is

$$t = \frac{150}{(0.5)(400)} = 0.75 \text{ mm}$$

The power required can be calculated by referring to Table 21.2 and taking an average value for stainless steel as $4\ \mathrm{W \cdot s/mm^3}$. Therefore, the power dissipated is

$$\text{Power} = \frac{(4)(1924)}{60} = 128\ \mathrm{W}.$$

Since $\mathrm{W} = 60\ \mathrm{N{\cdot}m/min}$, the power dissipated is 7680 N-m/min.

The cutting force, F_c, is the tangential force exerted by the tool. Power is the product of torque, T, and the rotational speed in radians per unit time; hence;

$$T = \frac{7680}{(2\pi)(400)} = 3.1\ \mathrm{N{\cdot}m}$$

since $T = F_c D_{\text{avg}}/2$,

$$F_c = \frac{(3.1)(1000)}{12.25/2} = 506\ \mathrm{N}.$$

23.3 Lathes and Lathe Operations

Lathes generally are considered to be the oldest machine tools. Although woodworking lathes originally were developed during the period from 1000 to 1 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 nearby engines on the factory floor. Modern lathes are all equipped with individual electric motors.

The maximum spindle speed of lathes typically is around 4000 rpm, but may be only about 200 rpm for large lathes. For special applications, speeds may range to 10,000 rpm, 40,000 rpm, or higher for very high speed machining. The cost of lathes ranges from about $2,000 for bench types to over $100,000 for larger units.

Although simple and versatile, an engine lathe requires a skilled machinist, because all controls are manipulated by hand. Consequently, lathes are inefficient for repetitive operations and for large production runs. The rest of this section will describe the various types of automation that usually are added to improve efficiency.

23.3.1 Lathe Components

Lathes are equipped with a variety of components and accessories, as shown in Fig. 23.2. Their features and functions are as follows:

Bed. The bed supports all major components of the lathe. Beds have a large mass and are built rigidly, usually from gray or nodular cast iron. (See also Section 25.3 on new materials for machine-tool structures.) The top portion of the bed has two **ways** with various cross sections that are hardened and machined for wear resistance and dimensional accuracy during turning. In *gap-bed lathes*, a section of the bed in front of the headstock can be removed to accommodate larger diameter workpieces.

Carriage. The carriage, or carriage assembly, slides along the ways and 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 adjustment. The *cross-slide* moves radially in and out, controlling the radial position of the cutting tool in operations such as facing (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 bed and is equipped with motors, pulleys, and V-belts that supply power to a *spindle* at various rotational speeds. The speeds 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

TABLE 23.6

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	4000
Turret	0.5/1.5	60	3000
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	1000
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.

provide a continuously variable range of speed to the spindle. Headstocks have a *hollow spindle* to which work-holding devices (such as *chucks* and *collets*; see Section 23.3.2) are mounted and long bars or tubing can be fed through them for various turning operations. The accuracy of the spindle is important for precision in turning, particularly in high-speed machining; preloaded tapered or ball bearings typically are used to rigidly support the spindle.

Tailstock. The tailstock, which can slide along the ways and be clamped at any position, supports the other end of the workpiece. It is equipped with a center that may be fixed (*dead center*), or it may be free to rotate with the workpiece (*live center*). Drills and reamers can be mounted on the tailstock *quill* (a hollow cylindrical part 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. The rod 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 it with the carriage; the split nut is also used for cutting threads accurately.

Lathe Specifications. A lathe generally is specified by the following parameters:

- Its *swing*, the maximum diameter of the workpiece that can be machined (Table 23.6).
- The maximum distance between the headstock and tailstock centers.
- The length of the bed.

For example, a lathe may have the following size: 360-mm swing by 760 mm between centers by 1830-mm length of bed. Lathes are available in a variety of styles and types of construction and power. Maximum workpiece diameters may be as much as 2 m.

23.3.2 Workholding Devices and Accessories

Workholding devices are important, particularly in machine tools and machining operations, as they must hold the workpiece securely. As shown in Fig. 23.3, one end

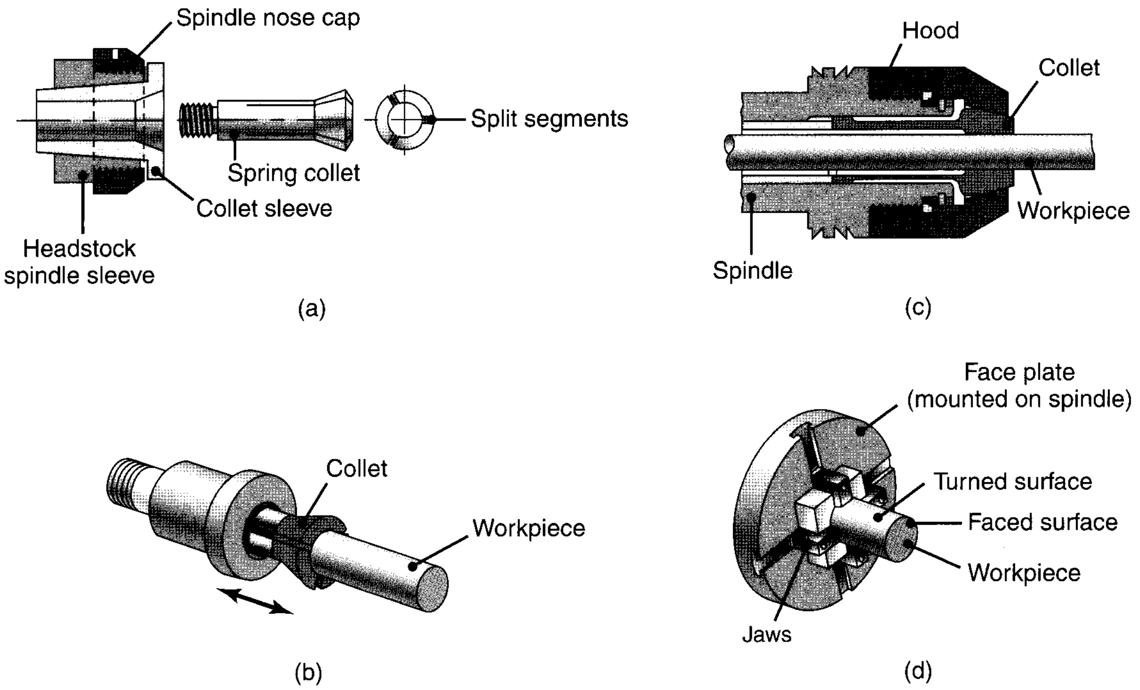


FIGURE 23.7 (a) and (b) Schematic illustrations of a draw-in type of collet. The workpiece is placed in the collet hole, and the conical surfaces of the collet are forced inward by pulling it with a draw bar into the sleeve. (c) A push-out type of collet. (d) Work holding of a workpiece on a face plate.

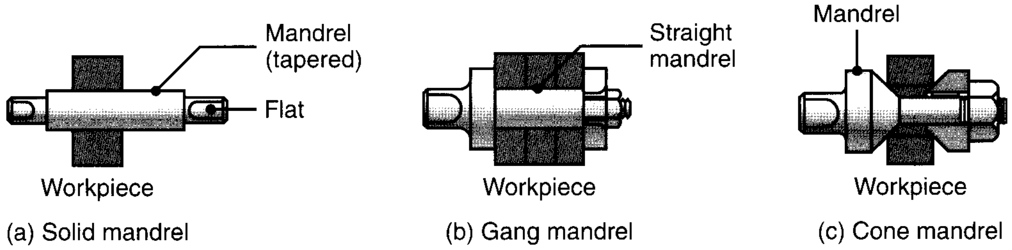
of the workpiece is clamped to the spindle of the lathe by a chuck, collet, face plate (see Fig. 23.7d), or mandrel.

A **chuck** usually is equipped with three or four *jaws*. *Three-jaw* chucks generally have a geared-scroll design that makes the jaws self-centering. They are used for round workpieces (such as bar stock, pipes, and tubing), which can be centered to within 0.025 mm. *Four-jaw* (independent) 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. Because they are constructed more ruggedly than three-jaw chucks, four-jaw chucks are used for heavy workpieces or for work requiring multiple chuckings where concentricity is important.

The jaws in some types of chucks can be reversed to permit clamping of hollow workpieces, such as pipes and tubing, either on the outside surfaces or on the inside surfaces. Also available are jaws made of low-carbon steel (*soft jaws*) that can be machined into desired shapes. Because of their low strength and hardness, soft jaws conform to small irregularities on workpieces and therefore result in better clamping. Chucks can be *power* or *manually actuated* with a chuck wrench. Because they take longer to operate, manually actuated chucks generally are used only for toolroom and limited production runs.

Power chucks, actuated pneumatically or hydraulically, are used in automated equipment for high production rates, including the loading of parts using industrial robots. Also available are several types of power chucks with lever- or wedge-type mechanisms to actuate the jaws; these chucks have jaw movements (stroke) that usually are limited to about 13 mm.

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 forces required. By controlling the magnitude of the jaw forces, an operator can ensure that the part does not slip in the chuck during



(a) Solid mandrel (b) Gang mandrel (c) Cone mandrel

FIGURE 23.8 Various types of mandrels to hold workpieces for turning. These mandrels usually are 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.

machining. High spindle speeds can reduce jaw (clamping) forces significantly due to the effect of *centrifugal forces*; this effect is particularly important in precision tube turning. Modern jaw-actuating mechanisms permit a higher clamping force for roughing, and lower force for finishing, operations.

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*; Figs. 23.7a and b) or pushed (*push-out collet*; Fig. 23.7c) mechanically into the spindle. The tapered surfaces shrink the segments of the collet radially, tightening onto the workpiece. Collets are used for round workpieces as well as for other shapes (e.g., square or hexagonal workpieces) and are available in a wide range of sizes.

One 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 the device well suited particularly for parts with small cross sections. Because the radial movement of the collet segments is small, workpieces generally should be within 0.125 mm of the nominal size of the collet.

Face plates are used for clamping irregularly shaped workpieces. The plates are round and have several slots and holes through which the workpiece is bolted or clamped (Fig. 23.7d). **Mandrels** (Fig. 23.8) are placed inside hollow or tubular workpieces and are used to hold workpieces that require machining on both ends or on their cylindrical surfaces. Some mandrels are mounted between centers on the lathe.

Accessories. Several devices are available as accessories and attachments for lathes. Among these devices are the following:

- Carriage and cross-slide stops, with various designs to stop the carriage at a predetermined distance along the bed.
- Devices for turning parts with various tapers.
- Milling, sawing, gear-cutting, and grinding attachments.
- Various attachments for boring, drilling, and thread cutting.

23.3.3 Lathe Operations

In a typical turning operation, the workpiece is clamped by any one of the workholding devices described previously. Long and slender parts must be supported by a *steady rest* and *follow rest* placed on the bed, as otherwise the part will deflect under the cutting forces. These 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 with the cross-slide and also clamping the carriage for better dimensional accuracy.

Form tools are used to produce 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 vibration and chatter may result and cause poor surface finish. As a rule, (a) the formed length of the part should not be greater than about 2.5 times the minimum diameter of the part, (b) the cutting speed should be set properly, and (c) cutting fluids should be used. The stiffnesses of the machine tools and work-holding devices also are important considerations.

Boring on a lathe is similar to turning. It is performed inside hollow workpieces or in a hole made previously by drilling or other means. Out-of-shape holes can be straightened by boring. The workpiece is held in a chuck or in some other suitable work-holding device. 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 in the tailstock quill. The workpiece is clamped in a work holder on the headstock, and the drill bit is advanced by rotating the handwheel of the tailstock. Holes drilled in this manner may not be sufficiently concentric because of the tendency for the drill to drift radially. The concentricity of the hole can be improved by subsequently boring the drilled hole. Drilled holes may be **reamed** (Section 23.6) on lathes in a manner similar to drilling, thus improving hole dimensional tolerances and surface finish.

The tools for *parting*, *grooving*, *thread cutting*, and various other operations are specially shaped for their particular purpose or are available as inserts. *Knurling* is performed on a lathe with hardened rolls (see Fig. 23.11), in which 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

There are a number of other lathe types, brief descriptions of which are given here.

Bench Lathes. As the name suggests, these lathes are placed on a workbench or a table. They have low power, are usually operated by hand feed, and are used to machine small workpieces. *Toolroom lathes* have high precision, enabling the machining of parts to close dimensional tolerances.

Special-purpose Lathes. These lathes are used for applications (such as railroad wheels, gun barrels, and rolling-mill rolls) with workpiece sizes as large as 1.7 m in diameter by 8 m in length and capacities of 450 kW.

Tracer Lathes. These lathes have special attachments that are capable of 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 plastic stencil. However, operations typically performed on a tracer lathe have been replaced largely by *numerical-control lathes* and *turning centers*, as described in Section 25.2.

Automatic Lathes. Lathes have become increasingly automated over the years; manual machine controls have been replaced by various mechanisms that enable machining operations to follow a certain prescribed sequence. In a *fully automatic lathe*, parts are fed and removed automatically, whereas in *semiautomatic* machines, these functions are performed by the operator. (The cutting remains automatic.)

Automatic lathes may have a horizontal or vertical spindle and are suitable for medium- to high-volume production.

Lathes that do not have tailstocks are called *chucking machines* or *chuckers*. They are used for machining individual pieces of regular or irregular shapes and are either single- or multiple-spindle types. In another type of automatic lathe, the bar stock is fed periodically into the lathe and a part is machined and cut off from the end of the bar stock.

Automatic Bar Machines. Also called **automatic screw machines**, these machine tools are designed for high-production-rate machining of screws and similar threaded parts. All operations on these machines are performed automatically with tools attached to a special turret. After each part or screw is machined to finished dimensions, the bar stock is fed forward automatically through the hole in the spindle and then cut off. Automatic bar machines may be equipped with single or multiple spindles. Capacities range from 3- to 150-mm diameter bar stock. Long stock is supported by special fixtures as it enters the spindle hole.

Single-spindle automatic bar machines are similar to turret lathes and are equipped with various cam-operated mechanisms. There are two types of single-spindle machines. In *Swiss-type automatics,* the cylindrical surface of the solid-bar stock is machined with 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, which minimizes deflections due to cutting forces. These machine tools are capable of high-precision machining of small-diameter parts.

The other single-spindle machine (called the *American type*) is similar to a small automatic turret lathe. The turret is on a vertical plane, and all motions of the machine components are controlled by cams. Automatic bar machines are now equipped with computer numerical control, eliminating the use of cams, and the operation is programmed for a specific product. (See Section 37.3.)

Multiple-spindle automatic bar machines typically have from four to eight spindles arranged in a circle on a large drum, with each 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. 23.9). Several cutting tools (usually as many as six) are mounted on the hexagonal *main turret*, which is rotated after each specific cutting operation is completed. The lathe usually has a *square turret* on the cross-slide, mounting as many as four cutting tools. The workpiece (generally a 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 (either the bar type or the chucking type) are versatile, and the operations may be carried out either by hand, using the turnstile (capstan wheel), or automatically. Once set up properly, these machines do not require highly skilled operators. *Vertical turret lathes* also are available; they are more suitable for short, heavy workpieces with diameters as large as 1.2 m.

The turret lathe shown in Fig. 23.9 is known as a **ram-type** turret lathe—one 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 style (called the **saddle type**),

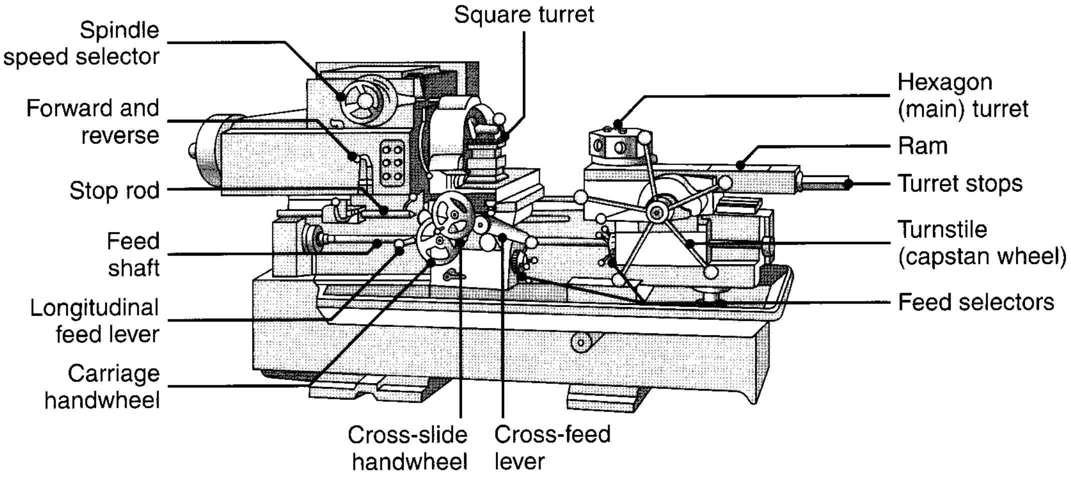


FIGURE 23.9 Schematic illustration of the components of a turret lathe. Note the two turrets: square and hexagonal (main).

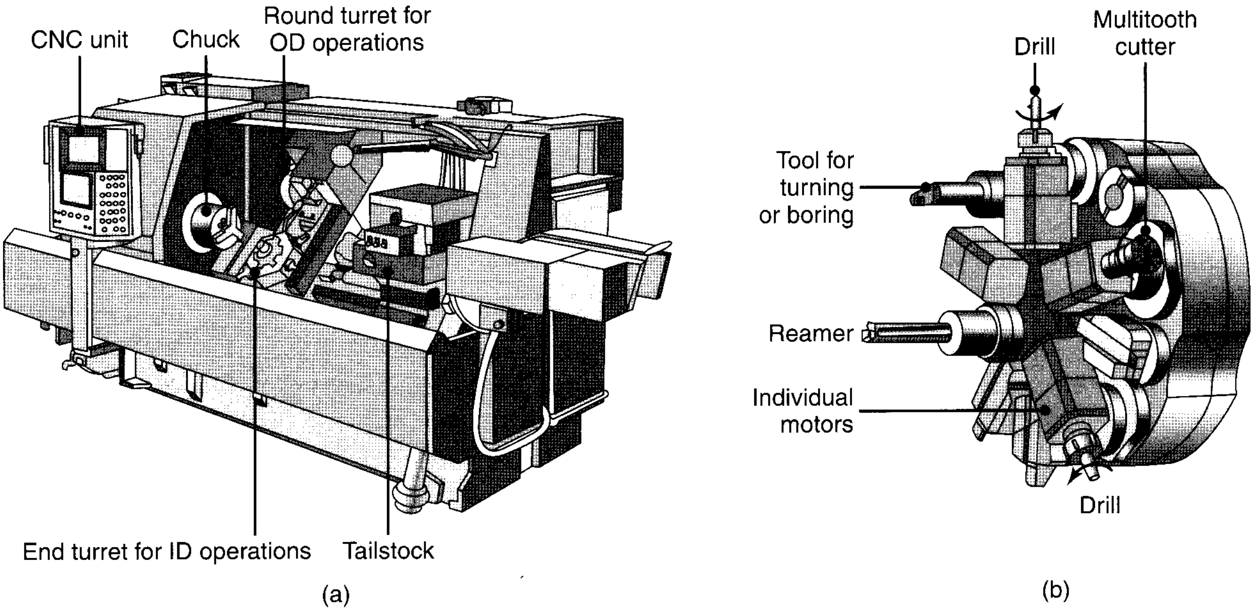


FIGURE 23.10 (a) A computer-numerical-control lathe. Note the two turrets. These machines have higher power and spindle speed than other lathes in order to take advantage of new cutting tools with enhanced properties. (b) A typical turret equipped with 10 tools, some of which are powered.

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 to machine large workpieces. Because of the heavy weight of the components, saddle-type lathe operations are slower than ram-type operations.

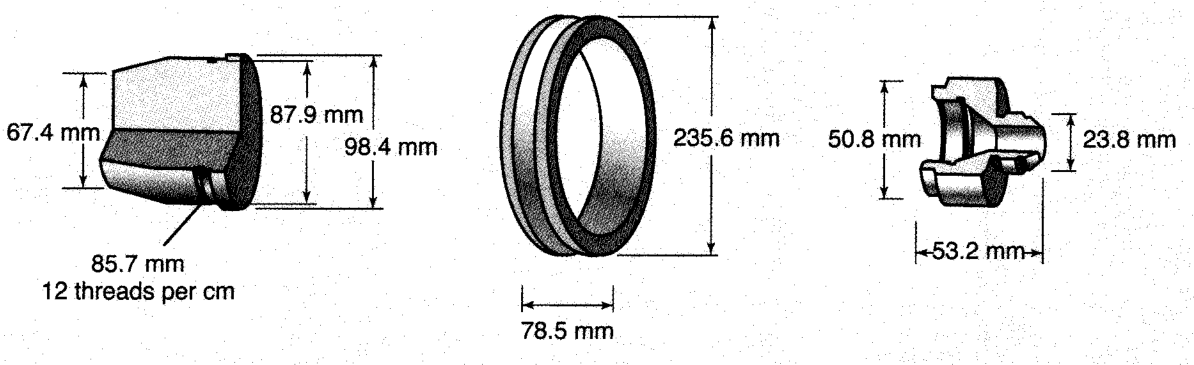
Computer-controlled Lathes. In the most advanced lathes, movement and control of the machine tool and its components are achieved by *computer numerical control* (CNC). The features of such a lathe are shown in Fig. 23.10a. These lathes generally are equipped with one or more turrets, and each turret is equipped with a variety of tools and performs several operations on different surfaces of the workpiece (Fig. 23.10b). Workpiece diameters may be as much as 1 m.

To take advantage of new cutting-tool materials, computer-controlled lathes are designed to operate faster and have higher power available compared with other lathes. They are equipped with *automatic tool changers* (ATCs). Their operations are reliably repetitive, maintain the desired dimensional accuracy, and require less skilled labor (once the machine is set up). They 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 Fig. 23.11. The material and number of cutting tools used and the machining times are indicated for each part. These parts also can be made on manual or turret lathes, although not as effectively or consistently.

Source: Courtesy of Monarch Machine Tool Company.



Material: Titanium alloy
Number of tools: 7
Total machining time
(two operations):
5.25 minutes

(a) Housing base

Material: 52100 alloy steel
Number of tools: 4
Total machining time
(two operations):
6.32 minutes

(b) Inner bearing race

Material: 1020 Carbon Steel
Number of tools: 8
Total machining time
(two operations):
5.41 minutes

(c) Tube reducer

FIGURE 23.11 Typical parts made on CNC lathes.

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. 23.12, 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, the machining of such parts consists of both roughing and finishing cuts):

1. *Pump shaft* (Fig. 23.12a). 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 shown in Fig. 23.10a. 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. 23.12b).

   The spindle turning angle is monitored directly, a processor performs a high-speed calculation, and the CNC issues a command to the cam turret in terms of that angle. The machine

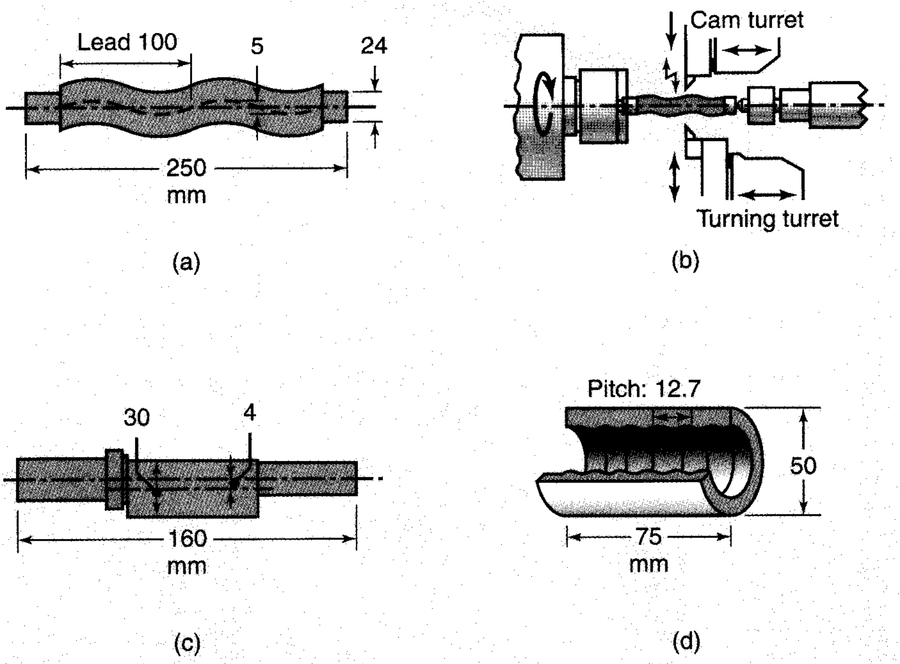


FIGURE 23.12 Examples of more complex shapes that can be produced on a CNC lathe.

has absolute-position feedback using a high-accuracy scale system. The CNC compares the actual value with the commanded one and 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 may be made of aluminum or stainless steel. The machining parameters for aluminum are given in Table 23.7 (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. 23.12c). This part is made of ductile (nodular) cast iron, and the machining parameters are shown in Part (b) of Table 23.7. The insert was K10 carbide. The machining time was 25 min; note that this time is of the same order of magnitude as that for the pump shaft described above.
3. *Tubular part with internal rope threads* (Fig. 23.12d). This part, made of 304 stainless steel, was machined under the conditions given for Part (c) in Table 23.7. 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, hence, improved dimensional accuracy and surface finish. For the threaded portion, the dimensional accuracy was ±0.05 mm, with a surface finish of $R_a = 2.5\ \mu\text{m}$.

   The machining time for this part was 1.5 min, which is 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.

TABLE 23.7

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 (C3)
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 the production rate among the processes listed in the table. These differences are due not only to the inherent characteristics of the processes and machine tools, but also to various other factors, such as the setup times and the types and sizes of the workpieces to be machined.

TABLE 23.8

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.

The ratings given in Table 23.8 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 at material-removal rates as high as 6000 cm^3/min. Also called *high-removal-rate machining*, this 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.

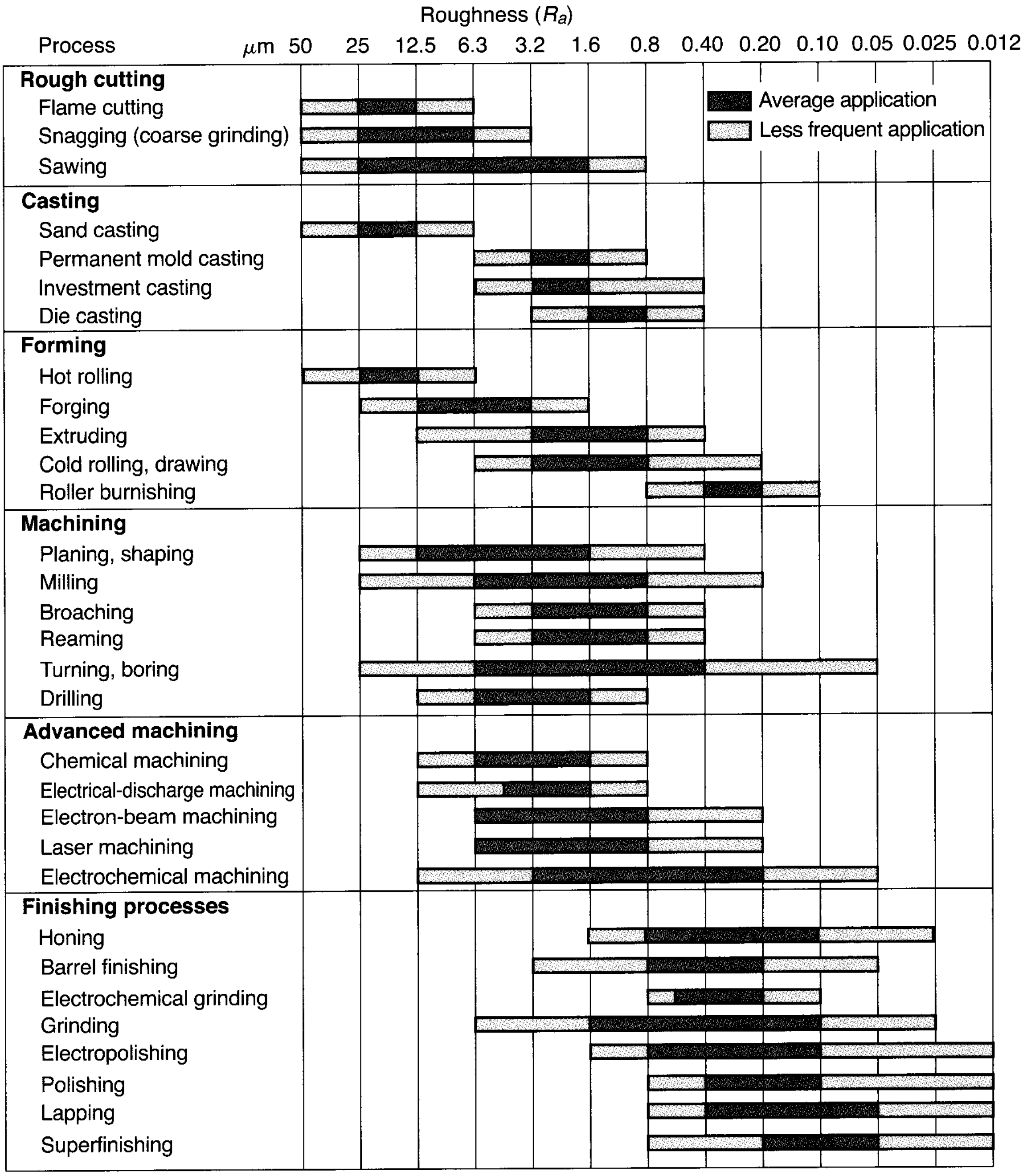


FIGURE 23.13 The range of surface roughnesses obtained in various processes. Note the wide range within each group, especially in turning and boring.

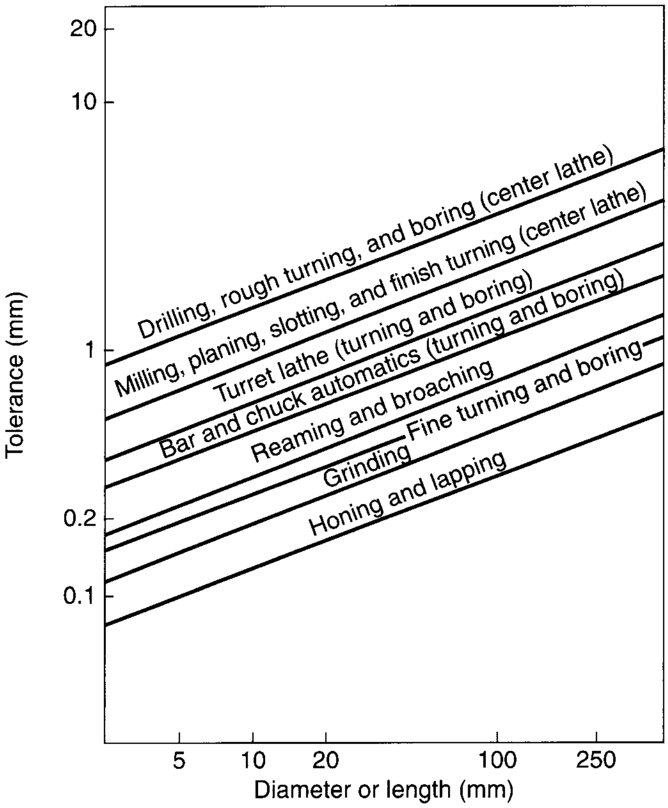


FIGURE 23.14 Range of dimensional tolerances obtained in various machining processes as a function of workpiece size. Note that there is an order-of-magnitude difference between small and large workpieces.

The surface finish (Fig. 23.13) and dimensional accuracy (Fig. 23.14) obtained in turning and related operations depend on several factors, such as the characteristics and condition of the machine tool, stiffness, vibration and chatter, process parameters, tool geometry and wear, the use of cutting fluids, the machinability of the workpiece material, and operator skill. As a result (and not surprisingly), a wide range of surface finishes can be obtained, as shown in Fig. 23.13. (See also Fig. 27.4.)

23.3.6 Design Considerations and Guidelines for Turning Operations

Several considerations are important in designing parts to be machined economically by turning operations. Because machining in general (a) takes considerable time, thus increasing the production cost, (b) wastes material, and (c) is not as economical as forming or shaping parts, it must be avoided as much as possible. When turning operations are necessary, follow these general design guidelines:

1. Parts should be designed so that they can be fixtured and clamped easily into work-holding devices. Thin, slender workpieces are difficult to support properly so that they can 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 for the part to still function properly.

TABLE 23.9

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

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 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 across the workpiece without obstruction.
6. Design features should be such that commercially available standard cutting tools, inserts, and toolholders can be used.
7. Workpiece materials should be selected for their machinability (Section 21.7) as much as possible.

Guidelines for Turning Operations. A general guide to the probable causes of problems in turning operations is given in Table 23.9. Recall that Chapters 21 and 22 described the factors influencing the parameters listed.

In addition to the various recommendations concerning tools and process parameters described thus far, an important factor is the presence of *vibration* and *chatter* (Section 25.4). Vibration during cutting can cause poor surface finish, poor dimensional accuracy, excessive tool wear, and premature tool failure. The following list outlines some generally accepted guidelines for turning operations (because of the complexity of the problem, however, some of the guidelines have to be implemented on a trial-and-error basis):

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, 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.)

23.3.7 Chip Collection Systems

The chips produced during machining 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 be in the range of 600 to 12,000 cm^3 depending on chip type. Likewise, the milling of 15 cm^3 of steel produces 450 to 750 cm^3 of chips, and cast iron produces 105 to 225 cm^3 of chips.

Also called **chip management**, the operation involves collecting chips from their source in the machine tool in an efficient manner and removing them from the work area. Long and stringy chips are more difficult to collect than short chips, which are produced by using tools with chipbreaker features (see Figs. 21.7 and 22.2). Thus, the type of chip produced is an integral aspect of the chip-collecting system.

Chips can be collected by any of the following methods:

- Allowing gravity to drop them onto a steel conveyor belt.
- Dragging the chips from a settling tank.
- Using augers with feed screws (similar to those in meat grinders).
- Using magnetic conveyors (for ferrous chips).
- Employing vacuum methods of chip removal.

Modern machine tools are designed with automated chip-handling features. Note that there may be a considerable amount of cutting fluid mixed with and adhering to the chips produced; hence, proper filtration or draining is important. The cutting fluid and sludge can be separated using *chip wringers* (centrifuges). Chip-processing systems usually require considerable floor space and can cost from $60,000 for small shops to over $1 million for large plants.

The collected chips may be recycled if it is economical to do so. Prior to their removal from a manufacturing plant, the large volume of chips can be reduced to as little as one-fifth of the loose volume by *compaction* (crushing) into briquettes or by *shredding*. Dry chips are more valuable for recycling because of reduced environmental contamination. The method chosen for chip disposal depends on economics, as well as on compliance with local, state, and federal regulations. The trend is to recycle all chips, the used cutting fluids, and the sludge.

23.3.8 Cutting Screw Threads

Screw threads are extremely common features. A *screw thread* may be defined as a ridge of uniform cross section that follows a helical or spiral path on the outside or inside of a cylindrical (*straight thread*) or tapered surface (*tapered thread*). Machine screws, bolts, and nuts have straight threads, as do threaded rods for applications such as the lead screw in lathes and various machinery components (Fig. 23.2). Threads may be *right handed* or *left handed*. Tapered threads commonly are used for water or gas pipes and plumbing supplies, which require a watertight or airtight connection.

Threads traditionally have been machined, but they are increasingly being formed by **thread rolling** (described in Section 13.5). Rolled threads now constitute the largest quantity of external threaded parts produced. It also may be possible to cast threaded parts, but there are limitations to dimensional accuracy, surface finish, and minimum dimensions. Production rates are not as high as those obtained in other processes.

Threads can be machined externally or internally with a cutting tool by a process called *thread cutting* or *threading*. Internal threads also can be produced with a special threaded tool, called a *tap*, and the process is called *tapping* (Section 23.7). External threads also may be cut with a *die* or by milling. Although the process adds considerably to the cost, threads subsequently may be ground for high dimensional accuracy and surface finish for applications such as screw drives in machines.

Screw-thread Cutting on a Lathe. A typical thread-cutting operation on a lathe is shown in Fig. 23.15a. 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 *lead screw* on the lathe. 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 rotation of the workpiece 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 the thread will be. In thread cutting, 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.

A number of passes in the sequence shown in Fig. 23.15b generally are required to produce threads with good dimensional accuracy and surface finish. Figure 23.15c shows a carbide insert for screw-thread cutting (*threading insert*) machining threads on a round shaft, and Fig. 23.15d shows an internal screw-thread cutting process. Cutting threads on a lathe is an old and versatile method, but it requires considerable operator skill and is a slow process. Consequently, except for small production runs, the process largely has been replaced by other methods, such as thread rolling, automatic screw machining, and the use of CNC lathes.

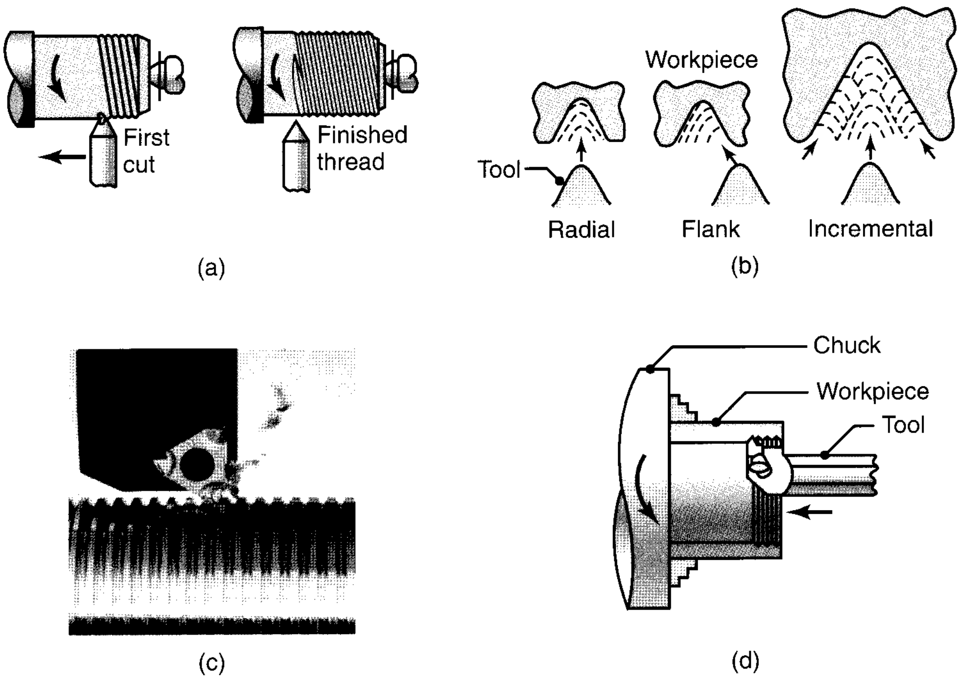


FIGURE 23.15 (a) Cutting screw threads on a lathe with a single-point cutting tool. (b) Cutting screw threads with a single-point tool in several passes, normally utilized for large threads. The small arrows in the figures show the direction of feed, and the broken lines show the position of the cutting 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.

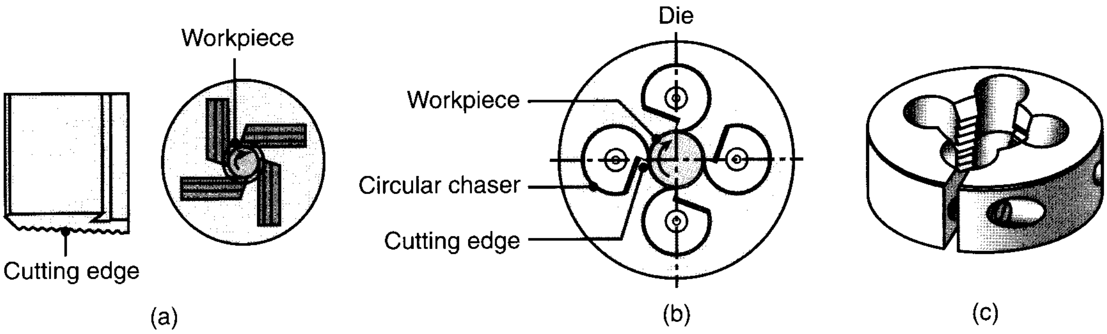


FIGURE 23.16 (a) Straight chasers for cutting threads on a lathe. (b) Circular chasers. (c) A solid threading die.

The production rate in cutting screw threads can be increased with tools called *die-head chasers* (Figs. 23.16a and b). These tools 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 around their axes to allow the part to be removed. *Solid-threading dies* (Fig. 23.16c) also are available for cutting straight or tapered screw threads. These dies are used mostly to thread the ends of pipes and tubing and are not suitable for production work.

Design Considerations for Screw Thread Machining. The design considerations that must be taken into account in order to produce high-quality and economical screw threads are as follows:

- Designs should allow for the termination of threads before they reach a shoulder. Internal threads in blind holes should have an unthreaded length at the bottom. The term *blind hole* refers to a hole that does not go through the thickness of the workpiece. (For example, see Fig. 23.1i.)
- Attempts should be made to eliminate shallow, blind tapped holes.
- 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 the diameter.
- Parts should be designed so that all cutting operations can be completed in one setup.

23.4 Boring and Boring Machines

Boring enlarges a hole made previously by some other process or 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. 23.17a) to reach the full length of the bore. The boring bar must be sufficiently stiff to minimize tool deflection and thus maintain dimensional accuracy and avoid vibration and chatter. For this reason, a material with a high elastic modulus (such as tungsten carbide) is

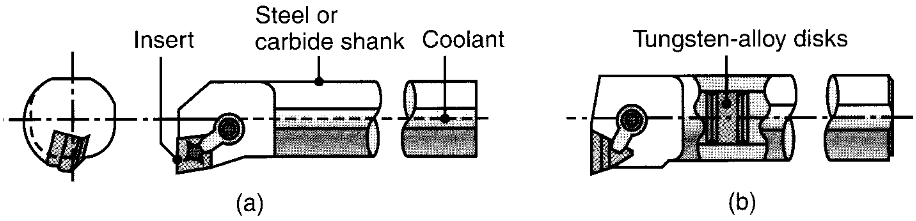


FIGURE 23.17 (a) Schematic illustration of a steel boring bar with a carbide insert. Note the passageway in the bar for cutting fluid application. (b) Schematic illustration of a boring bar with tungsten-alloy "inertia disks" sealed in the bar to counteract vibration and chatter during boring. This system is effective for boring-bar length-to-diameter ratios of up to 6.

desirable. Boring bars have been designed and built with capabilities for damping vibration (Fig. 23.17b).

Boring operations on relatively small workpieces can be carried out on lathes; large workpieces are machined on **boring mills.** These machine tools are either horizontal or vertical and are capable of performing various operations, such 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, which 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. 23.18) is similar to a lathe, has a vertical axis of workpiece rotation, and can accommodate workpieces with diameters as much as 2.5 m.

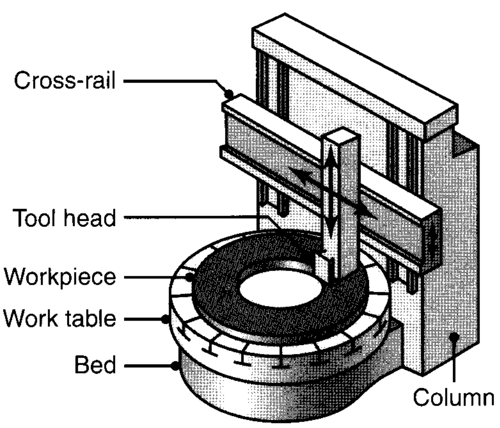


FIGURE 23.18 Schematic illustration of a vertical boring mill. Such a machine can accommodate workpiece sizes as large as 2.5 m in diameter.

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 radial movement (for facing), guided by the cross-rail. The head can be swiveled to produce conical (tapered) holes. Cutting speeds and feeds for boring are similar to those for turning. (For capabilities of boring operations, see Table 23.8.)

Boring machines are available with a variety of features. Machine capacities range up to 150 kW and are available with computer numerical controls, allowing all movements of the machine to be programmed. Little operator involvement is required, and consistency and productivity are improved.

Design Considerations for Boring. Guidelines for efficient and 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. Recall that the term *blind hole* refers to a hole that does not go through the thickness of the workpiece.
- The greater the length-to-bore-diameter ratio, the more difficult it is to hold dimensions, because of the deflections of the boring bar due to cutting forces as well as the higher 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

When inspecting various large or small products, note that the vast majority have several holes in them. Observe, for example, (a) the number of rivets on an airplane's wings and fuselage, (b) the bolts in engine blocks and heads, and (c) numerous consumer and industrial products. Holes typically are used for assembly with fasteners (such as bolts, screws, and rivets, each of which requires a hole), for design purposes (such as weight reduction, ventilation, or access to the inside of parts), or for appearance.

Hole making is among the most important operations in manufacturing, and **drilling** is a major and common hole-making process. Other processes for producing holes are punching (as described in Section 16.2) and various advanced machining processes (Chapter 27). The cost of hole making is among the highest machining costs in automotive engine production.

23.5.1 Drills

Drills typically have high length-to-diameter ratios (Fig. 23.19); hence, they are capable of producing relatively deep holes. However, drills are somewhat flexible

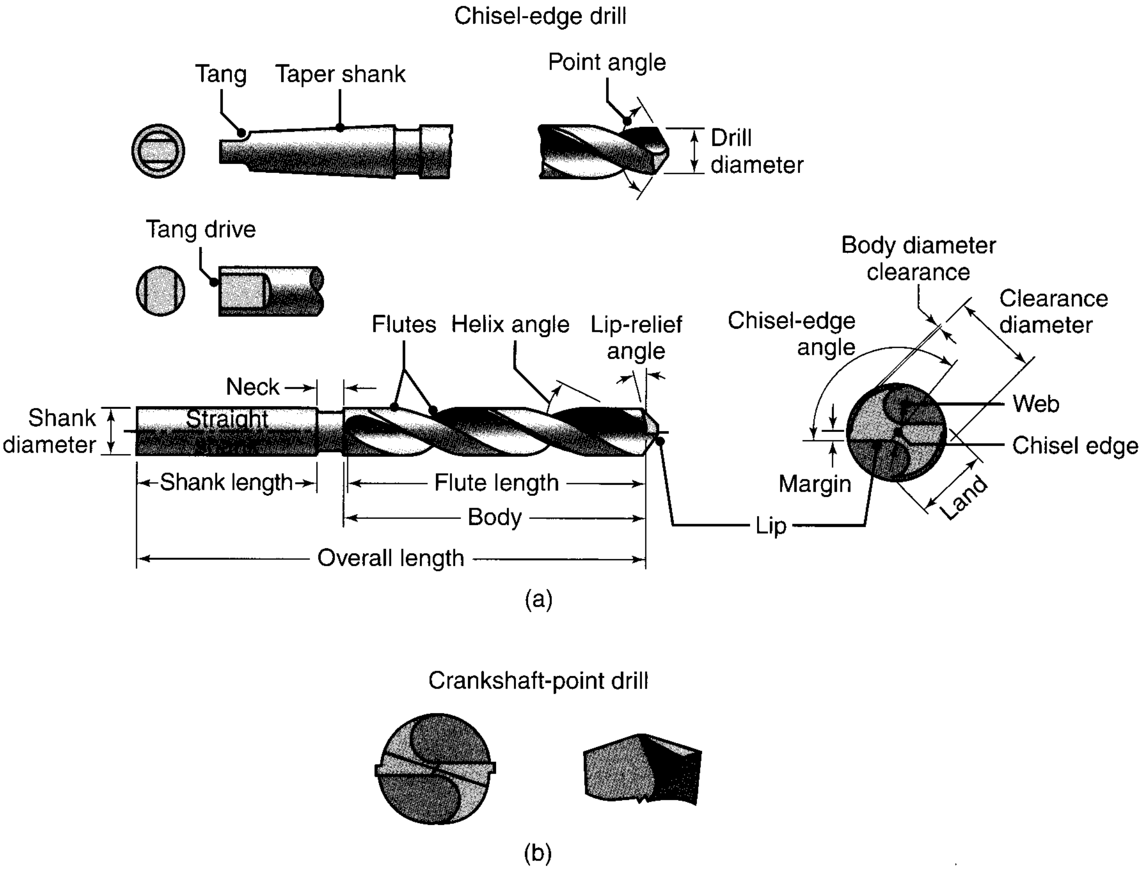


FIGURE 23.19 Two common types of drills: (a) Chisel-edge drill. The function of the pair of margins is to provide a bearing surface for the drill against walls of the hole as it penetrates into the workpiece. Drills with four margins (*double-margin*) are available for improved guidance and accuracy. Drills with chip-breaker features also are available. (b) Crankshaft drill. These drills have good centering ability, and because chips tend to break up easily, crankshaft drills are suitable for producing deep holes.

TABLE 23.10

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

and should be used with care in order to drill holes accurately and to prevent breakage. Furthermore, the chips that are produced within the hole move in a direction opposite to the forward movement of the drill. Thus, chip disposal and ensuring cutting-fluid effectiveness can present significant difficulties in drilling.

Drills generally leave a *burr* on the bottom surface upon breakthrough, necessitating deburring operations (Section 26.8). Also, because of its rotary motion, drilling produces holes with walls that have circumferential marks. In contrast, punched holes have longitudinal marks (see Fig. 16.5a). This difference is significant in terms of the hole's fatigue properties, as we describe in 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. The amount of oversize depends on the quality of the drill and of the equipment used, as well as on the machining practices employed. Furthermore, depending on their thermal properties, some metals and nonmetallic materials expand significantly due to the heat produced by drilling; thus, the final hole diameter could be smaller than the drill diameter. For better surface finish and dimensional accuracy, drilled holes may be subjected to subsequent operations, such as *reaming* and *honing*. The capabilities of drilling and boring operations are shown in Table 23.10.

Twist Drill. The most common drill is the conventional *standard-point twist drill* (Fig. 23.19a). 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 as follows (with typical ranges of angles 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 (*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, for example, the drill shown in Fig. 23.22a) through which cutting fluids are forced, thus improving lubrication and cooling as well as washing away the chips. Drills are available with a **chip-breaker** feature ground along the cutting edges. This feature is important in drilling with automated machinery, where a continuous removal of long chips without operator assistance is essential.

The various angles on a drill have been developed through experience and are 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. For example, too small a lip relief angle (Fig. 23.19a) increases the thrust force, generates excessive heat, and increases wear. By contrast, too large an angle can cause chipping or breaking of the cutting edge. In addition to conventional point drills, several other drill-point geometries have been developed to improve drill performance and increase the penetration rate. Special grinding techniques and equipment are used to produce these geometries.

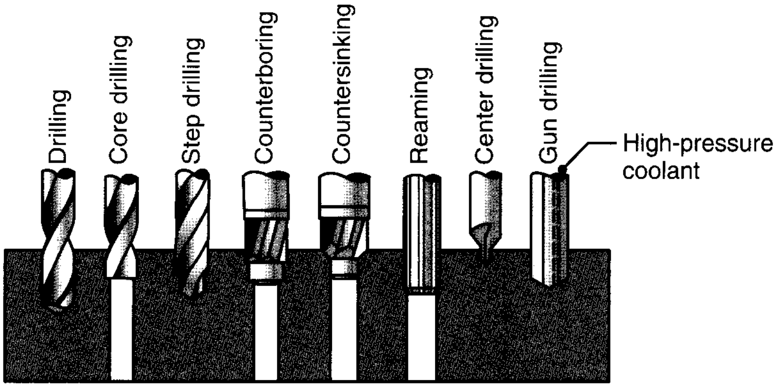


FIGURE 23.20 Various types of drills and drilling and reaming operations.

Other Types of Drills. Several types of drills are shown in Fig. 23.20. A *step drill* produces holes 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 surface to accommodate the heads of screws and bolts below the workpiece surface. A *center drill* is short and is used to produce a hole at the end of a piece of stock, so that it may be mounted between centers of the headstock and the tailstock of a lathe (Fig. 23.2). A *spot drill* is used to spot (to start) a hole at the desired location on a surface.

Spade drills (Fig. 23.21a) have removable tips or bits and are used to produce large-diameter and deep holes. These drills have the advantages of higher stiffness (because of the absence of flutes in the body of the drill), ease of grinding the cutting edges, and lower cost. A similar drill is the *straight-flute drill* (Fig. 23.21b).

Solid carbide and carbide-tipped drills (Figs. 23.21c and d) are available for drilling hard materials (such as cast irons), high-temperature metals, abrasive materials (such as concrete and brick—*masonry drills*), and composite materials with abrasive fiber reinforcements (such as glass and graphite).

Gun Drilling. Developed originally for drilling gun barrels, *gun drilling* is used for drilling deep holes and requires a special drill (Fig. 23.22). The depth-to-diameter

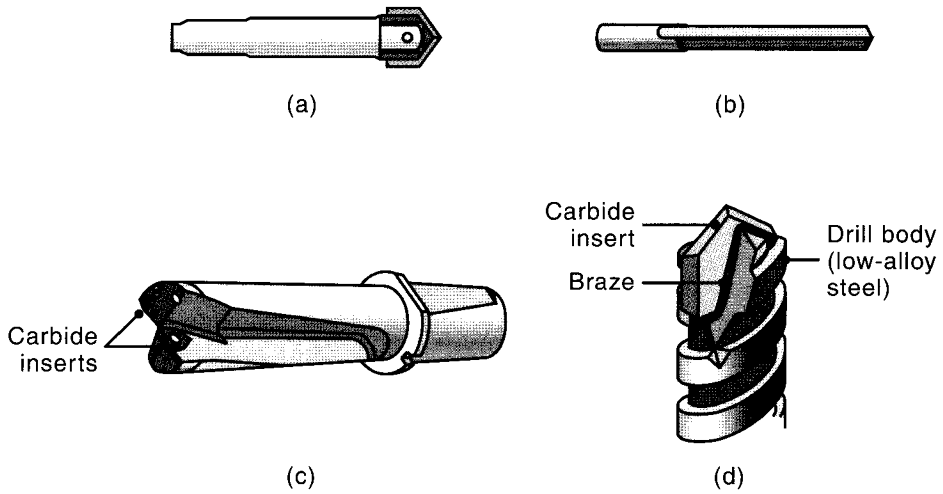


FIGURE 23.21 Various types of drills. (a) Spade drill; (b) Straight-flute drill; (c) Drill with indexable carbide inserts; (d) drill with brazed-carbide tip.

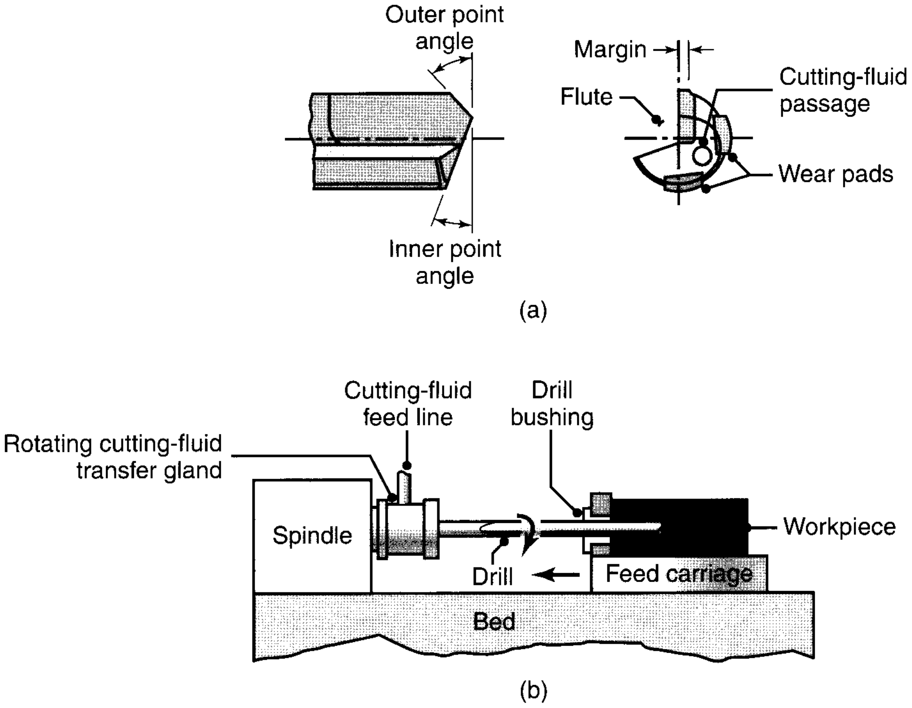
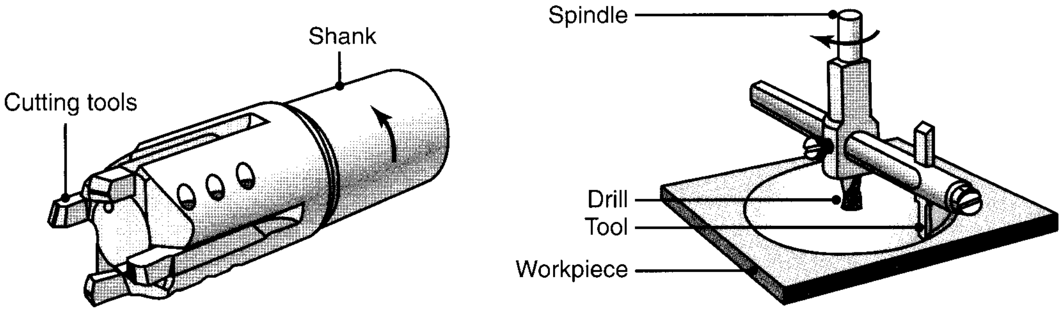


FIGURE 23.22 (a) A gun drill, showing various features. (b) Schematic illustration of the gun-drilling operation.

ratios of holes produced can be 300:1 or even higher. The thrust force (the radial force that tends to push 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** (see next subsection), 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. Tolerances typically are about 0.025 mm. The cutting fluid is forced under high pressure through a longitudinal hole (passage) 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 and thus interfere with the drilling operation. The tool does not have to be retracted to clear the chips, as is usually done with twist drills.

Trepanning. In *trepanning* (from the Greek *trypanon*, or "boring a hole" or "auger") the cutting tool (Fig. 23.23a) 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 that is removed 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 in which O-rings are to be placed (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. 23.23b.



(a) (b)

FIGURE 23.23 (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. For a drill with a 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,

$$\text{MRR} = \left(\frac{\pi D^2}{4}\right) f N. \tag{23.3}$$

The dimensional accuracy of this equation can be checked, as was done for Eq. (23.1), by noting that MRR = (mm^2)(mm/rev)(rev/min) = mm^3/min, which is the correct unit for volume removed per unit time.

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, particularly if it does not have sufficient stiffness (for example, thin sheet-metal structures), or it can cause the workpiece to slip into the workholding fixture.

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. Accurate calculation of the thrust force on the drill is difficult. Thrust forces typically range from a few newtons for small drills to as high as 100 kN for drilling high-strength materials with large drills. Experimental data are available as an aid in the design and use of drills and drilling equipment.

Torque. A knowledge of the *torque* in drilling is essential for estimating the power requirement; however, because of the many factors involved, it is difficult to calculate. Torque can be estimated from the data given in Table 21.2 by noting that the power dissipated during drilling is the product of torque and rotational speed and that we first have to calculate the material-removal rate. Torque in drilling can be as high as 4000 N • m.

EXAMPLE 23.4 Material-removal Rate and Torque in Drilling

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. Calculate the material-removal rate and the torque on the drill.

Solution The material-removal rate is calculated from Eq. (23.3):

$$\begin{aligned}\text{MRR} &= \left[\frac{(\pi)(10)^2}{4}\right](0.2)(800)\\ &= 12{,}570\ \text{mm}^3/\text{min} = 210\ \text{mm}^3/\text{s}.\end{aligned}$$

Referring to Table 21.2, let's take an average unit power of 0.5 $\text{W}\cdot\text{s/mm}^3$ 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 = $\text{N}\cdot\text{m}$, we find that

$$T = \frac{105}{83.8} = 1.25\ \text{N}\cdot\text{m}.$$

23.5.4 Drill Materials and Sizes

Drills usually are made of high-speed steels (M1, M7, and M10) and solid carbides or with carbide tips (typically made of K20 (C2) carbide), like those shown in Fig. 23.21c and d. Drills are now commonly coated with titanium nitride or titanium carbonitride for increased wear resistance (Section 22.5). Polycrystalline-diamond-coated drills are used for producing fastener holes in fiber-reinforced plastics. Because of their high wear resistance, several thousand holes can be drilled with little damage to the material.

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. (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. (in $\frac{1}{64}$-in. increments) to 3.5 in. (in $\frac{1}{16}$-in. increments)
- **Millimeter:** From 0.05 mm (0.002 in.) in increments of 0.01 mm.
  Note: 1 in. = 25.4 mm.

23.5.5 Drilling Practice

Drills and similar hole-making tools usually are held in *drill chucks*, which may be tightened with or without keys. 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 the operation. This problem is particularly severe with small-diameter long drills and can lead to failure. To start a hole properly, the drill should be guided, using fixtures (such as a bushing) to keep it from deflecting laterally. A small starting hole can be made with a *center drill* (usually with a point angle of 60°), or the drill point may be ground to an *S* shape (*helical* or *spiral point*). This shape has a self-centering characteristic—thus eliminating the need for center drilling—and produces accurate holes 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. Other alternatives for

TABLE 23.11

General Recommendations for Speeds and Feeds in Drilling

		Drill diameter			
		Feed, mm/rev		Speed, rpm	
Workpiece material	Surface speed m/min	1.5 mm	12.5 mm	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.

minimizing walking of the drill bit are to use a centering punch to produce an initial impression in which drilling starts or else to incorporate dimples or other features into the cast or forged blank.

Drilling Recommendations. Recommended ranges for drilling speeds and feeds are given in Table 23.11. 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

$$V = \left(\frac{12.7}{2}\text{mm}\right)(300\text{ rev/min})(2\pi\,\text{rad/rev})\left(\frac{1}{1000}\text{m/mm}\right) = 12\text{ m/min}$$

In drilling holes smaller than 1 mm in diameter, 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 drill 1.5 mm in diameter 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 during drilling can be difficult, especially for deep holes in soft and ductile workpiece materials. The drill should be retracted periodically (*pecking*) to remove chips that may have accumulated along the flutes. Otherwise, it may break because of excessive torque, or it may "walk" off location and produce a misshaped hole. A general guide to the probable causes of problems in drilling operations is given in Table 23.12.

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 computer numerical control 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. Coated drills can be recoated.

Measuring Drill Life. Drill life, as well as tap life (see Section 23.7) usually is measured by the number of holes drilled before they become dull and need to be re-

TABLE 23.12

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

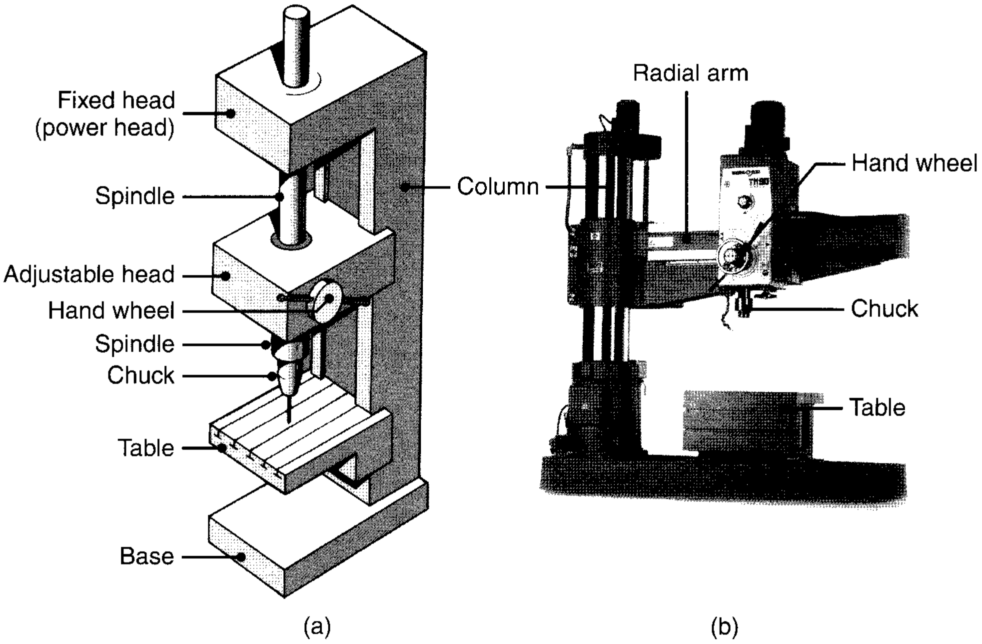


FIGURE 23.24 (a) Schematic illustration of the components of a vertical drill press. (b) A radial drilling machine. *Source*: (b) Courtesy of Willis Machinery and Tools.

worked or replaced. This can be determined experimentally by clamping a block of material on a suitable dynamometer or force transducer and drilling a number of holes while recording the torque or thrust force during each successive operation. After a number of holes have been drilled, the torque and force begin to increase because the tool is becoming dull. *Drill life* is defined as the number of holes drilled until this transition begins. Other techniques, such as monitoring vibration and acoustic emissions (Section 21.5.4), also may be used to determine drill life.

23.5.6 Drilling Machines

Drilling machines are used for drilling holes, tapping, reaming, and small-diameter boring operations. The most common machine is the **drill press**, the major components of which are shown in Fig. 23.24a. 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 in turn is clamped to the table. The drill is lowered manually by a handwheel or by power feed at preset rates. Manual feeding requires some skill in judging the appropriate feed rate.

Drill presses usually are 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 on drilling machines has to be adjustable to accommodate 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-diameter holes to large *radial drills* (Fig. 23.24b), which 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 be swiveled to drill holes at an angle. Developments in drilling machines include numerically controlled three-axis machines, in which the operations are performed automatically and in the desired sequence with the use of a turret (Fig. 23.25). Note that the turret holds several different drilling tools.

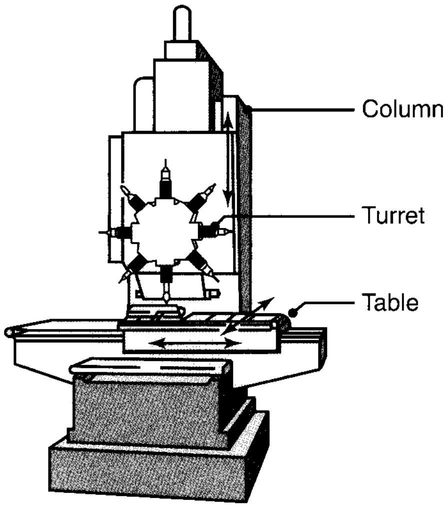


FIGURE 23.25 A three-axis computer numerical-control drilling machine. The turret holds as many as eight different tools, such as drills, taps, and reamers.

Drilling machines with multiple spindles (**gang drilling**) are used for high-production-rate operations. These machines are capable of drilling, in one cycle, as many as 50 holes of varying sizes, depths, and locations. They also are used for reaming and counterboring operations. However, with advances in machine tools, gang-drilling machines are now being replaced with *numerical-control turret drilling machines*. Special drilling machines, such as those which produce holes in continuous hinges (*piano hinges*), use twist drills 1 mm in diameter. These machines usually are horizontal and produce holes in up to 3-m long segments in one cycle.

Workholding devices for drilling are essential to ensure that the workpiece is located properly. They also keep the workpiece from slipping or rotating during drilling. Workholding devices are available in various designs; the important features are (a) three-point locating for accuracy and (b) three-dimensional work holding for secure fixtures. (See also Section 37.8.)

23.5.7 Design Considerations for Drilling

The basic design guidelines for drilling are as follows:

- Designs should allow holes to be drilled on flat surfaces and perpendicular to the drill motion. Otherwise, the drill tends to deflect and the hole will not be located accurately. Exit surfaces for the drill also should be flat.
- Interrupted hole surfaces should be avoided or minimized for improved dimensional accuracy, for extended drill life, and to avoid vibrations.
- Hole bottoms should match, if possible, standard drill-point angles; flat bottoms or odd shapes should be avoided.
- Through holes are preferred over blind holes. If holes with large diameters are required, the workpiece should have a preexisting hole, preferably made during fabrication of the part (such as by casting, powder metallurgy, or forming).
- Dimples should be provided when preexisting holes are not practical, in order to reduce the tendency of the drill to walk.
- Parts should be designed so that all drilling can be performed with a minimum of fixturing and without having to reposition the workpiece.
- Blind holes must be drilled deeper than subsequent reaming or tapping operations that may be performed.

23.6 Reaming and Reamers

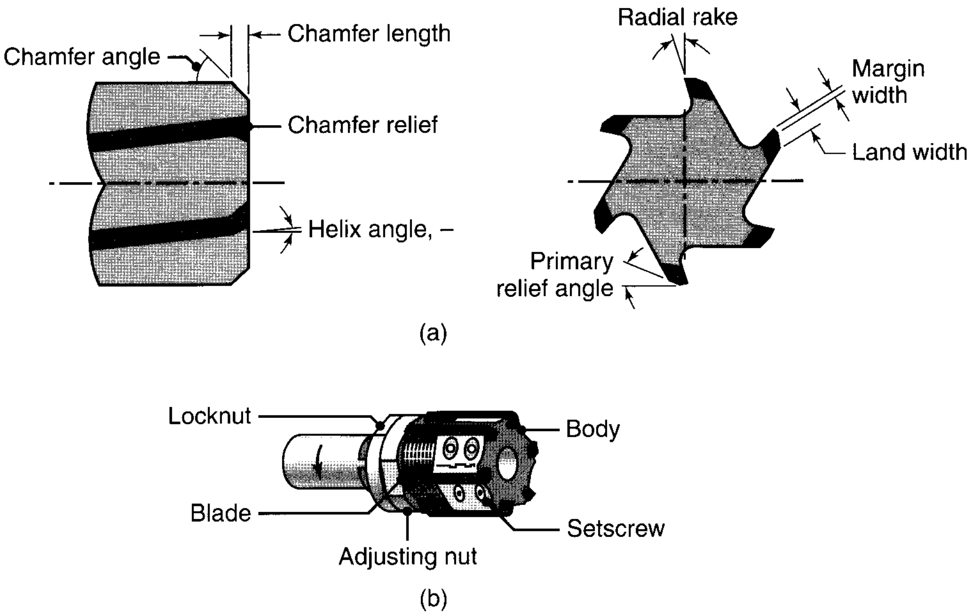
Reaming is an operation used to (a) make an existing hole dimensionally more accurate than can be achieved by drilling alone, and (b) improve its surface finish. The most accurate holes in workpieces generally are produced by the following sequence of operations:

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. 23.26a) 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 operated by a machine) are available in two types: (1) *Rose reamers* have cutting edges with wide margins and no relief (Fig. 23.26a). They remove considerable material and true up a hole for flute reaming. (2) *Fluted reamers* have small margins and relief, with a rake angle of about 5°. They usually are used for light cuts of about 0.1 mm on the hole diameter.



(a)

(b)

FIGURE 23.26 (a) Terminology for a helical reamer. (b) Inserted-blade adjustable reamer.

Shell reamers (which are hollow and are mounted on an arbor) generally are used for holes larger than 20 mm. *Expansion reamers* are adjustable for small variations in hole size and also to compensate for wear of the reamer's cutting edges. *Adjustable reamers* (Fig. 23.26b) can be set for specific hole diameters and therefore are versatile.

Reamers may be held rigidly (as in a chuck), or they may *float* in their holding fixtures to ensure alignment or 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 produces a hole by drilling, and the rest of the same tool performs a reaming operation. A similar development involves drilling and tapping in one stroke, using a single tool.

Reamers typically are 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 in workpieces can be produced by *tapping*. A *tap* is a chip-producing threading tool with multiple cutting teeth (Fig. 23.27a). Taps generally are available with two, three, or four flutes. The most common production tap is the two-flute spiral-point tap. The two-flute tap 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.

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 rotation.

Chip removal can be a significant problem during tapping because of the small clearances involved. If chips aren't removed properly, the excessive torque that results can break the tap. The use of a cutting fluid and periodic reversal and removal of the tap from the hole are effective means of chip removal and of improving the quality of the tapped hole. For higher tapping productivity, drilling and tapping can be combined in a single operation (*drapping*) with a single tool. The tool has a drilling section at its tip, followed by a tapping section.

Tapping may be done by *hand* or with machines such as (a) drilling machines, (b) lathes, (c) automatic screw machines, and (d) vertical CNC milling machines

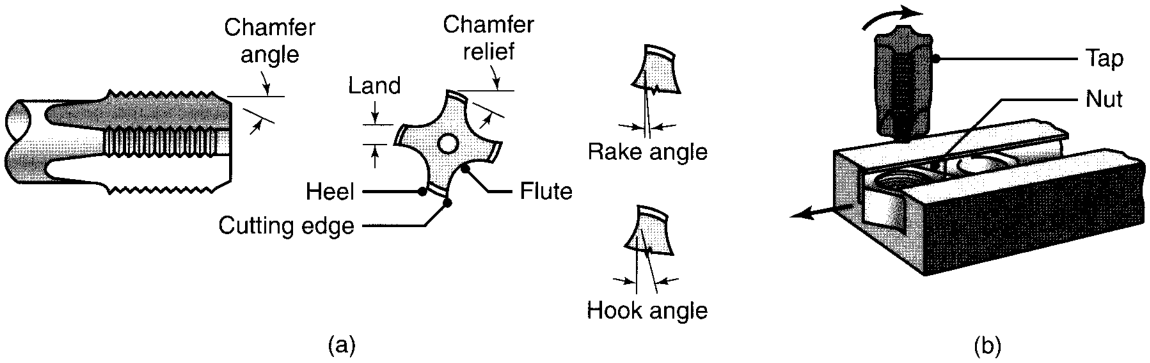


FIGURE 23.27 (a) Terminology for a tap. (b) Tapping of steel nuts in production.

combining 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 the tapping of holes. One system for the automatic tapping of nuts is shown in Fig. 23.27b.

With proper lubrication, tap life may be as high as 10,000 holes. Tap life can be determined with the same technique used to measure drill life. 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 high as 100 m/min. *Self-reversing* tapping systems also have been improved significantly and 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 now 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 23.1 Bone Screw Retainer

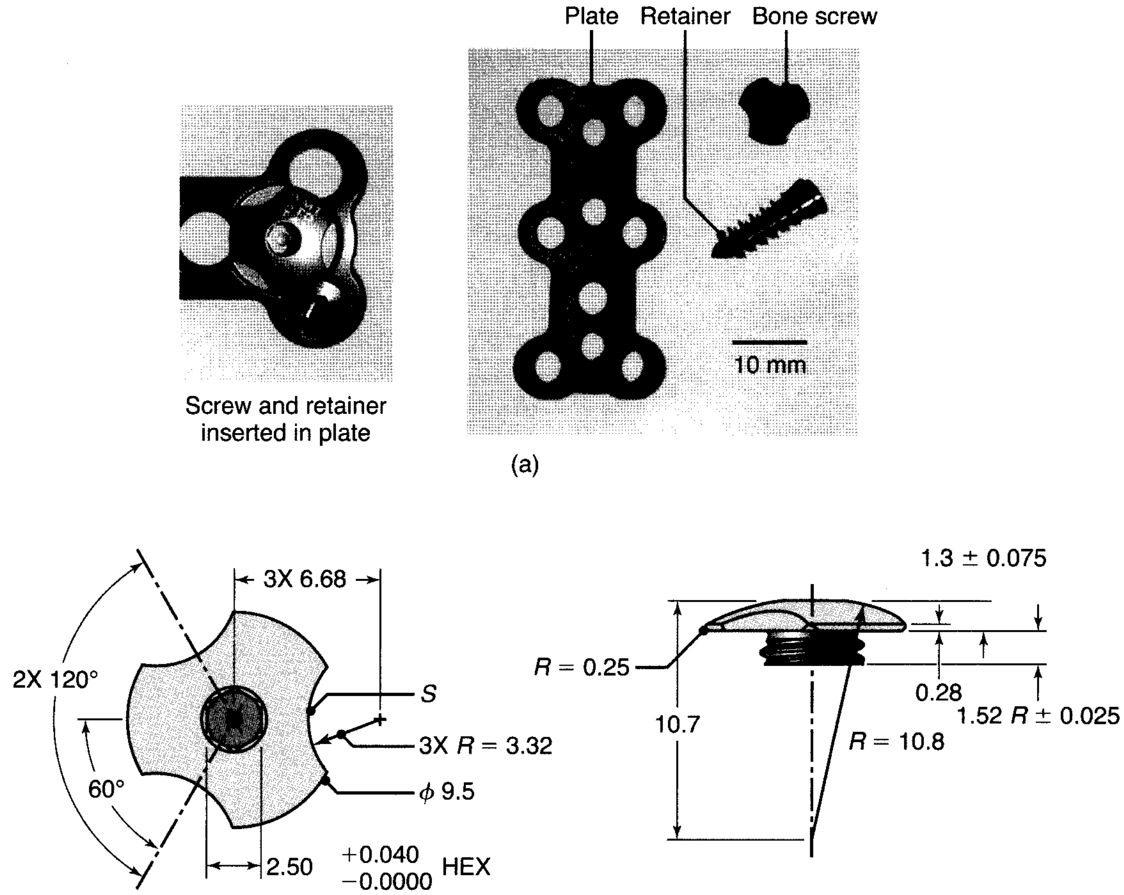
A cervical spine implant is shown in Fig. 23.28a. In the event that a patient requires cervical bone fusion at one or more vertebral levels, this implant can act as an internal stabilizer by decreasing the amount of motion in the region and thereby help promote a successful fusion. The plate affixes to the anterior aspect of the spine, with 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 time due to normal, repetitive loading from the patient. In extreme cases, this can result in a screw backing out, with the head of the screw no longer flush with the plate—a condition that obviously is undesirable. This implant uses a retainer to prevent the bone screw from backing out away from the plate. The part drawing for the retainer is shown in the left half of Fig. 23.28b.

The retainer has a number of 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. 23.28b.

The manufacturing steps followed to produce this part are shown in Fig. 23.28b. First, a 12-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 the 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 length. The spherical radius in the cap is machined, the center hole is drilled, and the hex head is broached. The cap is removed and inspected. 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 threaded, tapered 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 and tumbled to remove all sharp



Note: Thread must start at point *S* to ensure that retainer interferes with bone screw.

(b)

FIGURE 23.28 A cervical spine implant. All dimensions in mm.

corners, and the bottom is heavy-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 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 process requires an understanding of the interrelationships among design parameters (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
Dry machining
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

Arnone, M., **High Performance Machining**, Hanser, 1998.

ASM Handbook, Vol. 16: **Machining**, ASM International, 1989.

Boothroyd, G., and Knight, W.A., **Fundamentals of Machining and Machine Tools**, 3rd ed., Marcel Dekker, 2005.

Brown, J., **Advanced Machining Technology Handbook**, McGraw-Hill, 1998.

Byers, J.P. (ed.), **Metalworking Fluids**, Marcel Dekker, 1994.

Handbook of High-Speed Machining Technology, Chapman and Hall, 1985.

Hoffman, E.G., **Jigs and Fixture Design**, 5th ed., Industrial Press, 2003.

Joshi, P.H., **Machine Tools Handbook**, McGraw-Hill, 2008.

Krar, S.F., and Check, A.F., **Technology of Machine Tools**, 6th ed., Glencoe Macmillan/McGraw-Hill, 2009.

Lopez, L.N., and Lamikiz, A. (eds.), **Machine Tools for High Performance Machining**, Springer, 2009.

Machinery's Handbook, Industrial Press, revised periodically.

Machining Data Handbook, 3rd ed., 2 vols. Machinability Data Center, 1980.

Nachtman, E.S., and Kalpakjian, S., **Lubricants and Lubrication in Metalworking Operations**, Marcel Dekker, 1985.

Sluhan, C. (ed.), **Cutting and Grinding Fluids: Selection and Application**, Society of Manufacturing Engineers, 1992.

Stephenson, D.A., and Agapiou, J.S., **Metal Cutting Theory and Practice**, 2nd ed., CRC Press, 2006.

Stout, K.J., Davis, E.J., and Sullivan, P.J., **Atlas of Machined Surfaces**, Chapman and Hall, 1990.

Tool and Manufacturing Engineering Handbook, 4th ed., Vol. 1: **Machining**. Society of Manufacturing Engineers, 1983.

Walsh, R.A., **McGraw-Hill Machining and Metalworking Handbook**, 3rd ed., McGraw-Hill, 2006.

Weck, M., **Handbook of Machine Tools**, 4 vols. Wiley, 1984.

REVIEW QUESTIONS

23.1. Describe the types of machining operations that can be performed on a lathe.

23.2. Describe the operations that can be performed on a drill press.

23.3. Why were power chucks developed?

23.4. Explain why operations such as boring on a lathe and tapping are difficult.

23.5. Why are turret lathes typically equipped with more than one turret?

23.6. Describe the differences between boring a workpiece on a lathe and boring it on a horizontal boring mill.

23.7. How is drill life determined?

23.8. Why are reaming operations performed?

23.9. Explain the functions of the saddle on a lathe.

23.10. Describe the relative advantages of (a) self-opening and (b) solid-die heads for threading.

23.11. Explain how external threads are cut on a lathe.

23.12. What is the difference between a blind hole and a through hole? What is the significance of that difference?

QUALITATIVE PROBLEMS

23.13. Explain the reasoning behind the various design guidelines for turning.

23.14. 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.

23.15. List and explain the factors that contribute to poor surface finish in the processes described in this chapter.

23.16. Explain why the sequence of drilling, boring, and reaming produces a hole that is more accurate than drilling and reaming it only.

23.17. 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.

23.18. 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.

23.19. Describe the difficulties that may be encountered in clamping a workpiece made of a soft metal in a three-jaw chuck.

23.20. Does the force or torque in drilling change as the hole depth increases? Explain.

23.21. Explain the similarities and differences in the design guidelines for turning and for boring.

23.22. Describe the advantages and applications of having a hollow spindle in the headstock of a lathe?

23.23. 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.

23.24. 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?

23.25. Describe your observations concerning the contents of Tables 23.2 and 23.4, and explain why those particular recommendations are made.

23.26. The footnote to Table 23.11 states that as the hole diameter increases, speeds and feeds in drilling should be reduced. Explain why.

23.27. In modern manufacturing, which types of metal chips would be undesirable and why?

23.28. The operational severity for reaming is much lower than that for tapping, even though they both are internal machining processes. Why?

23.29. Review Fig. 23.6, and comment on the factors involved in determining the height of the zones (cutting speed) for various tool materials.

23.30. Explain how gun drills remain centered during drilling. Why is there a hollow, longitudinal channel in a gun drill?

23.31. Comment on the magnitude of the wedge angle on the tool shown in Fig. 23.4.

23.32. If inserts are used in a drill bit (see Fig. 23.21), how important is the shank material? If so, what properties are important? Explain.

23.33. Refer to Fig. 23.10b, 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

▶**23.34.** Calculate the same quantities as in Example 23.1 for high-strength titanium alloy and at $N = 700$ rpm.

▶**23.35.** Estimate the machining time required to rough turn a 0.50-m-long annealed copper-alloy round bar from a 60-mm diameter to a 58-mm diameter, using a high-speed steel tool. (See Table 23.4.) Estimate the time required for an uncoated carbide tool.

▶**23.36.** A high-strength cast-iron bar 200 mm in diameter is being turned on a lathe at a depth of cut $d = 1.25$ 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?

▶**23.37.** 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?

▶**23.38.** In Example 23.4, assume that the workpiece material is high-strength aluminum alloy and the spindle is running at $N = 500$ rpm. Estimate the torque required for this operation.

▶**23.39.** 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.

▶**23.40.** For the data in Problem 23.39, calculate the power required.

23.41. 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

23.42. 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.

23.43. Would it be difficult to use the machining processes described in this chapter on various soft nonmetallic 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.

23.44. 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.

23.45. 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.

23.46. Review Fig. 23.7d, and explain if it would be possible to machine eccentric shafts, such as that shown in Fig. 23.12c, on the setup illustrated. What if the part is long compared with its cross section? Explain.

23.47. Boring bars can be designed with internal damping capabilities to reduce or eliminate vibration and chatter during machining (see Fig. 23.17). Referring to the technical literature, describe details of designs for such boring bars.

23.48. 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.

Machining Processes: Milling, Broaching, Sawing, Filing, and Gear Manufacturing

CHAPTER 24

- This chapter begins with milling, one of the most versatile machining processes, in which a rotating cutter removes material while traveling along various axes with respect to the workpiece.
- Several other machining processes are then described, such as planing, shaping, and broaching, in which either the cutting tool or the workpiece travels along a straight path, producing flat or various shaped surfaces.
- Sawing processes are then covered, generally used for preparing blanks for subsequent operations such as forming, machining, and welding.
- Also briefly discussed is filing, used to remove small amounts of material, usually from edges and corners.
- The chapter ends with descriptions of gear-manufacturing methods by machining, the special cutters used, the equipment involved, and the quality and properties of the gears made.

Typical parts made: Parts with complex external and internal features, splines, and gears.

Alternative processes: Die casting, precision casting, precision forging, powder metallurgy, powder-injection molding, creep feed grinding, electrical discharge machining, rapid prototyping, and fabrication.

EXAMPLES:

CASE STUDY:

24.1 Introduction

In addition to producing parts with various external or internal round profiles, machining operations can produce many other parts with more complex shapes (Fig. 24.1). In this chapter, several cutting processes and machine tools that are capable of producing these shapes using single-point, multitooth, and cutting tools are described (see also Table 23.1).

It is not unusual that complex shapes need to be produced to very close tolerances and a fine surface finish. Although processes such as die casting and precision forging can achieve such goals to some degree, it is often necessary to perform complicated machining to achieve design requirements.

The preceding chapter described machining processes that produce round shapes. While the processes and machinery covered in this chapter also can produce round, axisymmetric shapes, it is generally advisable to use the processes of Chapter 23

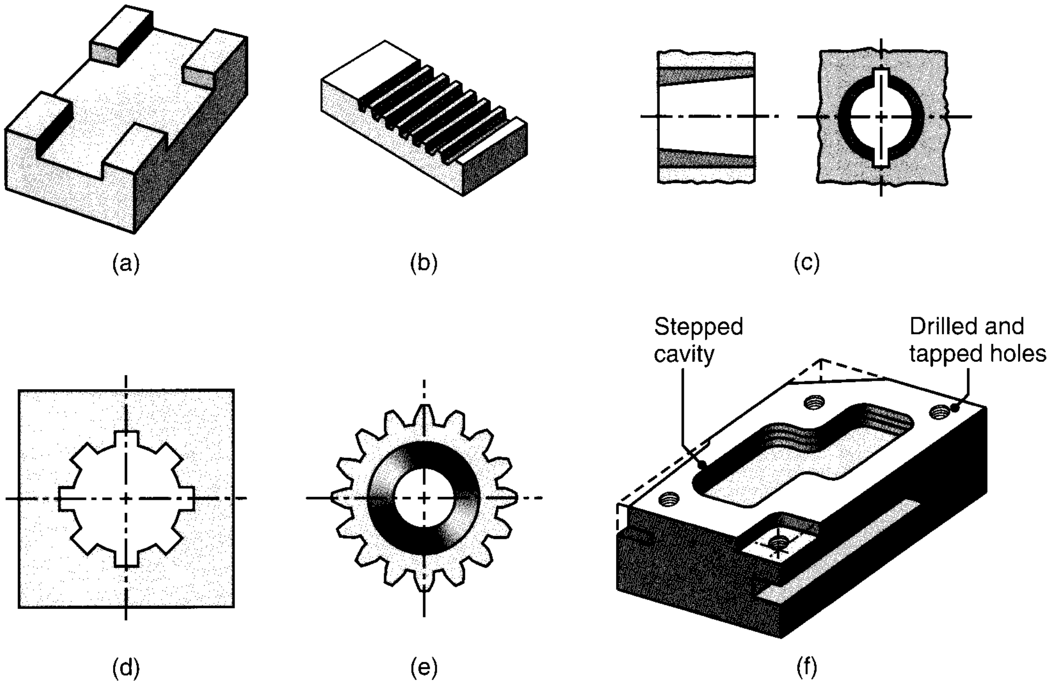


FIGURE 24.1 Typical parts and shapes that can be produced with the machining processes described in this chapter.

whenever possible, since the equipment is simpler, less expensive, and the processes are easier to set up and perform. This, however, cannot always be done, so the machining operations described in this chapter become essential.

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.

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. 24.2a. The cutter body, which generally is made of high-speed steel, 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 process is called *slab milling*. Cutters for peripheral milling may have *straight* or *helical teeth* (as shown in Fig. 24.2a), resulting in orthogonal or oblique cutting action, respectively. Helical teeth generally are preferred over straight teeth because the tooth is partially engaged with the workpiece as it rotates. Consequently, the cutting force and the torque on the cutter are lower, resulting in a smoother operation and reduced chatter. (See also Fig. 21.9.)

Conventional Milling and Climb Milling. Note in Fig. 24.3a that the cutter rotation can be either clockwise or counter-clockwise; this is significant in the operation.

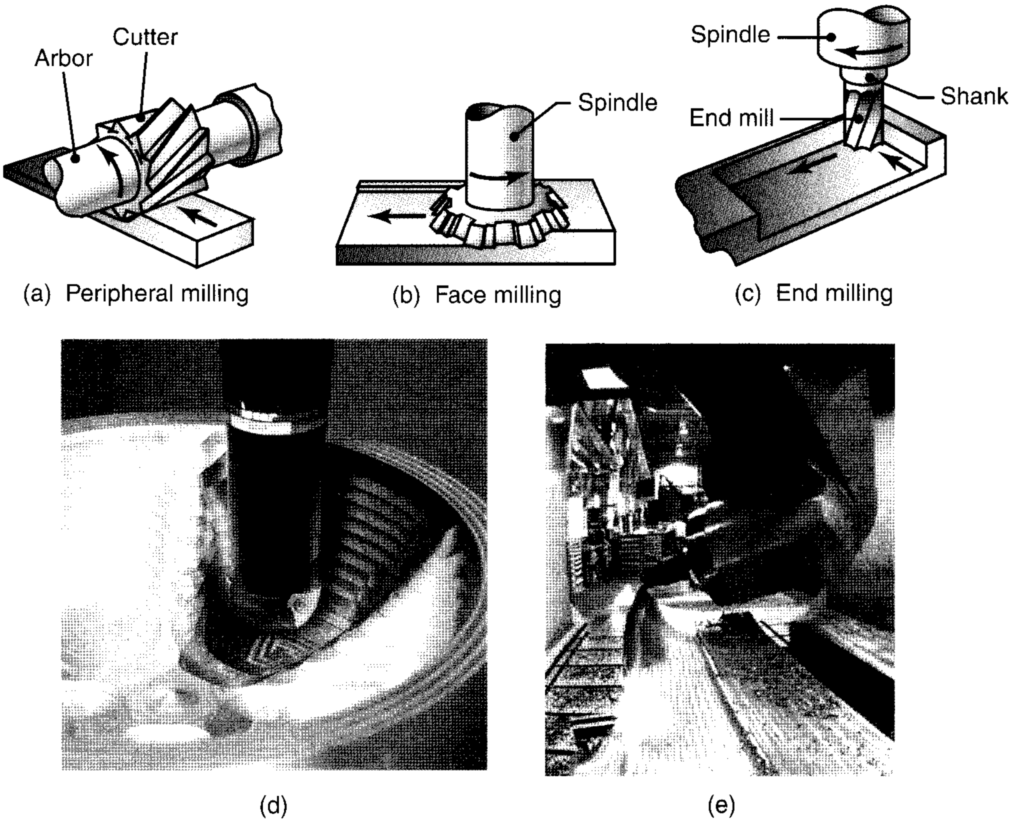


FIGURE 24.2 Some basic types of milling cutters and milling operations. (a) Peripheral milling. (b) Face milling. (c) End milling. (d) Ball-end mill with indexable coated-carbide inserts machining a cavity in a die block. (e) Milling a sculptured surface with an end mill using a five-axis numerical control machine. *Source*: (d) Courtesy of Iscar Metals, Inc. (e) Courtesy of The Ingersoll Milling Machine Co.

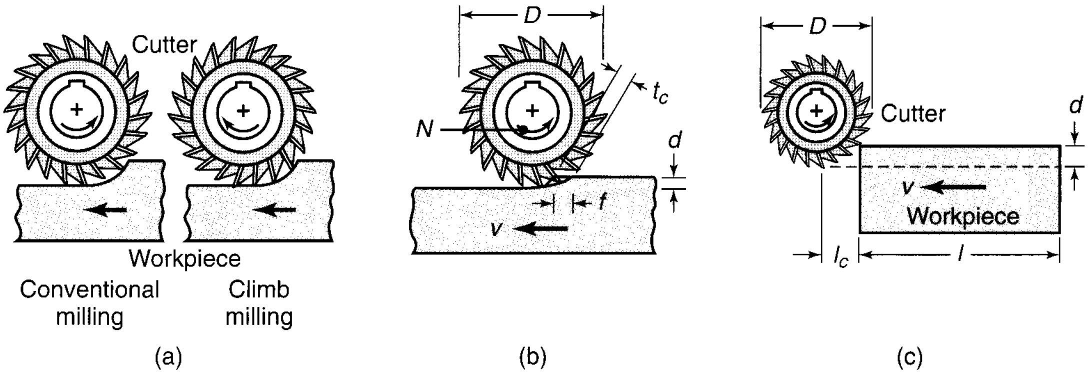


FIGURE 24.3 (a) Schematic illustration of conventional milling and climb milling. (b) Slab-milling operation showing depth of cut, d; feed per tooth, f; chip depth of cut, t_c; and workpiece speed, v. (c) Schematic illustration of cutter travel distance, l_c, to reach full depth of cut.

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. The advantages to conventional milling are that (a) tooth engagement is not a function of workpiece surface characteristics and (b) contamination or scale (oxide layer) on the surface does not adversely affect tool life. This is the more common method of milling. The cutting process is smooth, provided that the cutter teeth are sharp. Otherwise, the tooth will rub against and smear the surface for some distance before it begins to cut. Also, there may be a tendency for the tool to chatter, and the workpiece has a tendency to be pulled upward (because of the cutter rotation direction), necessitating proper clamping.

In *climb milling* (also called *down milling*), cutting starts at the surface of the workpiece where the chip is thickest. The advantage is that the downward component of the cutting force holds the workpiece in place, particularly for slender parts. However, because of the resulting impact forces when the teeth engage the workpiece, this operation must have a rigid work-holding setup, and gear backlash must be eliminated in the table feed mechanism. Climb milling is not suitable for the machining of workpieces having surface scale, such as hot-worked metals, forgings, and castings. The scale is hard and abrasive and causes excessive wear and damage to the cutter teeth, thus shortening tool life.

Milling Parameters. The cutting speed, V, in peripheral milling is the surface speed of the cutter, or

$$V = \pi DN, \tag{24.1}$$

where D is the cutter diameter and N is the rotational speed of the cutter (Fig. 24.4).

Note that the thickness of the chip in slab milling varies 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* (*chip depth of cut*),

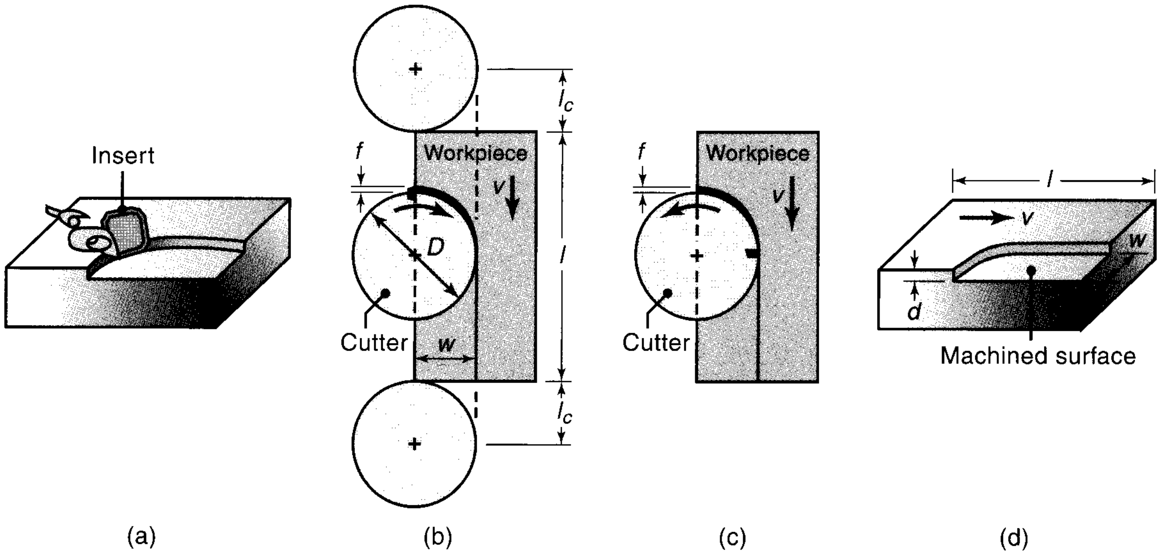


FIGURE 24.4 (a) Face-milling operation, showing action of an insert; (b) climb milling; (c) conventional milling; (d) dimensions in face milling. The width of cut, w, is not necessarily the same as the cutter radius.

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—that is, 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 increases.

Feed per tooth is determined from the equation

$$f = \frac{v}{Nn}, \tag{24.3}$$

where v is the linear speed (*feed rate*) of the workpiece and n is the number of teeth on the cutter periphery. The dimensional accuracy of this equation can be checked by using appropriate units for the individual terms; for instance,

$$(\text{mm/tooth}) = (\text{m/min})(10^3\ \text{mm/m})/(\text{rev/min})(\text{number of teeth/rev}).$$

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. 24.3c) 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 the same as 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 means such as faster travel of the machine tool components. The foregoing equations and the terminology used are summarized in Table 24.1.

TABLE 24.1

Summary of Peripheral Milling Parameters and Formulas

N = Rotational speed of the 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 = $\text{N} \cdot \text{m}$
= $F_cD/2$
Power = kW
= (Torque)(ω), where $\omega = 2\pi N$ radians/min

The *power requirement* in peripheral milling can be measured and calculated, but the *forces* acting on the cutter (tangential, radial, and axial; see also Fig. 23.5) are difficult to calculate because of the numerous variables involved, many pertaining to the cutting-tool geometry. These forces can be measured experimentally for a variety of conditions, and the *torque* on the cutter spindle can be calculated from the power. (See Example 24.1.) Although the torque is the product of the tangential force on the cutter and the cutter's radius, the tangential force per tooth will depend on how many teeth are engaged during the cut.

EXAMPLE 24.1 Material-removal Rate, Power, Torque, and Cutting Time in Slab Milling

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 $d = 3.0$ 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. 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):

$$v = fNn = (0.25)(100)(20) = 500 \text{ mm/min}.$$

From Eq. (24.5), the material-removal rate is calculated to be

$$\text{MRR} = (100)(3)(500) = 150{,}000 \text{ mm}^3/\text{min}.$$

Since the workpiece is annealed mild steel, let's estimate the unit power from Table 21.2 as $3 \text{ W}\cdot\text{s/mm}^3$. Hence, the power required can be estimated as

$$\text{Power} = (3)(150{,}000)\left(\tfrac{1}{60}\right) = 7.5 \text{ kW}$$

The torque acting on the cutter spindle also can be calculated by noting that power is the product of torque and the spindle rotational speed (in radians per unit time). Therefore,

$$\begin{aligned}\text{Torque} &= \frac{\text{Power}}{\text{Rotational speed}}\\ &= \frac{(7500)(60 \text{ N·m/min}\cdot\text{W})}{(100 \text{ rpm})(2\pi)}\\ &= 716 \text{ N·m}\end{aligned}$$

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

$$l_c = \sqrt{Dd} = \sqrt{(50)(3)} = 12.2 \text{ mm}.$$

Thus, the cutting time is

$$t = \frac{300 + 12.2}{500} = 0.62 \text{ min} = 37.2 \text{ s}.$$

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. 24.2b) and removes material in the manner shown in Fig. 24.4a. The cutter rotates at a rotational speed N, and the workpiece moves along a straight path at a linear speed v. When the cutter rotation is as shown in Fig. 24.4b, the operation is *climb milling*; when it is in the opposite direction (Fig. 24.4c), the operation is *conventional milling*. The cutting teeth, such as carbide inserts, are mounted on the cutter body as shown in Fig. 24.5 (see also Fig. 22.3c).

Because of the relative motion between the cutter teeth and the workpiece, face milling leaves *feed marks* on the machined surface (Fig. 24.6), similar to those left by turning operations (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 the various angles, is shown in Fig. 24.7. As can be seen from the side view of the insert in Fig. 24.8, the

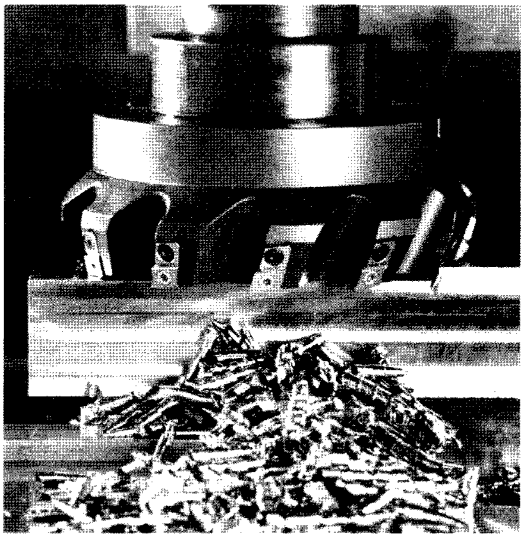


FIGURE 24.5 A face-milling cutter with indexable inserts. *Source*: Courtesy of Ingersoll Cutting Tool Company.

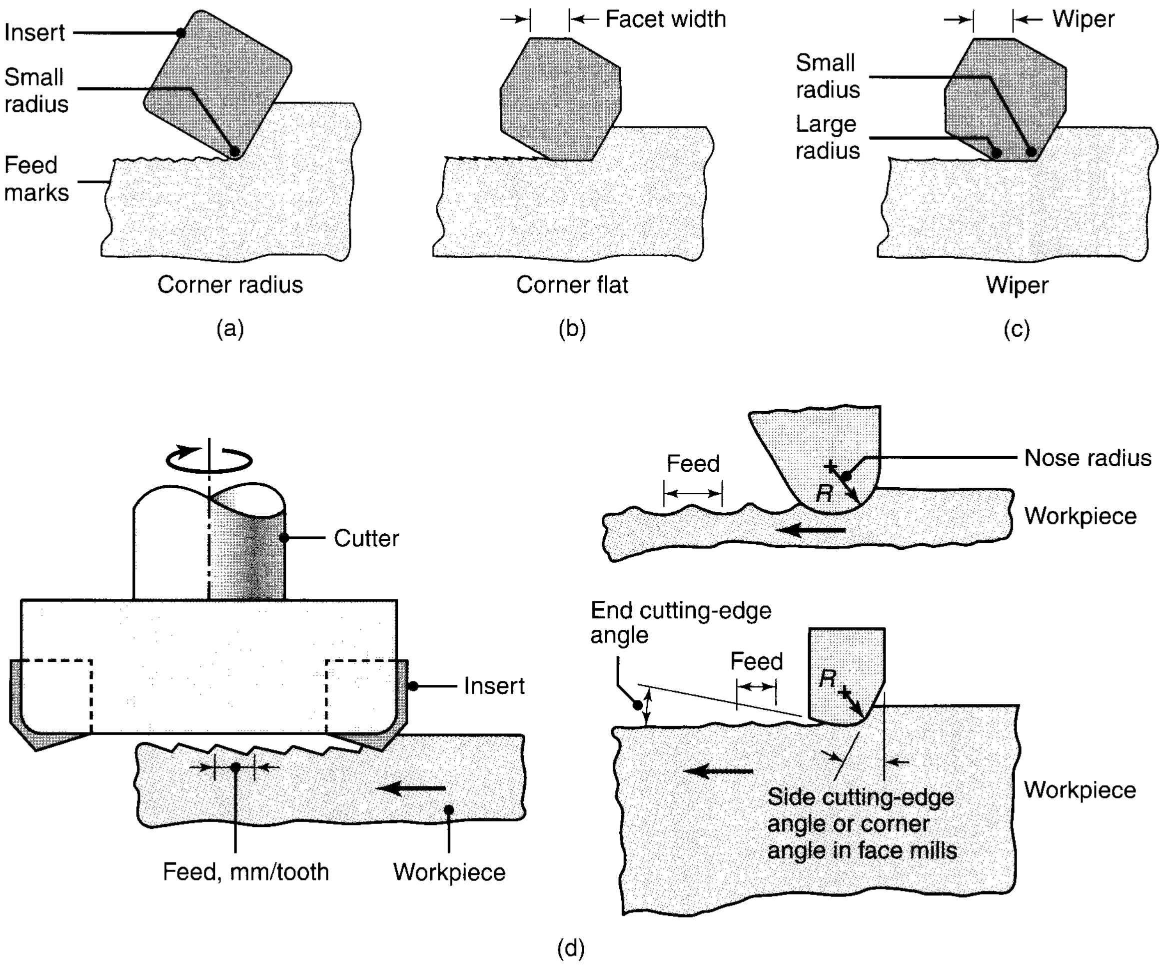


FIGURE 24.6 Schematic illustration of the effect of insert shape on feed marks on a face-milled surface: (a) small corner radius, (b) corner flat on insert, and (c) wiper, consisting of a small radius followed by a large radius, resulting in smoother feed marks. (d) Feed marks due to various insert shapes.

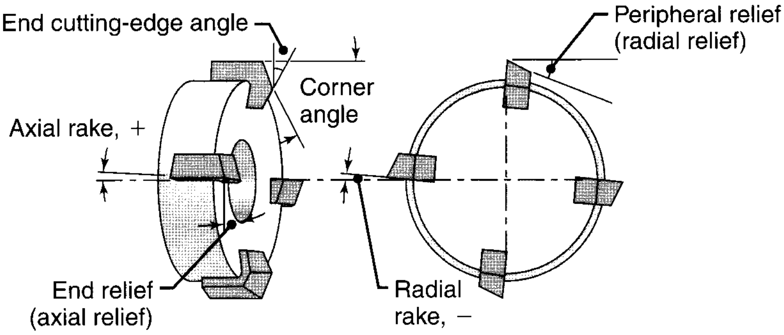


FIGURE 24.7 Terminology for a face-milling cutter.

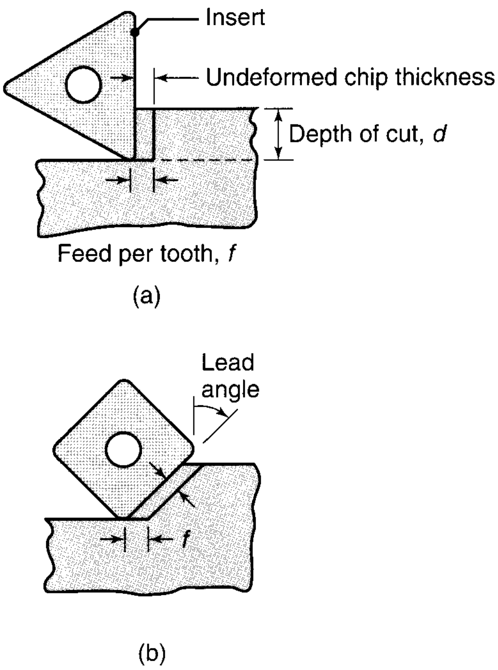


FIGURE 24.8 The effect of the lead angle on the undeformed chip thickness in face milling. Note that as the lead angle increases, the chip thickness decreases, but the length of contact (i.e., chip width) increases. The edges of the insert must be sufficiently large to accommodate the contact length increase.

lead angle of the insert in face milling has a direct influence on the *undeformed chip thickness*, as it does in turning operations. As the lead angle (positive, as shown in Fig. 24.8b) increases, the undeformed chip thickness decreases (as does chip thickness), and the length of contact (and hence chip width) increases. 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 (axial force on the cutter spindle). The *lead angles* for most face-milling cutters typically range from 0° to 45°. Note in Fig. 24.8 that the cross-sectional area of the undeformed chip remains constant.

A wide variety of milling cutters is available. The cutter diameter should be chosen such that it will not interfere with fixtures and other components in 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 is important in that it will determine the angle at which an insert *enters* and *exits* the workpiece. Note in Fig. 24.4b for climb milling that (if the insert has zero axial and radial rake angles, see Fig. 24.7) the rake face of the insert engages the workpiece directly. As seen in Figs. 24.9a 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. 24.9a that the tip of the insert makes the first contact, so there is a possibility for the cutting edge to chip off. In Fig. 24.9b, 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. Therefore, there is a lower tendency for the insert to fail, because the forces on the insert vary more slowly. Note from Fig. 24.7 that the radial and axial rake angles also will have an effect on this operation.

Figure 24.9c shows the exit angles for various cutter positions. Note that in the first two examples the insert exits the workpiece at an angle, thus causing the force on the insert to reduce to zero at a slower rate (which is desirable) than in the third example, where the insert exits the workpiece suddenly (which is undesirable for tool life).

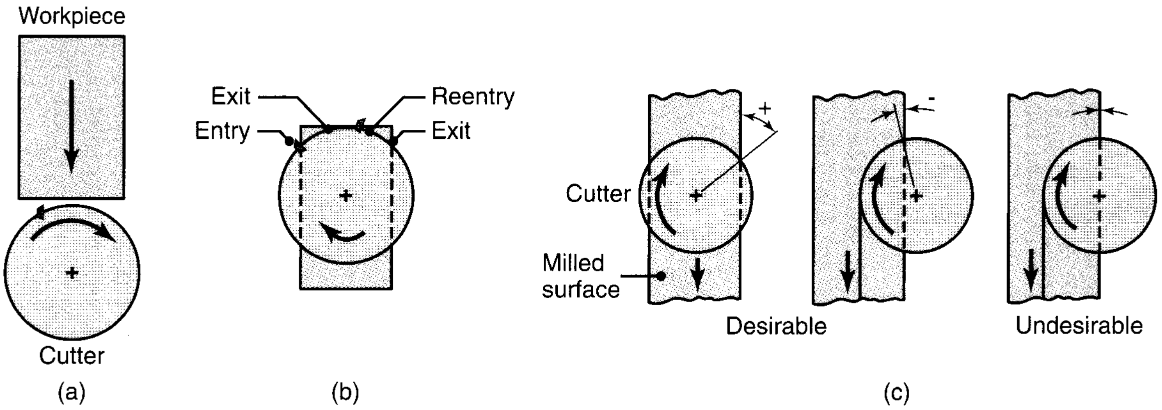


FIGURE 24.9 (a) Relative position of the cutter and insert as they first engage the workpiece in face milling. (b) Insert positions towards the end of cut. (c) Examples of exit angles of the insert, showing desirable (positive or negative angle) and undesirable (zero angle) positions. In all figures, the cutter spindle is perpendicular to the page and rotates clockwise.

EXAMPLE 24.2 Material-removal Rate, Power Required, and Cutting Time in Face Milling

Refer to Fig. 24.4 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. Calculate the material-removal rate, cutting time, and feed per tooth, and estimate the power required.

Solution First we note that the cross section of the cut is $wd = (60)(3) = 180\ \text{mm}^2$. 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. 24.4 that, for this problem, $l_c = \frac{D}{2} = 75$ mm. The cutting time is then

$$t = \frac{500 + 150}{10} = 65\ \text{s} = 1.08\ \text{min}.$$

The feed per tooth can be obtained from Eq. (24.3). Observing that $N = 100$ rpm $= 1.67$ rev/s, we find that

$$f = \frac{10}{(1.67)(10)} = 0.6\ \text{mm/tooth}.$$

For this material, let's estimate the unit power from Table 21.2 to be $1.1\ \text{W} \cdot \text{s/mm}^3$. Thus, the power is

$$\text{Power} = (1.1)(1800) = 1980\ \text{W} = 1.98\ \text{kW}.$$

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. 24.2c) 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 or with carbide inserts, similar to those for face milling. The cutter usually rotates on an axis perpendicular to the workpiece surface, and it also can be tilted to conform to machine-tapered or curved surfaces.

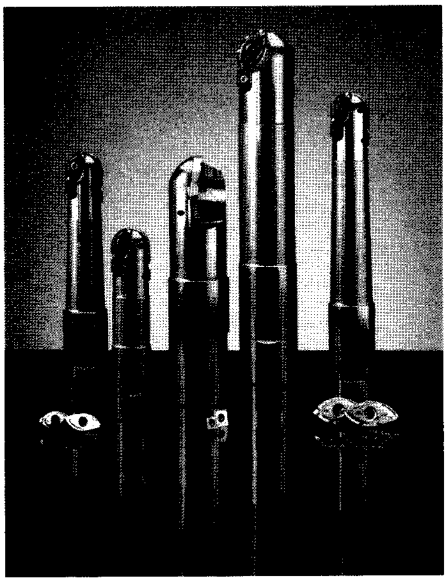
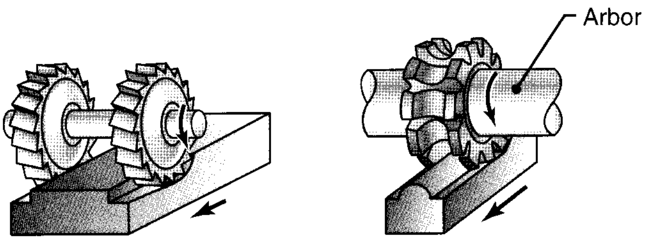


FIGURE 24.10 Ball nose end mills. These cutters are able to produce elaborate contours and are used often in the machining of dies and molds. (See also Fig. 24.2d.) *Source*: Courtesy of Dijet, Inc.

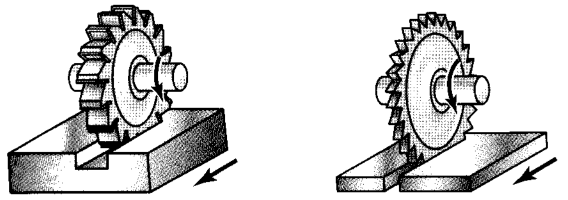
End mills are available with hemispherical ends (*ball nose mills*) for the production of sculptured surfaces, such as on dies and molds. *Hollow end mills* have internal cutting teeth and are used to machine the cylindrical surfaces of solid, round workpieces. End milling can produce a variety of surfaces at any depth, such as curved, stepped, and pocketed (Fig. 24.2d). The cutter can remove material on both its end and its cylindrical cutting edges, as can be seen in Fig. 24.2c. Both vertical spindle and horizontal spindle machines, as well as machining centers, can be used for end milling workpieces of various sizes and shapes. The machines can be programmed such that the cutter can follow a complex set of paths that optimize the whole machining operation for productivity and minimum cost.

High-speed End Milling. High-speed machining and its applications are described in Section 25.5. *High-speed end milling* has become an important process with numerous applications, such as the milling of large aluminum-alloy aerospace components and honeycomb structures with spindle speeds in the range from 20,000 to 60,000 rpm. The machines must have high stiffness and accuracy, usually requiring hydrostatic or air bearings, as well as high-quality work-holding devices. The spindles have a rotational accuracy of 10 μm; thus, the workpiece surface is also very accurate. At such high rates of material removal, chip collection and disposal can be a significant problem, as discussed in Section 23.3.7.

The production of cavities in metalworking dies (**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. 24.10). The machines have *four-axis* or *five-axis* movements (see, for example, Fig. 24.18) and are able to accommodate dies as large as 3 × 6 m and weighing 55 metric tons. It is not surprising, then, that such dies can cost 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 cutting tools (thus reducing the tendency for vibration), and (c) enable drilling of holes at various compound angles.



(a) Straddle milling (b) Form milling



(c) Slotting (d) Slitting

FIGURE 24.11 Cutters for (a) straddle milling, (b) form milling, (c) slotting, and (d) slitting with a milling cutter.

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 are used to machine two parallel surfaces on the workpiece (Fig. 24.11a). **Form milling** produces curved profiles using cutters that have specially shaped teeth (Fig. 24.11b). Such cutters are also used for cutting gear teeth, as described in Section 24.7. **Slotting** and **slitting** operations are performed with *circular cutters*, as shown in Figs. 24.11c and d, respectively. The teeth may be staggered slightly, like those in a saw blade (Section 24.5), to provide clearance for the cutter when making deep slots. *Slitting saws* are relatively thin, usually less than 5 mm. *T-slot cutters* are used to mill T-slots, such as those found in machine-tool worktables for clamping workpieces. As shown in Fig. 24.12a, a slot is first milled with an end mill; then the cutter machines the complete profile of the T-slot in one pass.

Key seat cutters are used to make the semicylindrical (or Woodruff) key seats for shafts. *Angle milling cutters* (single-angle or double-angle) are used to produce tapered surfaces with various angles. *Shell mills* (Fig. 24.12b) are hollow inside and are mounted on a shank; this allows the same shank to be used for different-sized cutters. The use of shell mills is similar to that of end mills. Milling with a single cutting tooth mounted on a high-speed spindle is known as *fly cutting*; generally, it is used in simple face-milling and boring operations. This 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. 23.23b.

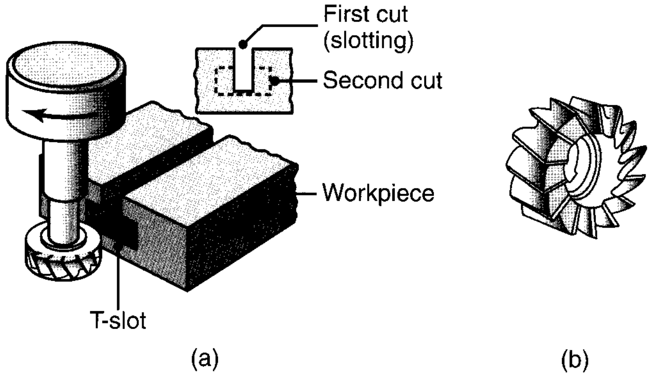


FIGURE 24.12 (a) T-slot cutting with a milling cutter. (b) A shell mill.

24.2.5 Toolholders

Milling cutters are classified as either arbor cutters or shank cutters. **Arbor cutters** are mounted on an *arbor* (see Figs. 24.11 and 24.15a), 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 end mills have tapered shanks for better mounting in the machine spindle in order to resist the higher forces and torque involved during cutting. Cutters with straight shanks are mounted in collet chucks or special end-mill holders; those with tapered shanks are mounted in tapered toolholders.

In addition to mechanical types, hydraulic toolholders and arbors are available. The stiffness of cutters and toolholders is important for surface quality and in reducing vibration and chatter during milling operations.

24.2.6 Milling Process Capabilities

In addition to the various characteristics of the milling processes described thus far, milling process capabilities include parameters such as surface finish, dimensional tolerances, production rate, and cost considerations. Data on process capabilities are presented in Tables 23.1 and 23.8, Figs. 23.13 and 23.14, and Chapter 40.

The conventional ranges of cutting speeds and feeds for milling are given in Table 24.2. Depending on the workpiece material and its condition, cutting-tool material, and process parameters, cutting speeds vary widely in the range from 30 to 3000 m/min. Feed per tooth typically ranges from about 0.1 to 0.5 mm, and depths of cut are usually 1 to 8 mm. For cutting-fluid recommendations, see Table 23.5.

A general **troubleshooting guide** for milling operations is given in Table 24.3; the last four items in this table are illustrated in Figs. 24.13 and 24.14. *Back striking* involves double feed marks 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).

TABLE 24.2

General Recommendations for Milling Operations

		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	120–180	0.085–0.38	90–425
Alloy steels					
Soft	Uncoated, coated, cermets	0.10–0.18	90–170	0.08–0.30	60–370
Hard	Cermets, PcBN	0.10–0.15	180–210	0.08–0.25	75–460
Cast iron, gray Soft	Uncoated, coated, cermets, SiN	0.10–10.20	120–760	0.08–0.38	90–1370
Hard	Cermets, SiN, PcBN	0.10–0.20	120–210	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	610–900	0.08–0.46	300–3000
High silicon	PCD	0.13	610	0.08–0.38	370–910
Copper alloys	Uncoated, coated, PCD	0.13–0.23	300–760	0.08–0.46	90–1070
Plastics	Uncoated, coated, 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 to 8 mm. PcBN: polycrystalline cubic-boron nitride. PCD: polycrystalline diamond. See also Table 23.4 for range of cutting speeds within tool material groups.

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
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

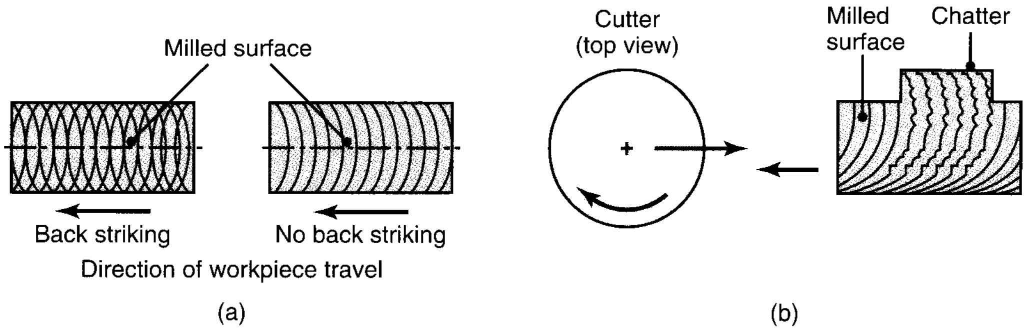


FIGURE 24.13 Machined surface features in face milling. (See also Fig. 24.6.)

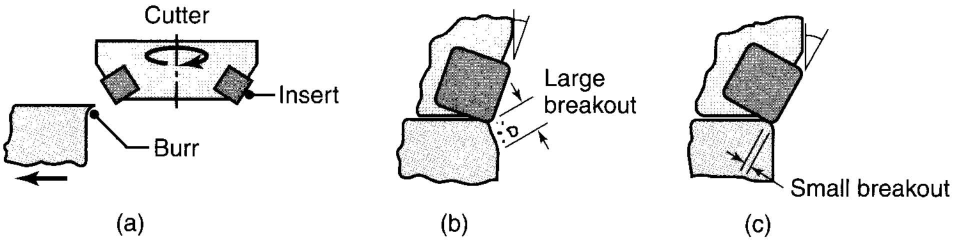


FIGURE 24.14 Edge defects in face milling: (a) burr formation along workpiece edge, (b) breakout along workpiece edge, and (c) how it can be avoided by increasing the lead angle (see also last row in Table 24.3).

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 operations include the following:

- Standard milling cutters should be used as much as possible, depending on part design features. Costly special cutters should be avoided.
- Chamfers should be specified instead of radii; it is difficult to smoothly match various intersecting surfaces if radii are specified.
- 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.
- Although small milling cutters allow the production of any surface, they are less rugged and more susceptible to chatter than large cutters. Thus, proper clearance should be provided in the design for milling cutters.
- Workpieces should be sufficiently rigid to minimize deflections that may result from clamping and cutting forces.

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 cases of vibration and chatter, tool shape and process conditions should be modified and cutters with fewer cutting teeth or with random tooth spacing should be used.

24.2.8 Milling Machines

Because they are capable of performing a variety of cutting operations, milling machines are among the most versatile and useful machine tools. The first milling machine was built in 1820 by Eli Whitney (1765–1825). A wide selection of milling machines with numerous features is now available. The features of typical standard milling machines are described next. Note, however, that many of these machines and operations are now being replaced with *computer controls* and *machining centers*. Inexpensive, manually controlled machines are still widely used, especially for small production runs.

Column-and-knee-type Machines. Used for general-purpose milling operations, *column-and-knee-type machines* are the most common milling machines. The spindle on which the milling cutter is mounted may be *horizontal* (Fig. 24.15a) for peripheral milling or *vertical* for face and end milling, boring, and drilling operations (Fig. 24.15b). The basic components of these machines are as follows:

- *Worktable:* on which the workpiece is clamped 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 can be adjusted vertically, and it can be swiveled in a vertical plane on the column for cutting tapered surfaces.

Plain milling machines have three axes of movement, with the motion usually imparted manually or by power. In *universal column-and-knee milling machines*, the table can be swiveled on a horizontal plane. In this way, complex shapes (such as helical grooves at various angles) can be machined to produce parts such as gears, drills, taps, and cutters.

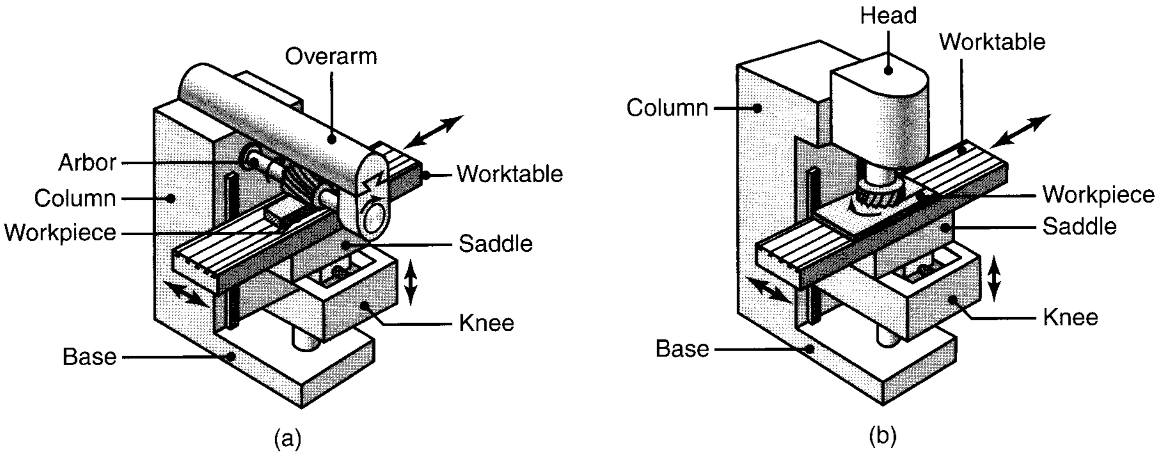


FIGURE 24.15 Schematic illustration of (a) a horizontal-spindle column-and-knee-type milling machine and (b) vertical-spindle column-and-knee-type milling machine. *Source*: After G. Boothroyd.

Bed-type Milling Machines. In *bed-type machines*, the worktable is mounted directly on the bed, which replaces the knee and can move only longitudinally (Fig. 24.16). These machines are not as versatile as other types, but they 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.

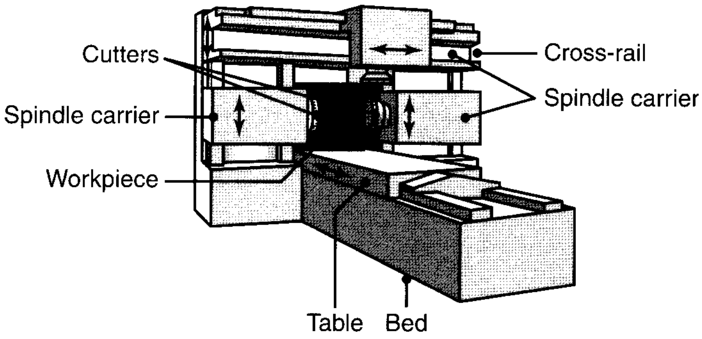


FIGURE 24.16 Schematic illustration of a bed-type milling machine.

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 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 milling machines and are equipped with one or more heads for face-milling operations.

Milling machines have been rapidly replaced by *computer numerical-control* (CNC) *machines* for all but the lowest production quantities. These machines are versatile and capable of milling, drilling, boring, and tapping with repetitive accuracy (Fig. 24.17). Also available are *profile milling machines*, which have five axes of

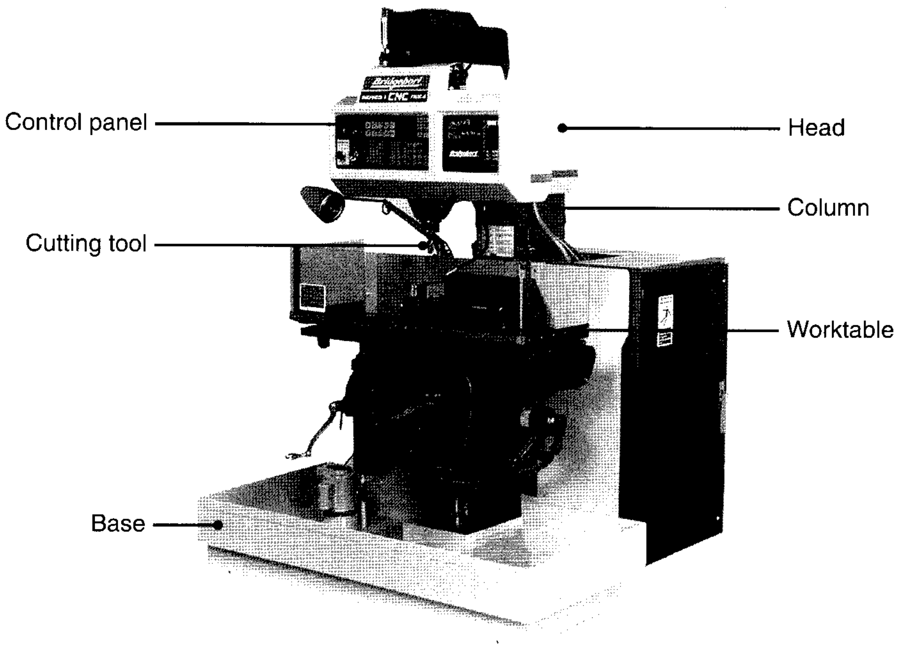


FIGURE 24.17 A computer numerical-control (CNC) vertical-spindle milling machine. This machine is one of the most versatile machine tools. The original vertical-spindle milling machine used in job shops is still referred to as a "Bridgeport," after its manufacturer in Bridgeport, Connecticut. *Source*: Courtesy of Bridgeport Machines Division, Textron, Inc.

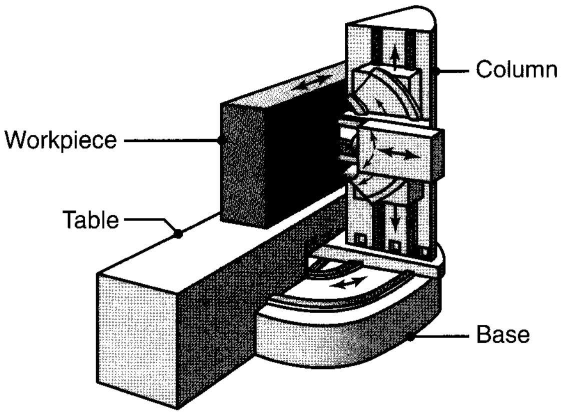


FIGURE 24.18 Schematic illustration of a five-axis profile milling machine. Note that there are three principal linear and two angular movements of machine components.

movement (Fig. 24.18); note the three linear and two angular movements of the machine components.

Work-holding Devices and Accessories. The workpiece to be milled must be clamped securely to the worktable in order to resist cutting forces and prevent slipping during milling. Various fixtures and vises generally are used for this purpose. (See also Section 37.8 on *flexible fixturing.*) Mounted and clamped to the worktable using the T-slots seen in Figs. 24.15a and b, vises are used for small production work on small parts and fixtures are used for higher production work and can be automated by various mechanical and hydraulic means.

Accessories for milling machines include various fixtures and attachments for the machine head (as well as the worktable) designed to adapt them to different milling operations. The accessory that has been used most commonly in the past, typically in job shops, 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. Dividing heads are now used only for low-volume job-shop quantities; they have been replaced by CNC controls and machining centers.

24.3 Planing and Shaping

Planing. This is a relatively simple machining operation by which flat surfaces, as well as cross sections with grooves and notches, are produced along the length of the workpiece (Fig. 24.19). Planing usually is done on large workpieces, as large as 25 × 15 m, although a length of 10 m is more typical. In a **planer**, the workpiece is mounted on a table that travels back and forth along a straight path. A horizontal cross-rail, which can be moved vertically along the ways of the column, is equipped with one or more tool heads. The cutting tools are mounted on the heads, and the machining is done along a straight path. In order to prevent tool cutting edges from chipping when they rub along a workpiece during the return stroke, tools are either tilted or lifted mechanically or hydraulically.

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. Otherwise, the chips produced can be very long, interfering with the planing operation and becoming a safety hazard.

Cutting speeds in planers can range up to 120 m/min, with power capacities of up to 110 kW. Recommended speeds for cast irons and stainless steels are in the range from 3 to 6 m/min and up to 90 m/min for aluminum and magnesium alloys. Feeds usually are in the range from 0.5 to 3 mm/stroke. The most common tool materials are M2 and M3 high-speed steels and K20 (C2) and P20 (C6) carbides.

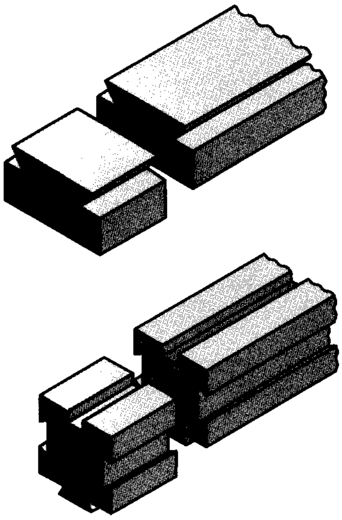


FIGURE 24.19 Typical parts that can be made on a planer.

Shaping. Machining by *shaping* is basically the same as by planing, except that (a) it is the tool, and not the workpiece, that travels and (b) workpieces are smaller, typically less than 1 × 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 (**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 with a long multiple-tooth cutter and is used to machine internal and external surfaces, such as holes of circular, square, or irregular section; keyways; the teeth of internal gears; multiple spline holes; and flat surfaces (Fig. 24.20). In a **broach** (Fig. 24.21a), the total depth 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 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. 24.21b. 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°;

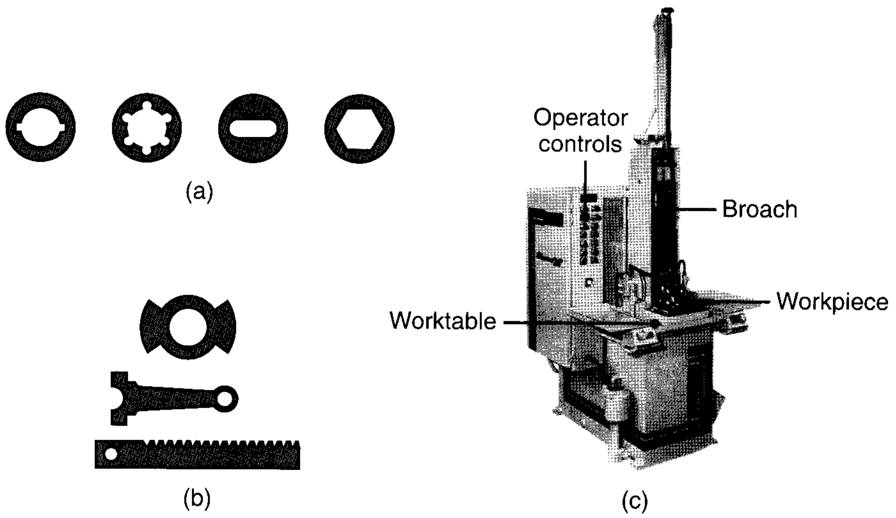


FIGURE 24.20 (a) Typical parts made by internal broaching. (b) Parts made by surface broaching. (c) Vertical broaching machine. *Source*: (a) and (b) Courtesy of General Broach and Engineering Company, (c) Courtesy of Ty Miles, Inc.

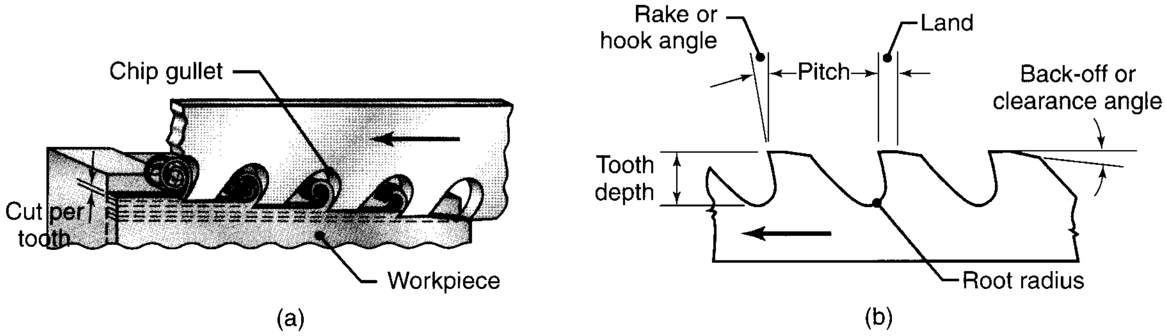


FIGURE 24.21 (a) Cutting action of a broach, showing various features. (b) Terminology for a broach.

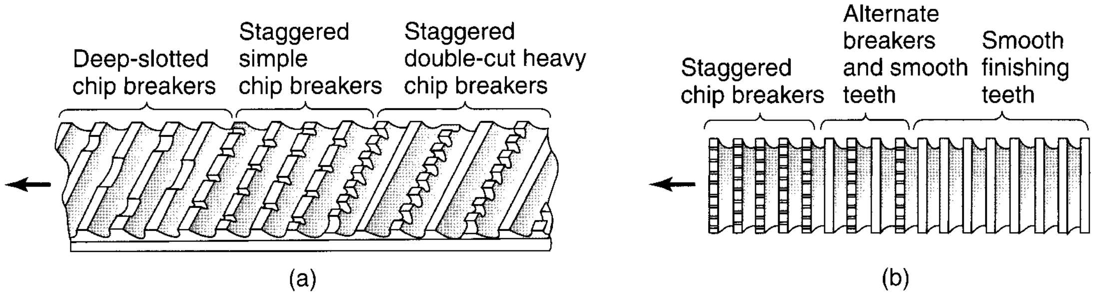


FIGURE 24.22 Chip breaker features on (a) a flat broach and (b) a round broach.

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. 24.22). 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. 24.23), *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. This process is typically used for broaching the bearing surfaces of crankshafts and similar parts. The crankshaft is rotated between centers, and the

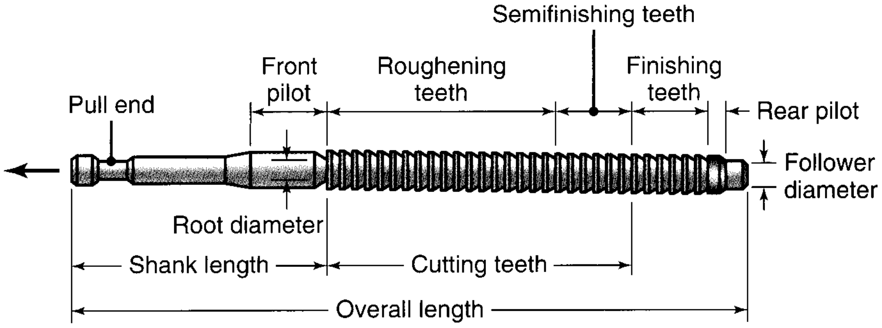


FIGURE 24.23 Terminology for a pull-type internal broach used for enlarging long holes.

broach, which is equipped with multiple carbide inserts, passes tangentially across the bearing surfaces and removes material. Turn broaching is a combination of *shaving* and *skiving* (removing a thin layer of material with a specially shaped cutting tool). Straight as well as circular broaches are used successfully in turn broaching. Machines that broach a number of crankshafts simultaneously have been built.

Broaching Machines. The machines for broaching are relatively simple in construction, have only linear motions, and usually are actuated hydraulically, although some are moved by crank, screw, or rack. Many styles of broaching machines are available, and sizes range from machines for making needle-like parts to those used for broaching gun barrels, including rifled (with internal spiral grooves) gun barrels. Broaching machines either pull or push the broaches and are either horizontal or vertical. *Push broaches* usually are shorter, generally in the range from 150 to 350 mm. *Pull broaches* tend to straighten the hole, whereas pushing permits the broach to follow any irregularity of the leader hole. Horizontal machines are capable of longer strokes. The *force* required to pull or push the broach depends on the (a) strength of the workpiece material, (b) total depth and width of cut, (c) cutting speed, (d) tooth profile, and (e) use of cutting fluids. The pulling force capacities of broaching machines are as high as 0.9 MN.

Process Parameters. Cutting speeds for broaching may range from 1.5 m/min for high-strength alloys to as much as 15 m/min for aluminum and magnesium alloys. The most common broach materials are M2 and M7 high-speed steels, as well as carbide inserts. The majority of broaches are now coated with titanium nitride for improved tool life and surface finish. Ceramic inserts also are used for finishing operations in some applications. Smaller, high-speed steel blanks for broaches can be made with powder-metallurgy techniques for better control of quality. Although carbide or ceramic inserts can be indexed after they are worn, high-speed steel broach teeth have to be resharpened by grinding, which reduces the size of the broach. 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 as follows:

- Parts should be designed so that they can be clamped securely in broaching machines. Parts should have sufficient structural strength and stiffness to withstand the cutting forces during broaching.

- The use of standardized parts is especially important for broaches. Keyways, splines, gear teeth, etc., all have standard sizes, and these dimensions allow the use of common broaches.
- Balanced cross sections are preferable to keep the broach from drifting, thus maintaining close tolerances.
- Radii are difficult to produce in broaching, and chamfers are preferred.
- Inverted or dovetail splines should be avoided.
- Broaching blind holes should be avoided whenever possible, but if necessary, there must be a relief at the end of the broached area.

EXAMPLE 24.3 Broaching Internal Splines

The part shown in Fig. 24.24 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 6,000 parts.

Source: ASM International.

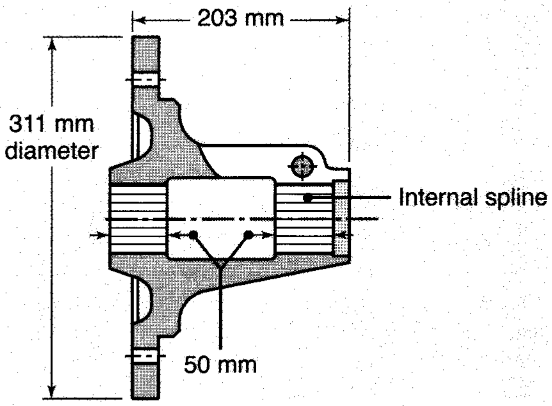


FIGURE 24.24 Example of a part with internal splines that were produced by broaching.

24.5 Sawing

Sawing is a common process dating back to around 1000 B.C. The cutting tool is a blade (saw) having a series of small teeth, each tooth removing a small amount of material with each stroke or movement of the saw. This process can be used for all metallic and nonmetallic materials and is capable of producing various shapes (Fig. 24.25). Sawing is an efficient bulk-removal process and can produce near-net shapes from raw materials. The width of cut (**kerf**) in sawing usually is small. Thus, the process wastes little material.

Typical saw-tooth and saw-blade configurations are shown in Fig. 24.26. Tooth spacing is generally in the range from 0.08 to 1.25 teeth per mm. A wide variety of sizes, tooth forms, tooth spacing, and blade thicknesses and widths are

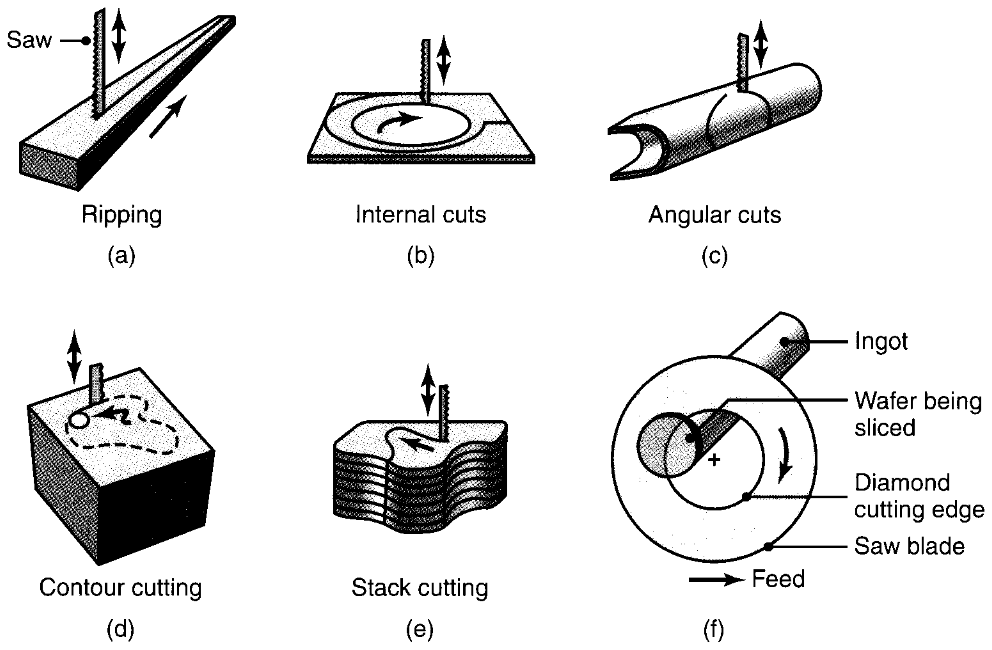


FIGURE 24.25 Examples of various sawing operations.

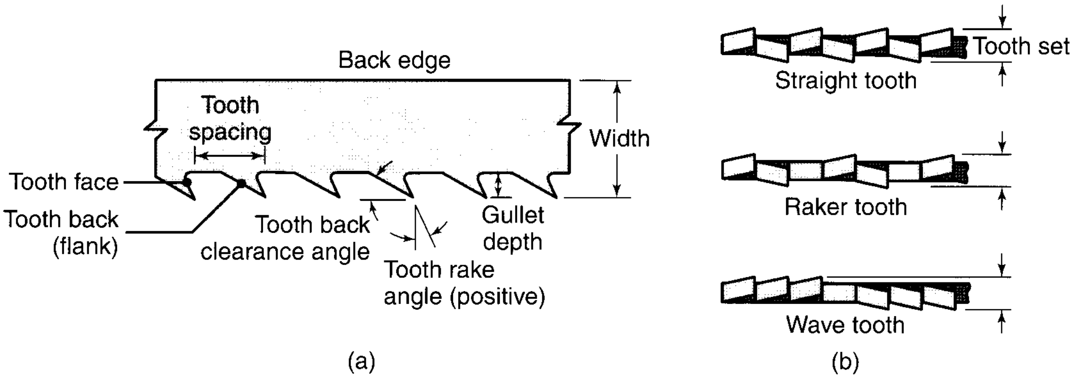


FIGURE 24.26 (a) Terminology for saw teeth. (b) Types of tooth sets on saw teeth staggered to provide clearance for the saw blade to prevent binding during sawing.

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 to saw harder materials (Fig. 24.27).

The **tooth set** in a saw (Fig. 24.26b) is important in providing a sufficiently wide kerf for the blade to move freely in the workpiece without binding or excessive frictional resistance, thus reducing the heat generated. Heat can have adverse effects on the cut, especially in cutting thermoplastics, which soften rapidly when heated. 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 always should be engaged with the workpiece in order to prevent *snagging* (catching of the saw tooth on the workpiece). This is 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 generally are used to improve the quality of the cut and the life of the saw.

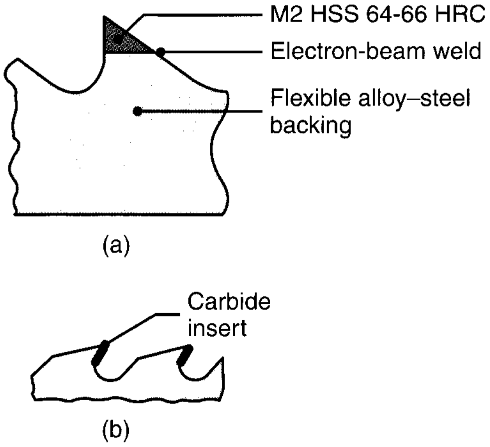


FIGURE 24.27 (a) High-speed-steel teeth welded onto a steel blade. (b) Carbide inserts brazed to blade teeth.

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 two reciprocating strokes, hacksaws are not as efficient as band saws (described later). *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* in cutting metal) generally are used for high-production-rate sawing, a process 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 large cross sections. Cold saws are available with a variety of tooth profiles and sizes and can be fed at any angle into the workpiece. In modern machines, cutting off with circular saws produces relatively smooth surfaces with good thickness control and dimensional accuracy due to the stiffness of the machines and of the saws. The inner-diameter cutting saw shown in Fig. 24.25f is used widely to cut single-crystal silicon wafers in microelectronic devices (see also Section 28.4).

Band saws have continuous, long, flexible blades and thus have a continuous cutting action. Vertical band saws are used for straight as well as *contour cutting* of flat sheets and other parts supported on a horizontal table (Fig. 24.25d). Also available are computer-controlled band saws with the capability of guiding the contour path automatically. Power band saws are available as well; they have higher productivity than power hacksaws because of their continuous cutting action. With high-speed steel blades, cutting 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**) so that the diamond particles act as cutting teeth (abrasive cutting); carbide particles also are used for this purpose. These 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. *Friction sawing* is a process in which a mild-steel blade or disk rubs against the workpiece at speeds of up to 7,600 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 in the workpiece produces a *heat-affected zone* (Section 30.9) on the cut surfaces. Thus, the workpiece properties along the cut edges can 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. This friction-sawing process is suitable for hard, ferrous metals and reinforced plastics but not for nonferrous metals, because of their tendency to stick to the blade. Friction-sawing disks as large as 1.8 m in diameter are used to cut off large steel sections. Friction sawing also is commonly used to remove flash from castings.

24.6 Filing

Filing involves the small-scale removal of material from a surface, corner, edge, or hole—including the removal of burrs. First developed around 1000 B.C., files usually are made of hardened steel and are available in a variety of cross sections, such as flat, round, half-round, square, and triangular. Files can have many tooth forms and grades of coarseness. Although filing usually is done by hand, filing machines with automatic features are available for high production rates, with files reciprocating at up to 500 strokes/min.

Band files consist of file segments, each about 75 mm long, that are riveted to a flexible steel band and are used in a manner similar to band saws. Disk-type files also are available.

Rotary files and **burs** (Fig. 24.28) 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 cutters generally are conical, cylindrical, or spherical in shape and have various tooth profiles. Their cutting action (similar to that of reamers) removes small amounts of material at high rates. The rotational speed of burs ranges from 1,500 rpm for cutting steels (large burs) to as high as 45,000 rpm for magnesium (small burs).

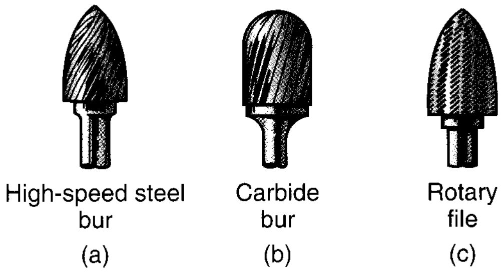


FIGURE 24.28 Types of burs used in burring operations.

24.7 Gear Manufacturing by Machining

Several processes for making gears or producing gear teeth on various components were described in Parts II and III: casting, forging, extrusion, drawing, thread rolling, and powder metallurgy, among others. 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 such processes as injection molding and casting.

Gears may be as small as those used in watches or as large as 9 m in diameter, for rotating mobile crane superstructures. The dimensional accuracy and surface finish required for gear teeth depend on the intended use. Poor gear-tooth quality contributes to inefficient energy transmission, increased vibration and noise, and adversely affects the gear's frictional and wear characteristics. Submarine gears, for example, have to be of extremely high quality so as to reduce noise levels, thus helping to avoid detection.

The standard nomenclature for an involute spur gear is shown in Fig. 24.29. Starting with a wrought or cast gear blank, there are two basic methods of making such gear teeth: *form cutting* and *generating*.

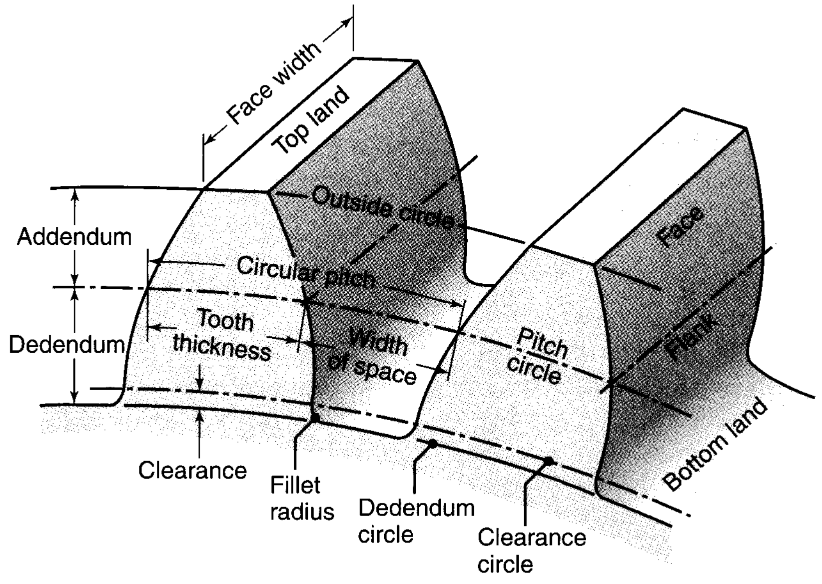


FIGURE 24.29 Nomenclature for an involute spur gear.

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. 24.30a). 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 until all of the teeth are machined.

Each cutter is designed to cut a range of numbers of teeth. The precision of the form-cut tooth profile depends on the accuracy of the cutter and on the machine and its stiffness. Although inefficient, form cutting can be done on milling machines with the cutter mounted on an arbor and the gear blank mounted in a dividing head. 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 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. 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 furthermore, 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.

Broaching also can be used to machine gear teeth and is particularly suitable for producing internal teeth. The broaching 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.

Gear teeth also may be cut on special machines with a single-point cutting tool that is guided by a *template* in the shape of the gear-tooth profile. Because the template can be made much larger than the gear tooth, dimensional accuracy is improved.

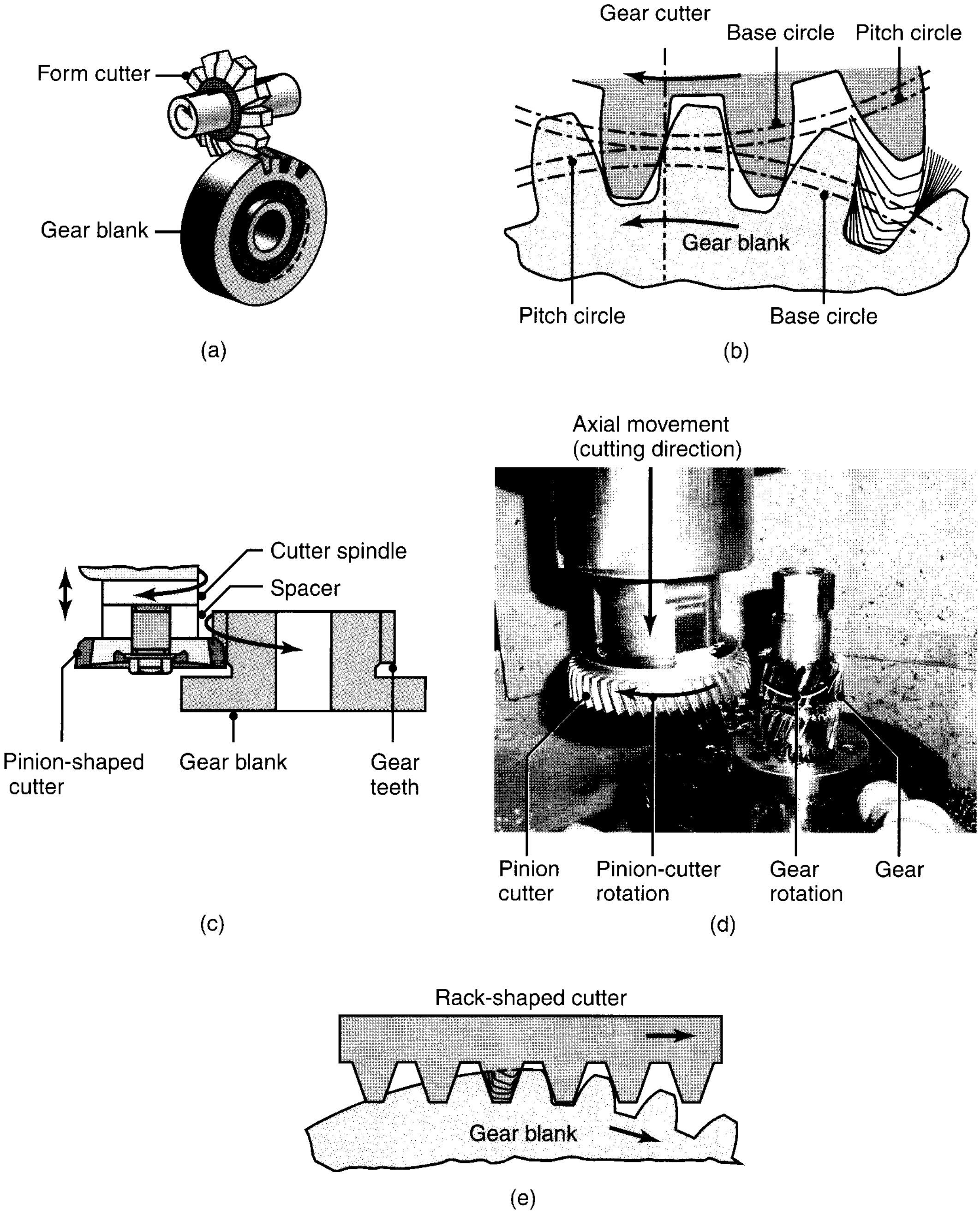


FIGURE 24.30 (a) Producing gear teeth on a blank by form cutting. (b) Schematic illustration of gear generating with a pinion-shaped gear cutter. (c) and (d) Gear generating in a gear shaper using a pinion-shaped cutter. Note that the cutter reciprocates vertically. (e) Gear generating with rack-shaped cutter. *Source*: (d) Courtesy of Schafer Gear Works, Inc.

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.

a. 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. 24.30b). This type of cutter is used on *gear shapers* (Figs. 24.30c and d). The cutter has an axis parallel

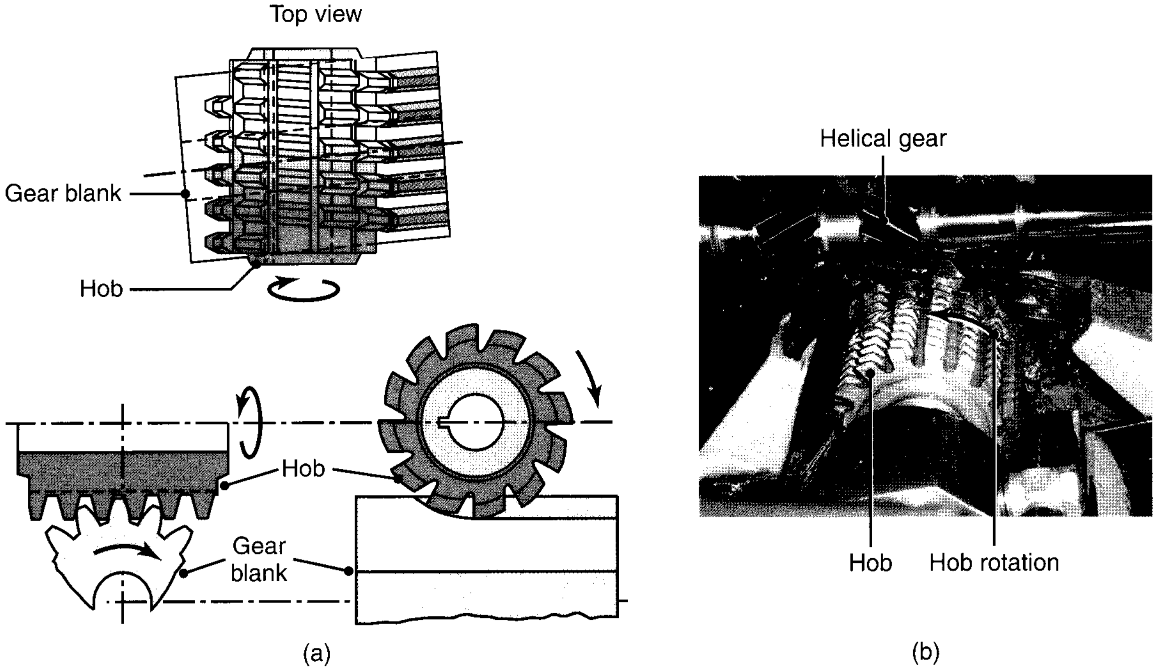


FIGURE 24.31 (a) Schematic illustration of gear cutting with a hob. (b) Production of a worm gear by hobbing. *Source*: Courtesy of Schafer Gear Works, Inc.

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 a flange in the gear blank in Figs. 24.30c and d. The process can be used for low-quantity as well as high-quantity production.

b. On a **rack shaper**, the generating tool is a *segment* of a rack (Fig. 24.30e), 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 the operation.

c. A **hob** (Fig. 24.31) 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 in 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. 24.29) 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 also is suitable for low-quantity production, it is most economical for medium- to high-quantity production.

Gear-generating machines also can produce spiral-bevel and hypoid gears. Like most other machine tools, modern gear-generating machines are computer controlled. *Multiaxis computer-controlled machines* are capable of generating many types and sizes of gears using indexable milling cutters.

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 to the proper shape on a gear generator. The generating method is analogous to the rack-generating method described previously. The cutters reciprocate across the face of the bevel gear as does the tool on a shaper (Fig. 24.32a). The machines for spiral bevel gears operate essentially on the same principle, and the spiral cutter is basically a face-milling cutter with a number of straight-sided cutting blades protruding from its periphery (Fig. 24.32b).

24.7.4 Gear-finishing Processes

As produced by any of the processes described previously, the surface finish and dimensional accuracy of gear teeth may not be sufficient for certain applications. Moreover, the gears may be noisy, or their mechanical properties (especially fatigue life and wear resistance) may not be sufficiently high. Several *finishing processes* are available to improve the surface quality of the gears. The choice of process is dictated by the method of gear manufacture 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 should be subjected to appropriate finishing operations.

Shaving. The gear-shaving process involves a cutter made in the exact shape of the finished tooth profile. The cutter removes small amounts of metal from the surface of the gear teeth. The cutter teeth are slotted or gashed at several points along its width, making the process similar to fine broaching. The cutter has a reciprocating motion. Shaving and burnishing (described next) can be performed only on gears with a hardness of 40 HRC or lower.

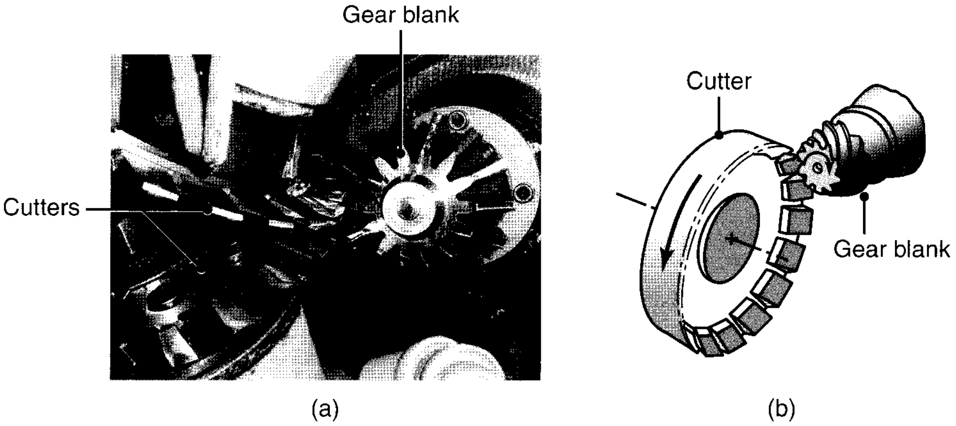


FIGURE 24.32 (a) Cutting a straight bevel-gear blank with two cutters. (b) Cutting a helical bevel gear. *Source*: Courtesy of Schafer Gear Works, Inc.

Although the tools are 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 a good dimensional accuracy of the tooth profile. Shaved gears subsequently may 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 special hardened, gear-shaped burnishing die that subjects the tooth surfaces to a surface-rolling action (**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.

Grinding, Honing, and Lapping. For the highest dimensional accuracy, tooth spacing and form, and surface finish, gear teeth subsequently may be ground, honed, and lapped, as described in Chapter 26. Specially dressed grinding wheels are used for either forming or 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. 24.33a). In **generating**, the grinding wheel acts in a manner similar to the gear-generating cutter described previously (Fig. 24.33b).

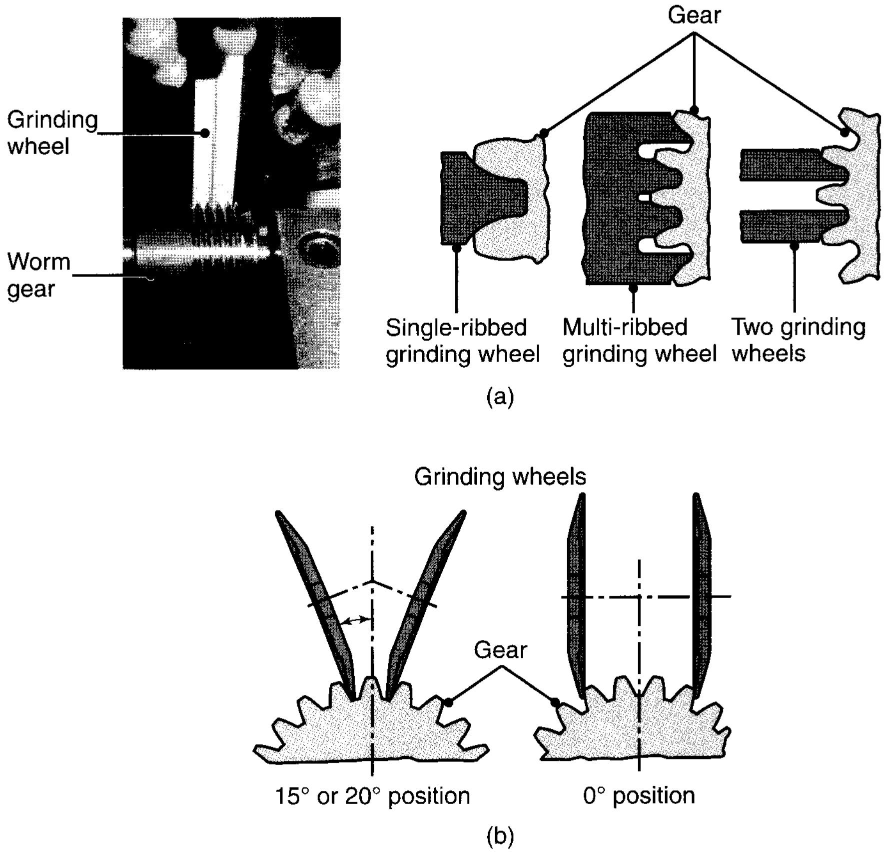


FIGURE 24.33 Finishing gears by grinding: (a) form grinding with shaped grinding wheels; (b) grinding by generating, using two wheels.

The **honing** tool is a plastic gear impregnated with fine abrasive particles. The honing process is faster than grinding and is used to improve surface finish. To further improve the surface finish, ground gear teeth are **lapped**, using abrasive compounds with either (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 suitable, particularly for producing 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 may be summarized as follows:

- Wide gears are more difficult to machine than narrow ones.
- Gears should be machined prior to their assembly on shafts.
- Sufficient clearance should be provided between gear teeth and flanges, shoulders, and other features, so that the cutting tool can machine without interference.
- Blank design is important for proper fixturing and to ease cutting operations. Machining allowances must be provided in blanks, and if machining is to be followed by subsequent finishing operations, the part must still be oversized after machining; that is, it has a finishing allowance after machining.
- Spur gears are easier to machine than helical gears, which in turn are easier to machine than bevel gears and 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 quality. Figure 24.34 shows the relative manufacturing cost of gears as a function of quality, as specified by AGMA (American Gear Manufacturers Association) and DIN (Deutsches Institut für Normung) numbers. 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.

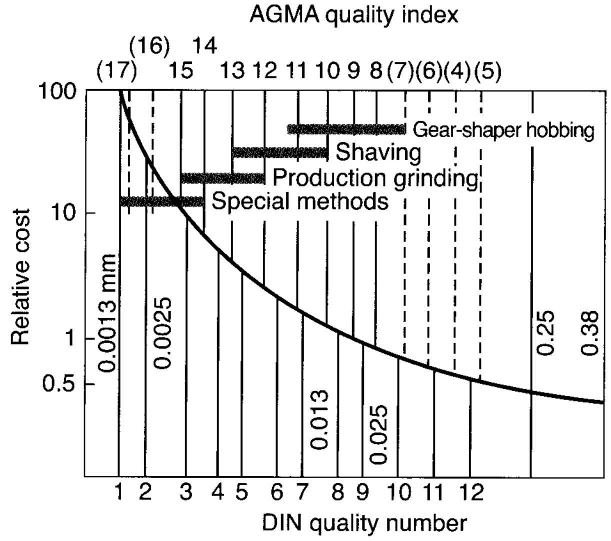


FIGURE 24.34 Gear manufacturing cost as a function of gear quality. The numbers along the vertical lines indicate tolerances.

CASE STUDY 24.1 Ping Golf Putters

In their efforts to develop high-end, top performing putters, engineers at Ping Golf, Inc., recently utilized advanced machining practices in their design and production processes for a new style of putter, the Anser® series. (Fig. 24.35). Following a unique set of design constraints, they had the task and goal of creating putters that would both be practical for production quantities and 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 unique to golf equipment). Among the materials evaluated, 303 stainless steel was chosen because it is a free-machining material (Section 21.7), indicating smaller chips, lower power consumption, a better surface finish, and an improved tool life, thus allowing for increased machining speeds and higher productivity.

The next step of the project involved determining the optimum blank type and the sequence of operations to be performed in its production. In 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; thus, dimensions were specified approximately 1.25–1.9 mm larger in all directions than that of the final part.

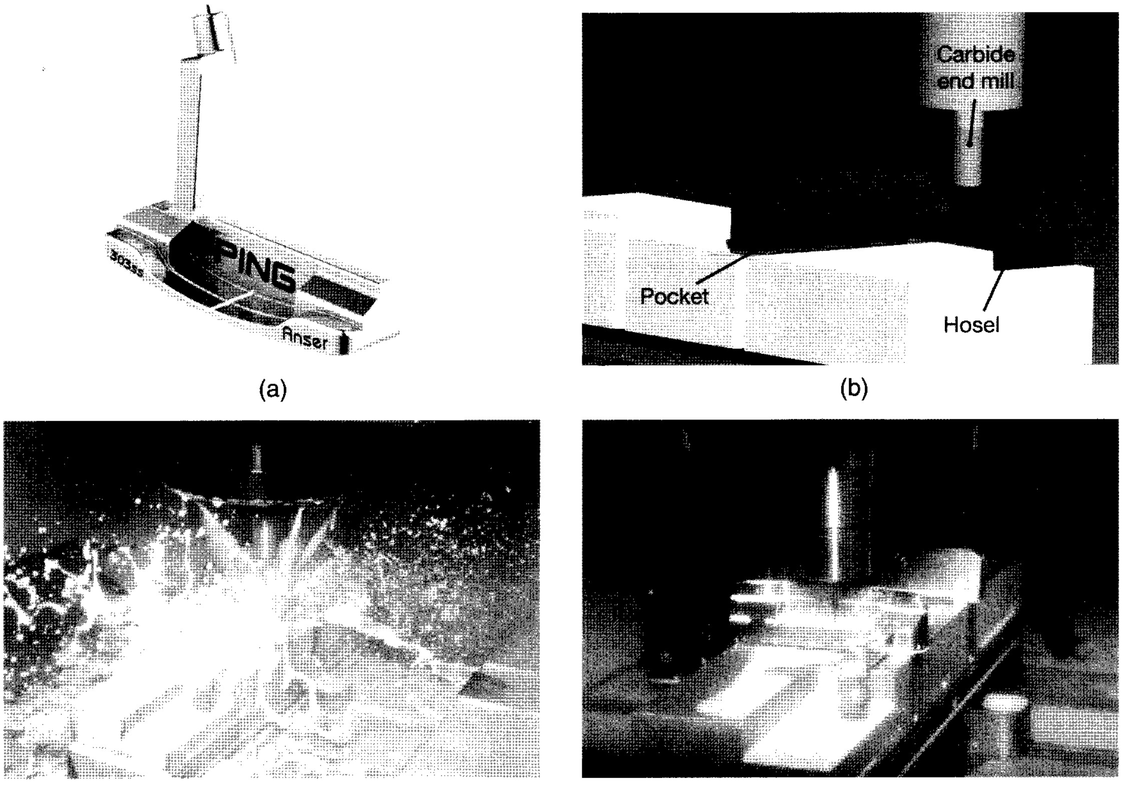


FIGURE 24.35 The Ping Anser® golf putter; (b) CAD model of rough machining of the putter outer surface; (c) rough machining on a vertical machining center; (d) machining of the lettering in a vertical machining center; the operation was paused to take the photo, as normally the cutting zone is flooded with a coolant. *Source*: Courtesy of Ping Golf, Inc., Phoenix, Arizona.

The most challenging and longest task of the project was developing the necessary programming and fixtures for each part. 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 injection-mold making) was used to machine most of the finished geometry. Although this operation required additional machining time, 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 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 (see 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 tungsten-carbide cutting tools coated with TiAlN (Section 22.5.2) allowed the quick changeover between right- and left-handed parts as well as different putter models. After the initial locating operation was complete, parts were transferred to a three-axis vertical machining center (see, for example, Fig. 25.7) to create 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 38 mm inside the "cavity" of the putter was the most demanding milling operation (see Figs. 24.35b and c). The putter has small inside radii with a comparatively long depth (7× diameter or greater).

A four-axis horizontal machining center (see, e.g., Fig. 25.2) was used to reduce the number of setups in the 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 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 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 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.

SUMMARY

- Various 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 and sizes from a large number of workpiece materials.
- Although these processes are similar to turning, drilling, and boring, and involve similar metal-cutting mechanics, tool materials, and cutting fluids, most of the processes described in this chapter utilize multitooth tools and cutters at various axes with respect to the workpiece.
- Machine tools used to produce complex shapes are mostly computer controlled, having various dedicated features, and imparting much more flexibility in their application.

- Broaching is a method of accurately enlarging a round hole or other profile in a workpiece. Broaching uses either pull or push broaches. Sawing is the gradual removal of material by small teeth spaced on a saw and is very common and 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 the various forming and shaping processes described in preceding chapters, gears 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 improved further by finishing processes, such 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

Arnone, M., **High Performance Machining,** Hanser, 1998.

ASM Handbook, Vol. 16: **Machining,** ASM International, 1989.

Boothroyd, G., and Knight, W.A., **Fundamentals of Machining and Machine Tools,** 3rd ed., Marcel Dekker, 2005.

Brown, J., **Advanced Machining Technology Handbook,** McGraw-Hill, 1998.

Davis, J.R., (ed.), **Gear Materials, Properties and Manufacture,** ASM International, 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., and Check, A.F., **Technology of Machine Tools,** 6th ed., Glencoe Macmillan/McGraw-Hill, 2009.

Lopez, L.N., and Lamikiz, A. (eds.), **Machine Tools for High Performance Machining,** Springer, 2009.

Machinery's Handbook, Industrial Press, revised periodically.

Machining Data Handbook, 3rd ed., 2 vols. Machinability Data Center, 1980.

Nachtman, E.S., and Kalpakjian, S., **Lubricants and Lubrication in Metalworking Operations,** Marcel Dekker, 1985.

Stephenson, D.A., and Agapiou, J.S., **Metal Cutting: Theory and Practice,** 2nd ed., Marcel Dekker, 2006.

Tool and Manufacturing Engineers Handbook, 4th ed., Vol. 1: **Machining,** Society of Manufacturing Engineers, 1983.

Walsh, R.A., **McGraw-Hill Machining and Metalworking Handbook.** 3rd ed., McGraw-Hill, 2006.

Weck, M., **Handbook of Machine Tools,** 4 vols., Wiley, 1984.

REVIEW QUESTIONS

▶**24.1.** Explain why milling is such a versatile machining operation.

▶**24.2.** Describe the different types of cutters used in milling operations and give an application of each type.

▶**24.3.** Explain the relative characteristics of climb milling and up milling.

▶**24.4.** Describe the geometric features of a broach and explain their functions.

▶**24.5.** Why is sawing a commonly used process? Why do some saw blades have staggered teeth? Explain.

▶**24.6.** What advantages do bed-type milling machines have over column-and-knee-type machines for production operations?

▶**24.7.** Explain why the axis of a hob is tilted with respect to the axis of the gear blank.

▶**24.8.** What is a shell mill? Why is it used?

▶**24.9.** Why is it difficult to saw thin sheet metals?

QUALITATIVE PROBLEMS

▶**24.10.** Would you consider the machining processes described in this chapter to be near-net or net-shape processing? Explain with appropriate examples.

▶**24.11.** Why is end milling such an important versatile process? Explain with examples.

▶**24.12.** List and explain factors that contribute to poor surface finish in the processes described in this chapter.

▶**24.13.** Explain why broaching crankshaft bearings is an attractive alternative to other machining processes.

▶**24.14.** Several guidelines are presented in this chapter for various cutting operations. Discuss the reasoning behind these guidelines.

▶**24.15.** What are the advantages of helical teeth over straight teeth on cutters for slab milling?

▶**24.16.** Explain why hacksaws are not as productive as band saws.

▶**24.17.** What similarities and differences are there in slitting with a milling cutter and with a saw?

▶**24.18.** Why do machined gears have to be subjected to finishing operations? Which of the finishing processes are not suitable for hardened gear teeth? Why?

▶**24.19.** How would you reduce the surface roughness shown in Fig. 24.6? Explain.

▶**24.20.** Why are machines such as the one shown in Fig. 24.17 so useful?

▶**24.21.** Comment on your observations concerning the designs illustrated in Fig. 24.20b and on the usefulness of broaching operations.

▶**24.22.** Explain how contour cutting could be started in a band saw, as shown in Fig. 24.25d.

▶**24.23.** In Fig. 24.27a, 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.

▶**24.24.** Describe the parts and conditions under which broaching would be the preferred method of machining.

▶**24.25.** With appropriate sketches, explain the differences between and similarities among shaving, broaching, and turn-broaching operations.

▶**24.26.** Explain the reason that it is difficult to use friction sawing on nonferrous metals.

▶**24.27.** Would you recommend broaching a keyway on a gear blank before or after machining the gear teeth? Why?

QUANTITATIVE PROBLEMS

▶**24.28.** In milling operations, the total cutting time can be significantly influenced by (a) the magnitude of the noncutting distance, l_c, shown in Figs. 24.3 and 24.4, 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.

▶**24.29.** 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.

▶**24.30.** Show that the distance l_c in slab milling is approximately equal to $\sqrt{Dd}$ for situations where $D \gg d$. (See Fig. 24.3c.)

▶**24.31.** In Example 24.1, which of the quantities will be affected when the feed is increased to $f = 0.5$ mm/tooth?

▶**24.32.** Calculate the chip depth of cut, t_c, and the torque in Example 24.1.

▶**24.33.** 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.

24.34. A 300-mm-long, 25-mm-thick plate is being cut on a band saw at 45 m/min. The saw has 12 teeth per 250 mm. If the feed per tooth is 0.075 mm, how long will it take to saw the plate along its length?

24.35. 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.

▶**24.36.** Assume that in the face-milling operation shown in Fig. 24.4 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 $5\text{W} \cdot \text{s/mm}^3$ and that only 75% of the cutter diameter is engaged during cutting. Calculate (a) the power required and (b) the material-removal rate.

▶**24.37.** A slab-milling operation will take place on a part 300 mm long and 40 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 0.75 m/s, find the machining time and metal-removal rate for removing 6 mm from the surface of the part.

24.38. Explain whether the feed marks left on the workpiece by a face-milling cutter (as shown in Fig. 24.13a) are segments of true circles. Describe the parameters you consider in answering this question.

24.39. In describing the broaching operations and the design of broaches, we have not given equations regarding feeds, speeds, and material-removal rates, as we have done in turning and milling operations. Review Fig. 24.21 and develop such equations.

SYNTHESIS, DESIGN, AND PROJECTS

24.40. The part shown in Fig. 24.1f is 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.

24.41. In Problem 24.40, would you prefer to machine this part 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?

24.42. If expanded honeycomb panels (see Section 16.12) 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.

24.43. 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.

24.44. Suggest methods whereby milling cutters of various designs (including end mills) can incorporate carbide inserts.

24.45. 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.

24.46. On the basis of the data developed in Problem 24.45, describe your thoughts regarding the procedure to be followed in determining what type of machine tool to select when machining a particular part.

24.47. 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.

Machining Centers, Machine-Tool Structures, and Machining Economics

CHAPTER 25

- This chapter presents the characteristics, types, and advantages of machining centers, and the concept of reconfigurable machine tools.
- The design and materials for machine tools are also presented.
- Emphasis is placed on the importance of understanding the performance of these machines, their modules, and their components, particularly with regard to their stiffness, vibration, chatter, and damping characteristics, important not only for quality and dimensional accuracy, but also because of their influence on tool life, productivity, and overall machining economics.
- Also described are high-speed, hard, and ultraprecision machining operations, topics that are tied strongly to the economics of machining.
- The chapter ends with a discussion of 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 have described machining operations and machine tools, but have not emphasized the widespread integration of computer technology into machining and the flexibility it allows in manufacturing operations. Computers have become common during the past 40 years and have dramatically improved the capabilities of machine tools. As described in this chapter, modern machine tools have the capability of rapidly producing extremely complex part geometries. The program controlling a machine tool can incorporate changes in cutting conditions, compensate for tool wear, automatically change tools, and machine a workpiece without refixturing or transferring it to another machine tool, as had been the practice for many years.

In addition to the aforementioned computer technologies, advanced machining concepts, such as 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 the 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 type of operation, such as turning, boring, drilling, milling, broaching, planing, or shaping. Also, it was shown that most parts manufactured by the methods described throughout this book require further 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 that all of the surfaces on these parts require a different type of machining operation (such as milling, facing, boring, drilling, reaming, or threading) to meet certain specific requirements concerning shapes, features, dimensional tolerances, and surface finish. A brief review will lead to 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 need some additional finishing operations. For example, deep small-diameter holes, threaded holes, flat surfaces for sealing with gaskets, very close dimensional tolerances, sharp corners and edges, and flat or curved surfaces with different surface-finish requirements will require further machining.
- If some form of machining is required or if it is shown to be more economical to finish machine these parts to their final shapes, then it is obvious that none of the machine tools described in Chapters 23 and 24 *individually* could produce the parts completely. Also, note that, traditionally, machining operations are performed by moving the workpiece from one machine tool to another until all of the required machining operations are completed.

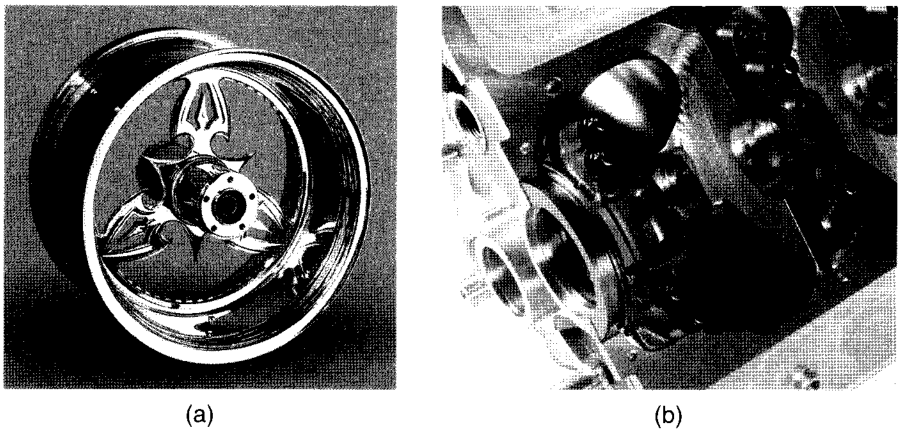

(a) (b)

FIGURE 25.1 Examples of parts that can be machined on machining centers using various processes such as turning, facing, milling, drilling, boring, reaming, and threading. Such parts ordinarily would require the use of a variety of machine tools to complete. (a) Forged motorcycle wheel, finish machined to tolerance and subsequently polished and coated; (b) detailed view of an engine block, showing complex cavities, threaded holes, and planar surfaces. *Source*: (a) Courtesy of R.C. Components, (b) Courtesy of Donovan Engineering, programming by N. Woodruff, and Photography by E. Dellis, Powersports Photography.

The Concept of Machining Centers. The traditional method of machining parts by using different types of machine tools has been, and continues to be, a viable and efficient manufacturing method. It can be highly automated in order to increase productivity, and it is indeed the principle behind **transfer lines** (also called *dedicated manufacturing lines,* or DML), as described in Section 37.2.4. Commonly used in *high-volume* or *mass production,* transfer lines consist of several specific machine tools arranged in a logical sequence. The workpiece (such as an automotive engine block) is moved from station to station, with a specific machining operation performed at each station, after which it is transferred to the next machine for further specific machining operations.

There are situations, however, in which 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 changes in product shape or style. It is a very expensive and time-consuming process to rearrange these machine tools to respond to the needs for the next 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 different 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 milling, drilling, honing, tapping, and similar operations. Whereas, in transfer lines or in typical shops and factories, the workpiece is brought *to the machine,* in machining centers it is the machining operation that is brought *to the workpiece.*

In using the word *workpiece,* we should point out that the workpiece in a machining center also can consist of all types of **tooling.** Tooling includes forming and cutting tools, cutters and toolholders, tool shanks for holding tool inserts, molds for casting, male and female dies for forming, punches for metalworking and powder metallurgy, rams for extrusion, work-holding devices, and fixturing—all of which also have to be manufactured. Since the geometries often are quite complicated and a variety of machining operations are necessary, these tools commonly are produced in machining centers.

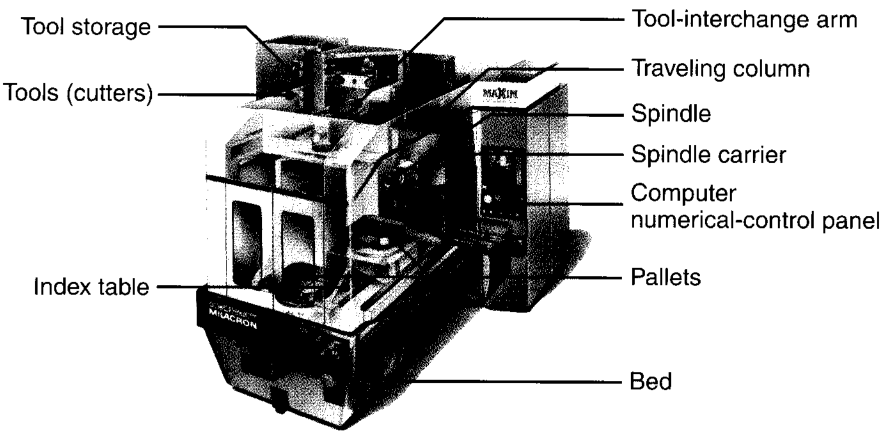


FIGURE 25.2 A horizontal-spindle machining center equipped with an automatic tool changer. Tool magazines can store up to 200 cutting tools of various functions and sizes. *Source*: Courtesy of Cincinnati Milacron, Inc.

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 (**turning center**) with two turrets carrying several cutting tools.

Components of a Machining Center. The workpiece in a machining center is placed on a **pallet**, or *module*, that 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 to be machined) is brought into position by an **automatic pallet changer** (Fig. 25.4). All movements are computer controlled, and pallet-changing cycle times are on the order of only 10 to 30 seconds. Pallet stations are available with several pallets serving the machining center. The machines also can be equipped with various automatic features, such as part loading and unloading devices.

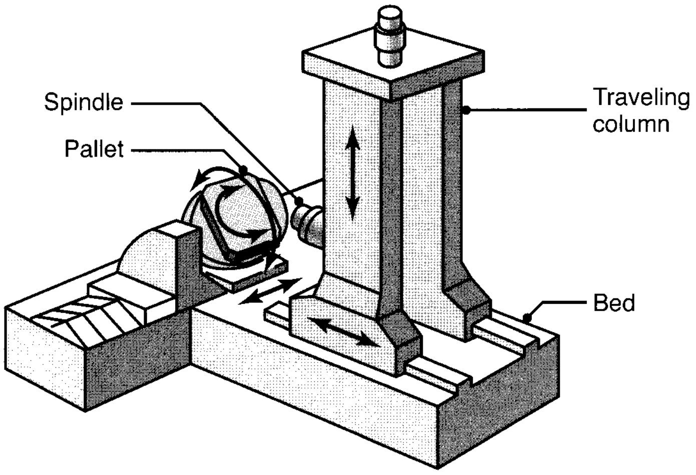


FIGURE 25.3 Schematic illustration of the principle of a five-axis machining center. Note that, in addition to possessing three linear movements (three axes), the pallet, which supports the workpiece, can be swiveled around two axes (hence a total of five axes), allowing the machining of complex shapes, such as those shown in Fig. 25.1. *Source*: Courtesy of Toyoda Machinery.

A machining center is equipped with a programmable **automatic tool changer** (ATC). Depending on the particular design, up to 200 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. 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 known as the **work envelope**—a term

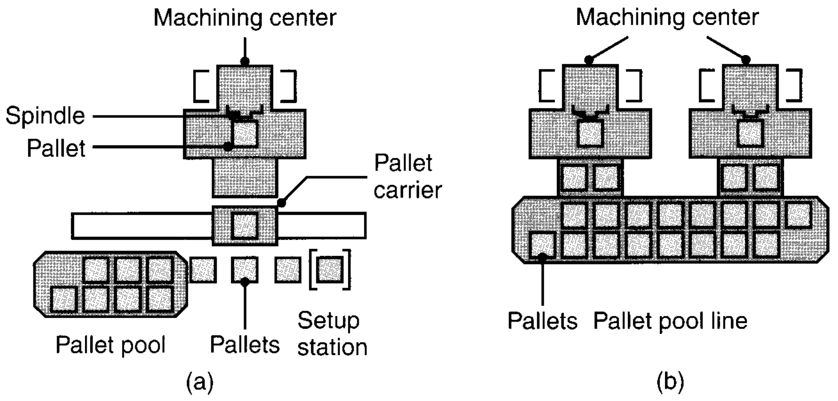


FIGURE 25.4 (a) Schematic illustration of the top view of a horizontal-spindle machining center showing the pallet pool, setup station for a pallet, pallet carrier, and an active pallet in operation (shown directly below the spindle of the machine). (b) Schematic illustration of two machining centers with a common pallet pool. Various other pallet arrangements are possible in such systems. *Source*: Courtesy of Hitachi Seiki Co., Ltd.

that first was used in connection with industrial robots, as will be described in Section 37.6.

The **tool-exchange arm** shown in Fig. 25.5 is a common design; it swings around to pick up a particular 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, coded tags, or memory chips attached directly to their toolholders. *Tool-changing times* are typically between 5 and 10 seconds, but may be up to 30 seconds for tools weighing up to 110 kg. Because tool changing is a noncutting operation, the continuing trend is to reduce the times even more.

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 the

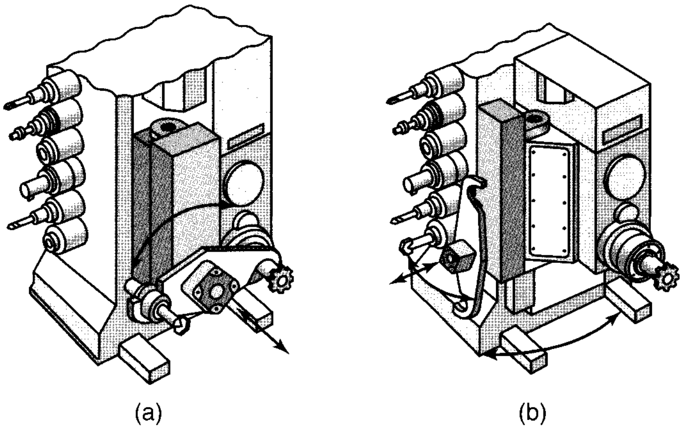
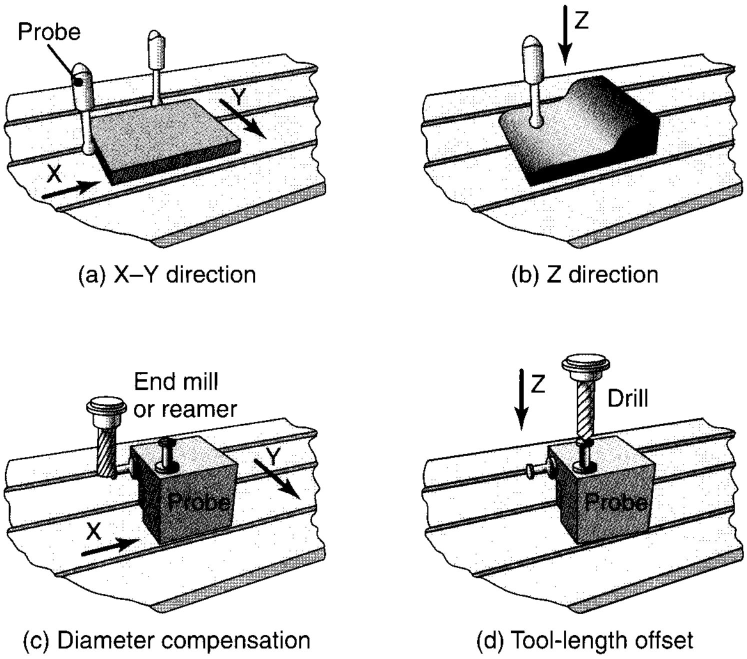

(a) (b)

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 movement 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.



(a) X–Y direction (b) Z direction

(c) Diameter compensation (d) Tool-length offset

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.

selection of tool settings and for the 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 are determined and stored in the database of the computer software. The data are then used to program tool paths and to compensate for tool length, tool diameter, and for tool wear in more advanced machine tools.

25.2.1 Types of Machining Centers

There are various designs for machining centers. The two basic types are vertical spindle and horizontal spindle, but many machines are capable of operating along both axes.

Vertical-spindle Machining Centers. Also called *vertical machining centers* (VMC), these are capable of performing various machining operations on parts with deep cavities, as in mold and die making. 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 figure, and all operations and movements are directed and modified through the computer control panel shown on the right. Because the thrust forces in vertical machining are directed downward, such machines have high stiffness and produce parts with good dimensional accuracy. These machines generally are less expensive than horizontal-spindle machines.

Horizontal-spindle Machining Centers. Also called *horizontal machining centers* (HMC), these are suitable for large as well as tall workpieces that require machining on a number of their surfaces. The pallet can be swiveled on different axes

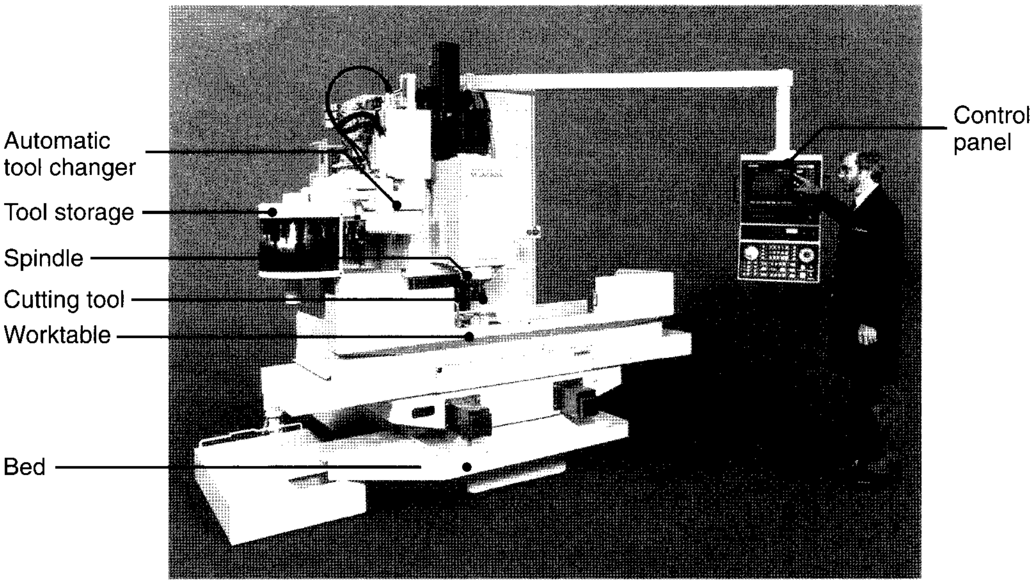


FIGURE 25.7 A vertical-spindle machining center. The tool magazine is on the left of the machine. The control panel on the right can be swiveled by the operator. *Source*: Courtesy of Cincinnati Milacron, Inc.

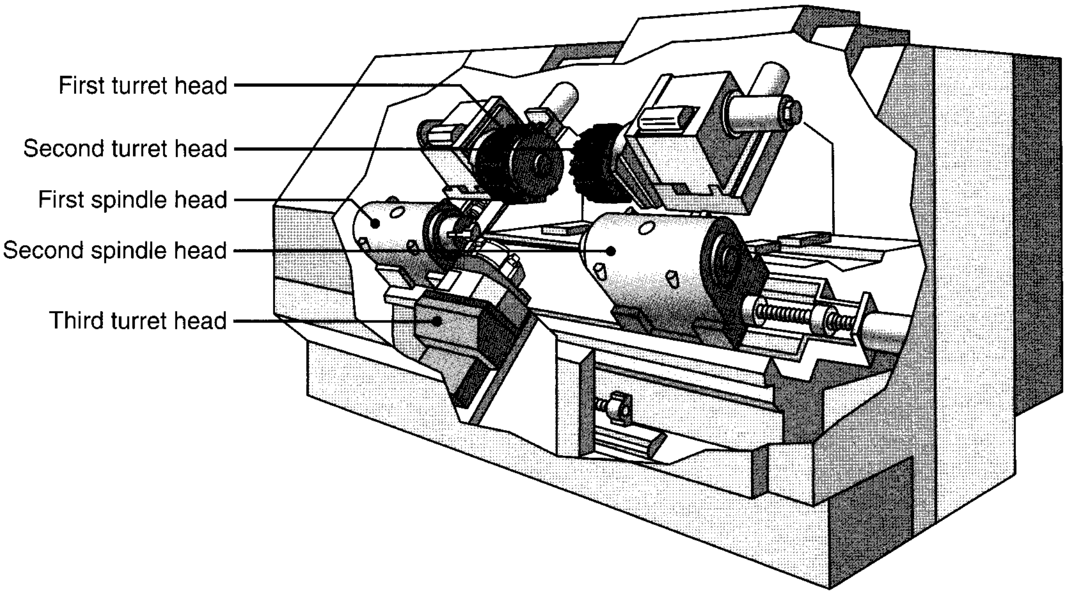


FIGURE 25.8 Schematic illustration of a computer numerical-controlled turning center. The two spindle heads and three turret heads make the machine very flexible in its machining capabilities. *Source*: Courtesy of Hitachi Seiki Co., Ltd.

(for example, see Fig. 25.3) to various angular positions. Another category of horizontal-spindle machines is **turning centers**, which are computer-controlled *lathes* with several features. A three-turret turning center is shown in Fig. 25.8. It is constructed with two horizontal spindles and three turrets equipped with a variety of cutting tools used to perform several operations on a rotating workpiece.

Universal machining centers are equipped with both vertical and horizontal spindles. They have a variety of features and are capable of machining all of the surfaces of a workpiece (i.e., vertical and horizontal and at a wide range of angles).

25.2.2 Characteristics and Capabilities of Machining Centers

The major characteristics of machining centers are summarized as follows:

- Machining centers are capable of handling a wide variety of part sizes and shapes efficiently, economically, repetitively, and with high dimensional accuracy—with tolerances on the order of ±0.0025 mm.
- These machines are versatile and capable of quick changeover from one type of product to another.
- The time required for loading and unloading workpieces, changing tools, gaging of the part, and troubleshooting is reduced. Therefore productivity is improved, thus reducing labor requirements (particularly skilled labor) and minimizing production costs.
- These machines are equipped with tool-condition monitoring devices for the detection of tool breakage and wear, as well as probes for tool-wear compensation and for tool positioning.

- In-process and postprocess gaging and inspection of machined workpieces are now features of machining centers.
- These machines are relatively compact and highly automated and have advanced control systems, so 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 (Section 23.3.7); this is referred to as *chip management.* Several designs are available for **chip collection,** with one or more spiral (screw) or chain conveyors that collect the chips along troughs in the machine and deliver them to a collecting point.

Machining centers are available in a wide variety of sizes and features, and their costs range from about \$50,000 to \$1 million and higher. Typical capacities range up to 75 kW. Maximum spindle speeds are usually in the range from 4,000 to 8,000 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 only 1.5 seconds. Some pallets are capable of supporting workpieces weighing as much as 7,000 kg, although even higher capacities are available for special applications.

25.2.3 Selection of Machining Centers

Machining centers generally require significant capital expenditure, so, to be cost effective, they may have to be used 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 *just-in-time manufacturing,* as will be described in Section 39.5.

The selection of the type and size of machining centers depends on several factors, especially the following:

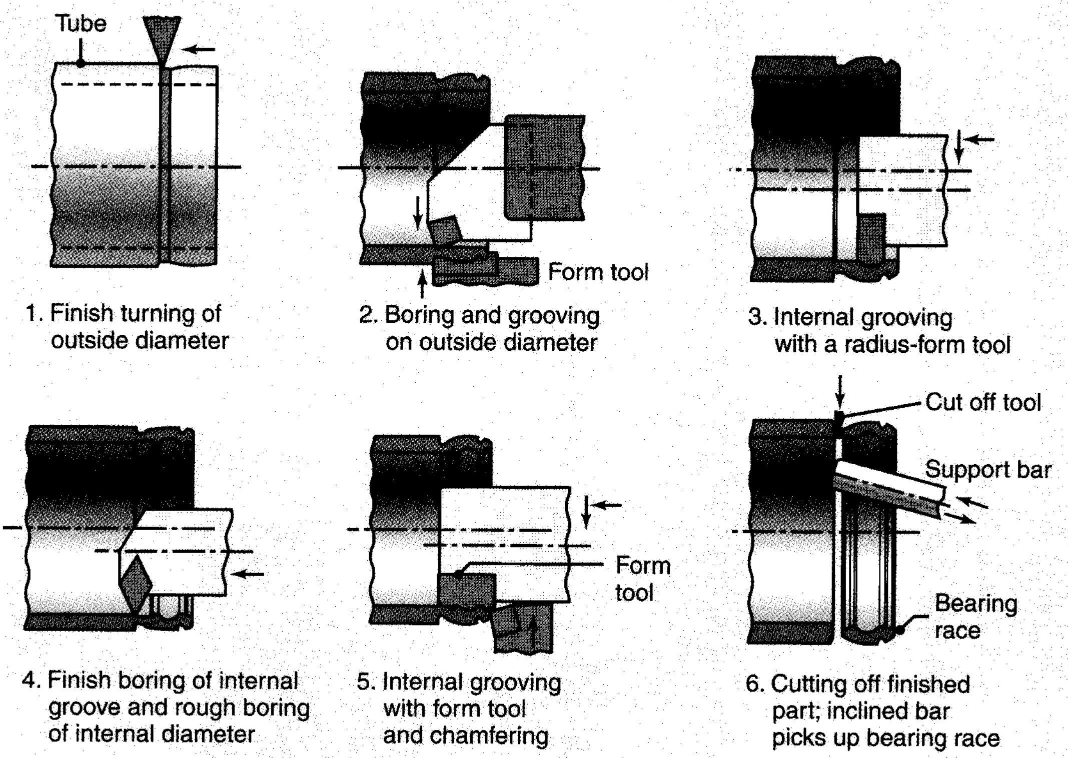
- Type of products, their size, and their shape complexity
- Type of machining operations to be performed and the type and number of cutting tools required
- Dimensional accuracy required
- Production rate required.

EXAMPLE 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), which is 3.18 mm, instead of 4.76 mm for the high-speed steel cutoff tool that formerly was used.

The material saved by this change is significant because the race width is small. The turning center was able to machine these races at high speeds and with repeatable tolerances of ±0.025 mm. (See also Example 23.2.)

Source: Courtesy of McGill Manufacturing Company.



1. Finish turning of outside diameter
2. Boring and grooving on outside diameter
3. Internal grooving with a radius-form tool
4. Finish boring of internal groove and rough boring of internal diameter
5. Internal grooving with form tool and chamfering
6. Cutting off finished part; inclined bar picks up bearing race

FIGURE 25.9 Machining of outer bearing races.

25.2.4 Reconfigurable Machines and Systems

The need for the flexibility of manufacturing processes has led to the more recent concept of *reconfigurable machines,* consisting of various modules. The term *reconfigurable* stems from the fact that (using advanced computer hardware and reconfigurable controllers and utilizing advances in information management technologies) the machine components can be arranged and rearranged quickly in a number of configurations to meet specific production demands. Figure 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 machining 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 movements) machine can be reconfigured by assembling different modules.

Reconfigurable machines have the promise 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). These capabilities are significant, especially in view of the frequent introduction of new products into a highly competitive global marketplace, fluctuations in product demand and product mix, and unpredictable modifications in product design.

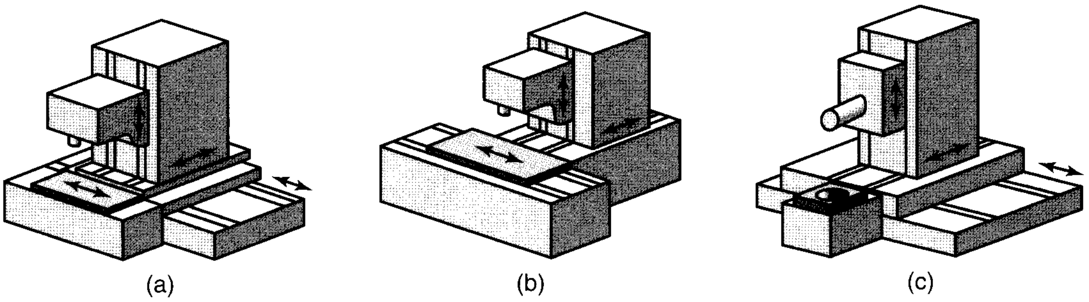


FIGURE 25.10 Schematic illustration of a reconfigurable modular machining center capable of accommodating workpieces of different shapes and sizes and requiring different machining operations on their various surfaces. *Source*: After Y. Koren.

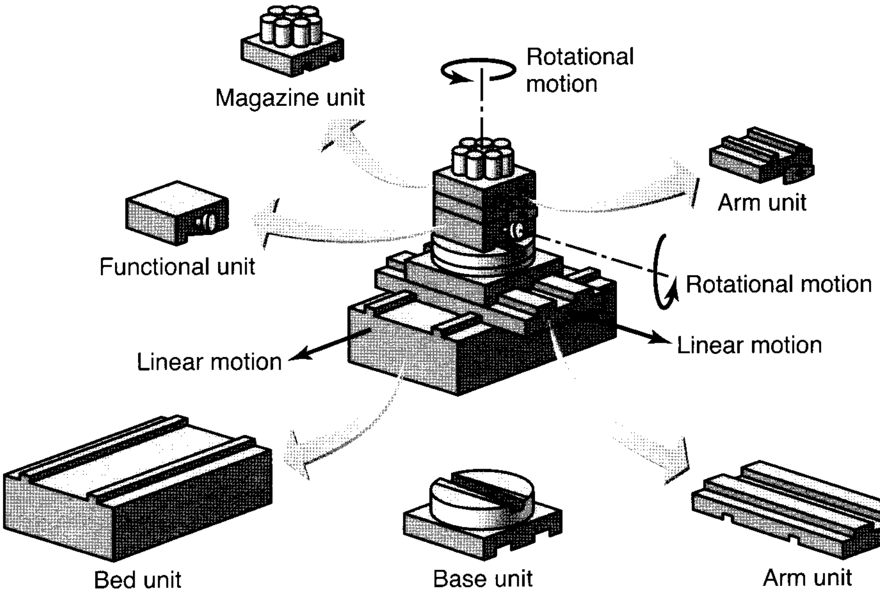


FIGURE 25.11 Schematic illustration of the assembly of different components of a reconfigurable machining center. *Source*: After Y. Koren.

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.

25.3.1 Materials

The following is a list of the materials that have been used commonly or are suitable for machine-tool structures, together with a description of their characteristics.

- **Gray cast iron**, the first material used in machine tool structures, has the advantages of a good damping capacity and low cost, but the disadvantage of being heavy. Most machine-tool structures are made of class 40 cast iron;

some are made of class 50 (see Table 12.4). Each casting requires a pattern, the cost of which increases significantly with the size of the part.

- **Welded steel** structures are lighter than cast-iron structures. Wrought steels (a) are available in a wide range of section sizes and shapes (such as channels, angles, and tubes), (b) have desirable mechanical properties, (c) possess good manufacturing characteristics (such as formability, machinability, and weldability), and (d) have low cost. Structures made of steels can have high stiffness-to-weight ratios, using cross sections such as tubes and channels. In contrast, their damping capacity is very low.
- **Ceramic** components are used in advanced machine tools for their strength, stiffness, corrosion resistance, surface finish, and good thermal stability. Ceramic machine-tool components were first introduced in the 1980s. Spindles and bearings now can be made of silicon nitride, which has better friction and wear characteristics than traditional metallic materials. Furthermore, the low density of ceramics makes them suitable as the components of high-speed machinery that undergo rapid reciprocating or rotating movements in which 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** may consist of a polymer matrix, metal matrix, or ceramic matrix with various reinforcing materials. The compositions can be tailored to provide appropriate mechanical properties in selected axes of the machine tool. Although they are expensive and (presently) limited in use, composites are likely to become significant 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 precision centerless and internal grinders in the early 1980s. These composite materials have several favorable properties: (a) good castability (which allows design versatility in machine tools), (b) high stiffness-to-weight ratio, (c) thermal stability, (d) resistance to environmental degradation, and (e) good damping capacity.
- **Polymer concrete** is a mixture of crushed concrete and plastic (typically polymethylmethacrylate) and can be cast easily into desired shapes for machine bases and various components. It was first introduced in the 1980s. Although it has good damping capacity, polymer concrete has low stiffness (about one-third that of class 40 cast iron) and poor thermal conductivity. It also can be used for sandwich construction 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-tool bases with loose sand also has been shown to 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 construction.
- Thermal distortion of machine components.
- Error compensation and the control of moving components along slideways.

Stiffness. Stiffness is a major factor in the dimensional accuracy of a machine tool. It 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 stiffness can be enhanced by design improvements, such 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 (e.g., welded versus bolted) in the machine-tool structure. As described previously, cast irons and polymer-matrix composites have much better damping capacity than metals or ceramics. Also, the greater the number of joints in a structure, the more damping there is.

Thermal Distortion. An important factor in machine tools is the thermal distortion of their components, which contributes significantly to their lack of precision. There are two sources of heat in machine tools:

1. *Internal sources*—such as from bearings, ballscrews, machine ways, spindle motors, pumps, and servomotors, as well as from the cutting zone during machining.
2. *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.

These considerations are important, particularly in **precision** and **ultraprecision machining**, where dimensional tolerances and surface finish are now approaching the nanometer range. The machines used for these applications are equipped with the following features:

- Various thermal and geometric real-time error-compensating features, including (1) the modeling of heating and cooling and (2) electronic compensation for accurate ballscrew positions.
- Gas or fluid hydrostatic spindle bearings; these low-friction bearings allow tools to more easily achieve precise motions without encountering stick-slip or high friction.
- New designs for traction or friction drives for linear motion.
- Extremely fine feed and position controls using microactuators.
- Fluid-circulation channels in the machine-tool base for the maintainance of thermal stability.

The structural components of the machine tool can be made of materials with high dimensional stability and a low coefficient of thermal expansion, such as Super-Invar, 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 welding. 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 by a hybrid casting technology. *Resin bonding* is being used to assemble machine tools, replacing mechanical fastening. Adhesives (described in 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 components.

Guideways. The preparation of guideways in machine tools traditionally has required significant effort. The plain cast-iron ways in machines (which is the most common) require much care to achieve the required precision and service life. The movements of various components in a machine tool along various axes usually have utilized conventional *ballscrews, rotating-screw drives,* and *rotary motors.* This system of mechanical and electrical components has several disadvantages, such as speed limitations, length restrictions, inertia effects, gear backlash and other errors, wear of the components, and low efficiency.

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. The sliding surfaces in these drives are separated by an air gap and, as a result, have very low friction.

Linear motor drives in machine tools have the following important advantages:

- Simplicity and minimal maintenance—since there is one moving part and no mechanical linkages.
- Smooth operation, better positioning accuracy, and repeatability—at as low as submicron ranges.
- 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.
- The moving components encounter no wear, because there is no physical contact between the sliding surfaces of the machine.

Machine Foundations. Foundation materials, their mass, and the manner in which they are installed in a plant are important considerations, as they help reduce vibration and do not adversely affect the performance of nearby machinery. 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. The large mass of concrete, combined with the machine base, reduced the amplitude of vibrations. Even better results can be obtained when the machine is installed on an *independent* concrete slab that is isolated from the rest of the plant floor with shock-isolation devices.

25.3.3 Hexapod Machines

Developments in the design and materials used for machine-tool structures and components are taking place continually. The important goals are (a) imparting machining flexibility to machine tools, (b) increasing their *machining envelope* (i.e., the space within which machining can be done), and (c) making them lighter. A truly innovative 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 based on a mechanism called the *Stewart platform* (after D. Stewart)—an invention developed in 1966 and first used to position aircraft cockpit simulators. The main advantage is that the links in the hexapod are loaded axially; the bending stresses and deflections are minimal, resulting in an extremely stiff structure.

The workpiece is mounted on a stationary table. Three pairs of *telescoping tubes* (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.

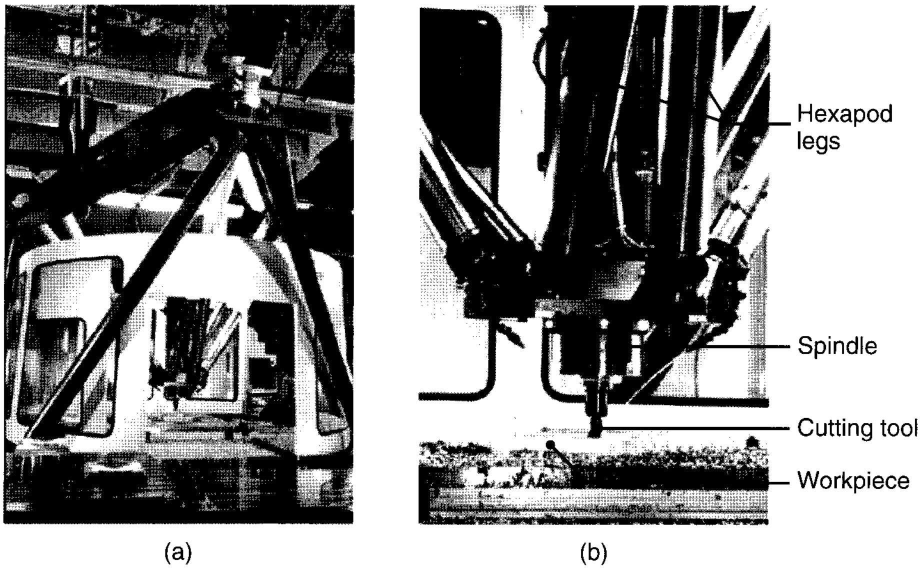


FIGURE 25.12 (a) A hexapod machine tool, showing its major components. (b) A detailed view of the cutting tool in a hexapod machining center. *Source*: Courtesy of National Institute of Standards and Technology.

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 with high inertial forces are involved.

These 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 (hence greater access to the work zone), (e) are capable of maintaining the cutting tool perpendicular to the surface being machined (thus improving the machining operation), and (f) have high flexibility (with six degrees of freedom) in the production of parts with various geometries and sizes without the need to refixture the work in progress. Unlike most machine tools, these are basically portable. In fact, *hexapod attachments* are now available so that 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. The cost of a hexapod machine is currently about $500,000, but is likely to come down significantly as they become more accepted and used.

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 as important as any other parameter in machining. Low stiffness can cause *vibration and chatter* of the cutting tools and the machine components, and thus have adverse effects on product quality. Uncontrolled vibration and chatter can result in the following problems:

- 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.

Vibration and chatter in machining are complex phenomena. There are two basic types of vibration in machining: forced vibration and self-excited vibration.

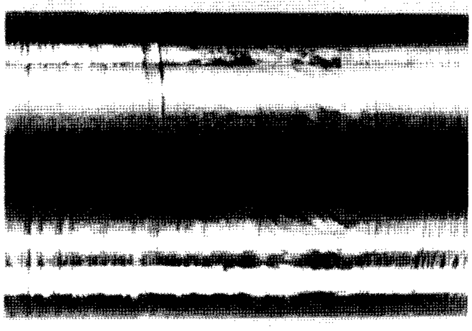


FIGURE 25.13 Chatter marks (right of center of photograph) on the surface of a turned part. *Source*: Courtesy of General Electric Company.

Forced Vibration. Forced vibration generally is caused by some *periodic* applied force present in the machine tool, such as that from gear drives, imbalance of the machine-tool components, misalignment, and motors and pumps. In processes such as the milling or turning of a splined shaft or a shaft with a keyway or hole, forced vibrations are caused by the periodic engagement of the cutting tool with the workpiece surface. (For example, see Figs. 24.9 and 24.14.)

The basic solution to forced vibration is to isolate or remove the forcing element. For example, if the forcing frequency is at or near the natural frequency of a machine-tool system 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. Although changing the cutting parameters generally does not appear to greatly influence the magnitude of forced vibrations, changing the cutting speed and the tool geometry can be helpful. Also, it is recognized that the source of vibrations 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, reduce deflections, and thus improve accuracy.

Self-excited Vibration. Generally called **chatter**, self-excited vibration is caused by the interaction of the chip-removal process with the structure of the machine tool. Self-excited vibrations usually have a very high amplitude. Chatter typically begins with a disturbance in the cutting zone; among its possible causes are (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 (e.g., see Figs. 21.2 and 21.21). Consequently, the depth of cut varies, 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 be observed easily while driving a car over a rough road (the so-called *washboard effect*).

Self-excited vibrations generally can be controlled by

- Increasing the **stiffness** and, especially, the **dynamic stiffness** of the system.
- **Damping.**

Dynamic stiffness is defined as the ratio of the applied-force amplitude to the vibration amplitude. For example, recall that in the trepanning operation (Fig. 23.23b) there are four components involved in the deflections and hence causing the vibration: the spindle, the supporting arm for the cutting tool, the drill, and the cutting tool. Experience and analysis would indicate that, unless all of the components are sufficiently stiff, such an operation likely will lead to chatter, beginning with torsional vibration around the spindle axis and twisting of the arm. Two examples are long and slender drills that may undergo torsional vibrations and 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 cutting forces increase with strength (hence with hardness of the workpiece material), the tendency to chatter generally increases as hardness increases. Thus, aluminum and magnesium alloys 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 cutting operations. Continuous chips involve steady cutting forces. Consequently, such chips generally do not cause chatter. Discontinuous chips and serrated chips (Fig. 21.5), however, may do so. These types of chips are periodic, and the resulting force variations during cutting can cause chatter. Other factors that may contribute to chatter are the use of dull cutters, a 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 soon the vertical motion stops. Damping is a major factor in controlling machine-tool vibration and chatter, and consists of internal and external damping.

Internal damping results from the energy loss in materials during vibration. For example, steels have a lower damping capacity, and composite materials have a higher damping capacity, than gray cast iron (Fig. 25.14). The difference in the damping capacity of materials can be observed by striking them with a gavel and listening to the sound. For example, try striking a brass cymbal, a piece of concrete, and a piece of wood, and listen to the variations in sound.

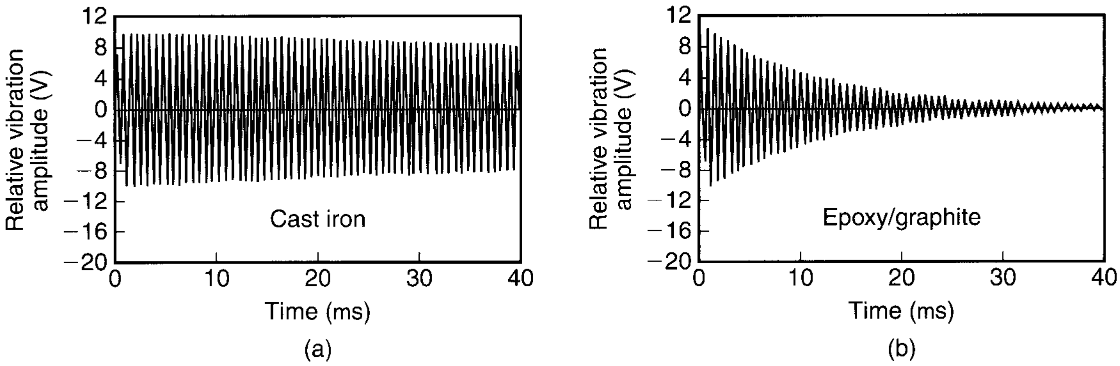


FIGURE 25.14 The relative damping capacity of (a) gray cast iron and (b) an epoxy–granite composite material. The vertical scale is the amplitude of vibration and the horizontal scale is time.

Bolted joints in the structure of a machine tool also are a source of damping; their effectiveness depends on size, position, and the number of joints. Because friction dissipates energy, small relative movements along dry (unlubricated) joints do so and thus improve damping. Because all 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 how overall damping increases as the number of components on a lathe and their contact areas increase. However, as expected, the overall system stiffness of the machine generally is reduced as the number of joints increases.

Another method of damping in order to reduce or eliminate vibration and chatter is implemented in some boring bars and toolholders. As described and illustrated in Fig. 23.17b, damping can be accomplished by mechanical means—dissipating energy by the frictional resistance of the components within the structure of the boring bar.

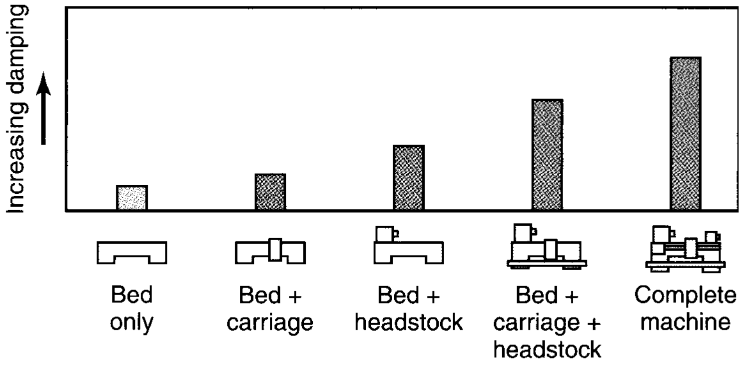


FIGURE 25.15 The damping of vibrations as a function of the number of components on a lathe. Joints dissipate energy; the greater the number of joints, the higher is the damping capacity of the machine. (See also Fig. 23.2.) *Source*: After J. Peters.

External damping is accomplished with external dampers that are similar to shock absorbers on automobiles or machines. Special vibration absorbers have been developed and installed on machine tools for this purpose. Furthermore, machine tools can be installed on specially prepared floors and foundations to isolate forced vibrations, such as those from nearby machines on the same floor.

Guidelines for Reducing Vibration and Chatter. It is evident from the foregoing discussion that 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. In various sections of Chapters 23 and 24, we outlined general guidelines for reducing vibration and chatter in machining operations. These basic guidelines may be summarized as follows:

- Minimize tool overhang.
- Improve the stiffness of work-holding devices and support workpieces rigidly.
- Modify tool and cutter geometry to minimize forces or make them uniform.
- Change process parameters, such as cutting speed, feed, depth of cut, and cutting fluids.
- Increase the stiffness of the machine tool and its components by improving design and using larger cross sections and materials with a higher elastic modulus.
- Improve the damping capacity of the machine tool.

25.5 High-speed Machining

With continuing demands for higher productivity and lower production costs, investigations have been carried out since the late 1950s to increase the cutting speed and the material-removal rate in machining, particularly with applications in the aerospace and automotive industries.

The term "high" in *high-speed machining* (HSM) is somewhat relative; as a general guide, an approximate range of cutting speeds may be defined as follows:

- High speed: 600 to 1,800 m/min
- Very high speed: 1,800 to 18,000 m/min
- Ultrahigh speed: >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 should a failure occur. The *spindle power* required in high-speed machining is generally on the order of 0.004 W/rpm —much less than in traditional machining, which is typically in the range from 0.2 to 0.4 W/rpm. Feed rates in high-speed machining are now up to 1 m/s, and the acceleration rates of machine-tool components are very high.

Spindle designs for high speeds require *high stiffness* and *accuracy* and generally involve 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 elements or hydrostatic; the latter is more desirable because it requires less space than the former. Because of *inertia effects* during the acceleration and decelaration of machine-tool components, the use of lightweight materials (including ceramics and composite materials) is an important consideration.

The selection of appropriate cutting-tool materials is, of course, a major consideration. On the basis of our 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 this operation.

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 time in the overall machining operation. As will be described in Section 38.6 and Chapter 40, **noncutting time** and various other factors are important considerations in the overall assessment of the benefits of high-speed machining for a particular application.

Research and development work continues on high-speed machining, especially in the turning, milling, boring, and drilling of aluminum alloys, titanium alloys, steels, and superalloys. Much data have been collected regarding the effect of high cutting speed on (a) the type of chips produced, (b) cutting forces and power, (c) temperatures generated, (d) tool wear and failure, (e) surface finish and integrity, and (f) the economics of the process.

These studies have indicated that high-speed machining is economical for certain specific applications. For example, 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, and weighing 55,000 kg; 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 regarding *machining centers* 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 in cutting operations. As we have seen in Fig. 21.14, as the cutting speed increases, more and more of the heat generated is removed by the chip; thus, the tool and (more importantly) the workpiece remain close to ambient temperature. This is beneficial, because there is no thermal expansion or warping of the workpiece during machining.

The important machine-tool characteristics in high-speed machining may be summarized as follows:

1. Spindle design for stiffness, accuracy, and balance at very high rotational speeds.
2. Bearing characteristics.
3. Inertia of the machine-tool components.
4. Fast feed drives.
5. Selection of appropriate cutting tools.
6. Processing parameters and their computer control.
7. Work-holding devices that can withstand high centrifugal forces.
8. Chip-removal systems that are effective at very high rates of material removal.

25.6 Hard Machining

We have 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. As will be described in Chapters 26 and 27, there are several other mechanical (including abrasive machining), nonmechanical, and advanced techniques of removing material economically from hard or hardened metals and alloys. However, it is still possible to apply traditional machining processes to hard metals and alloys by selecting an appropriate hard-tool material and using machine tools with high stiffness, power, and precision.

A common 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,* this process produces machined parts with good dimensional accuracy, surface finish (as low as 0.25 μm), and surface integrity. The (a) available power, (b) static and dynamic stiffness of the machine tool and its spindle, and (c) work-holding devices and fixturing are important factors.

As we described in Section 25.3, trends in the design and construction of modern machine tools include the use of hydrostatic bearings for the spindles and slideways. The headstock and the slanted bed in the machines (see Fig. 23.10a) can be made of *granite composite materials* with unique properties. Cutting-tool selection and edge preparation also are important to avoid premature failure in hard machining.

From technical, economic, and ecological considerations, hard turning can compete successfully with the *grinding* process, and the quality and characteristics of products produced by hard turning can compete with the quality and characteristics approaching those produced in grinding machines. For instance, in some specific cases, hard turning has been shown to be three times faster than grinding, requires fewer operations to finish the part, and utilizes five times less energy. A detailed comparative example of hard turning versus grinding will be presented in Example 26.4.

25.7 Ultraprecision Machining

Beginning in the 1960s, increasing demands were being made for the precision manufacturing of components for computer, electronic, nuclear, and defense applications. Some examples include optical mirrors, computer memory disks, and drums

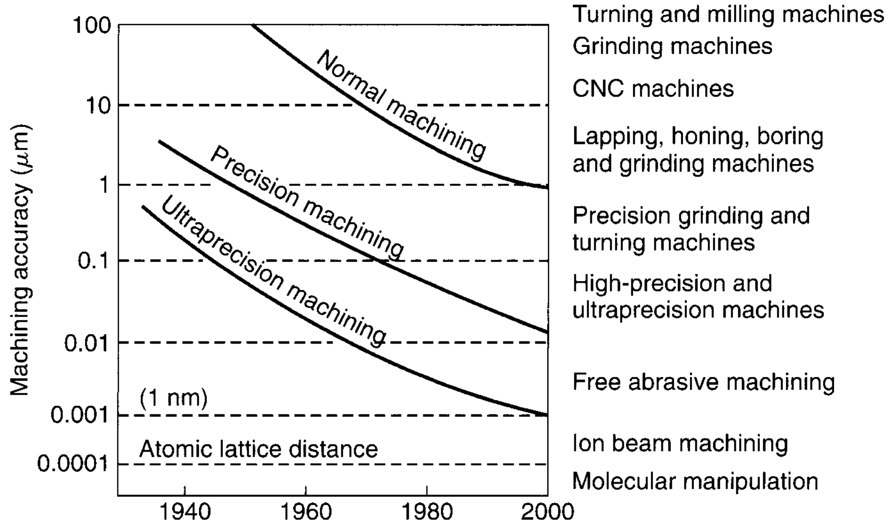


FIGURE 25.16 Improvements in machining accuracy over the years, using ultraprecision machining technologies. *Source*: After C.J. McKeown, N. Taniguchi, G. Byrne, D. Dornfeld, and B. Denkena.

for photocopying machines. Surface-finish requirements are in the tens-of-*nanometer* (10^{-9} m, or 0.001 μm) range, and dimensional tolerances and form accuracies are in the micrometer (μm) and submicrometer range.

This trend toward ultraprecision manufacturing continues to grow. Modern **ultraprecision machine tools** with advanced computer controls can position a cutting tool within an accuracy approaching 1 nm, as can be seen from Fig. 25.16. Also, note in this figure that higher precision is now being achieved by processes such as abrasive machining, ion-beam machining, and molecular manipulation. We will describe *abrasive machining* and *advanced machining processes* in Chapters 26 and 27, respectively. We will cover the topics of *micromachining*, the *fabrication of microelectronic* and *microelectromechanical systems* (MEMS), and *nanomanufacturing* in Chapters 28 and 29.

The cutting tool for *ultraprecision machining* applications is almost exclusively a single-crystal diamond, and the process is called **diamond turning**. The diamond tool has a polished cutting edge with a radius as small as a few nm. Wear of the diamond can be a significant problem, and recent advances include *cryogenic diamond turning*, in which the tooling system is cooled by liquid nitrogen to a temperature of about −120°C.

The workpiece materials for ultraprecision machining include copper alloys, 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 known as **ductile-regime cutting**. (See also *ductile-regime grinding* in Section 26.3.4.) Deeper cuts in brittle materials produce discontinuous chips.

The machine tools for these applications are built with very high precision and high machine, spindle, and work-holding-device stiffnesses. These ultraprecision machines (parts of which are made of structural materials with low thermal expansion and good dimensional stability; see Section 25.3) 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 machines on the same floor—must be avoided as much as possible. Laser metrology is used for feed and position control, 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 machine tools, somewhat similar to those in high-speed machining:

1. Machine-tool design, construction, and assembly—providing stiffness, damping, and geometric accuracy.
2. Motion control of various components—both linear and rotational.
3. Spindle technology.
4. Thermal growth of the machine tool, compensation for thermal growth, and control of the machine-tool environment.
5. Cutting-tool selection and application.
6. Machining parameters.
7. Real-time performance and control of the machine tool and implementation of a tool-condition monitoring system.

25.8 Machining Economics

In the Introduction to Part IV, it was stated that the important limitations of machining operations include (a) the relatively longer time required to machine a part compared with forming and shaping it, (b) the need to reduce noncutting time, and (c) the fact that material is inevitably wasted. Despite these drawbacks, machining is indispensable, particularly for producing complex workpiece shapes with external and internal features and for obtaining high dimensional accuracy and surface finish.

We have outlined the material and process parameters that are relevant to efficient machining operations. In analyzing the *economics* of machining, however, several other factors have to be considered. As will be described in greater detail in Chapter 40, these factors include the costs involved in (a) machine tools, work-holding devices and fixtures, and cutting tools; (b) labor and overhead; (c) the time consumed 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 surface finish; and (f) cutting times and noncutting times.

Actual machining time is an important consideration. Recall also the discussion in Section 25.5 on the importance of noncutting time in assessing the economic relevance of high-speed machining. Unless noncutting time is a significant portion of the floor-to-floor time, high-speed machining should not be considered—except if its other benefits are relevant.

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. Various methods and approaches have been developed to accomplish this goal. With the increasing use of software and user-friendly computers, this task now has become easier. However, in order for the results to be reliable, it is essential that input data be accurate and up to date. Described next is one of the simpler and more common 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 cutting operation. Consequently, using the least expensive tool is not always an effective way of 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, indirect labor, and the like. 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 so on. The **tooling cost** is

$$C_t = \frac{1}{N_i}[T_c(L_m + B_m) + D_i] + \frac{1}{N_f}[T_i(L_m + B_m)], \tag{25.4}$$

where N_i is the number of parts machined per insert, N_f is the number of parts that can be produced per insert face, 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 produce 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. For example, let's consider a turning operation; the machining time (see Section 23.2) is

$$T_m = \frac{L}{fN} = \frac{\pi LD}{fV}, \tag{25.6}$$

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 the Taylor tool-life equation [Eq. (21.20b)], we have

$$T = \left(\frac{C}{V}\right)^{1/n}, \tag{25.7}$$

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. The number of pieces per insert face is thus simply

$$N_f = \frac{T}{T_m}, \tag{25.8}$$

and the number of pieces per insert is given by

$$N_i = mN_f = \frac{mT}{T_m}. \tag{25.9}$$

Sometimes not all of the faces are used before the insert is discarded, so it should be recognized that m corresponds to the number of faces that are actually used, not the number of faces provided per insert. The combination of Eqs. (25.6)–(25.8) gives

$$N_i = \frac{mfC^{1/n}}{\pi LDV^{(1/n)-1}}. \tag{25.10}$$

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**, we differentiate C_p with respect to V and set it to zero. Thus,

$$\frac{\partial C_p}{\partial V} = 0. \tag{25.11}$$

Thus, the optimum cutting speed, V_o, is

$$V_o = \frac{C(L_m + B_m)^n}{\left(\frac{1}{n} - 1\right)^n \left\{\frac{1}{m}[T_c(L_m + B_m) + D_i] + T_i(L_m + B_m)\right\}^n}, \tag{25.12}$$

and the optimum tool life, T_o, is

$$T_o = \left(\frac{1}{n} - 1\right)\frac{\frac{1}{m}[T_c(L_m + B_m) + D_i] + T_i(L_m + B_m)}{L_m + B_m}. \tag{25.13}$$

To find the optimum cutting speed and the optimum tool life for **maximum production**, we differentiate T_p with respect to V and set the result to zero. Thus, we have

$$\frac{\partial T_p}{\partial V} = 0. \tag{25.14}$$

The **optimum cutting speed** now becomes

$$V_o = \frac{C}{\left[\left(\frac{1}{n} - 1\right)\left(\frac{T_c}{m} + T_i\right)\right]^n}, \tag{25.15}$$

and the **optimum tool life** is

$$T_o = \left(\frac{1}{n} - 1\right)\left(\frac{T_c}{m} + T_i\right). \tag{25.16}$$

A qualitative plot of minimum cost per piece and minimum time per piece (i.e., the maximum production rate) is given in Fig. 25.17. The cost of a machined surface also depends on the finish required: The machining cost increases rapidly with finer surface finish.

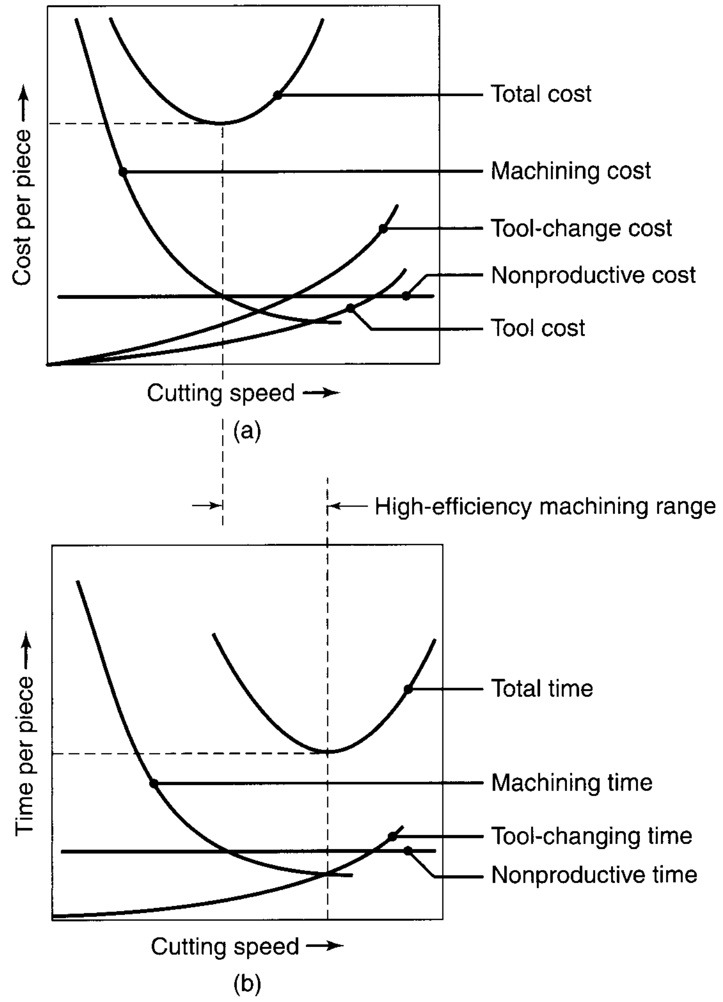


FIGURE 25.17 Graphs showing (a) cost per piece and (b) time per piece in machining. Note the optimum speeds for both cost and time. The range between the two is known as the high-efficiency machining range.

The preceding analysis indicates the importance of identifying all relevant parameters in a machining operation, determining various cost factors, obtaining relevant tool-life curves for the particular operation, and properly measuring the various time intervals involved in the overall operation. The importance of obtaining accurate data is shown clearly in Fig. 25.17, as small changes in cutting speed can have a significant effect on the minimum cost or time per piece.

Such an analysis is typically done for all manufacturing processes, 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 tolerances can be calculated from an equation similar to Eq. (25.1), but including costs associated with sand casting (the cost of mold production, pattern depreciation, etc.). A similar calculation can be performed on a processing approach that uses powder metallurgy (thus increasing die and machinery costs), but requires less machining because of its ability to achieve tighter tolerances, thereby reducing machining costs. A comparison of cost estimates can then help determine a processing strategy, as discussed in greater detail in Section 40.9.

SUMMARY

- Because they are versatile and capable of performing a variety of machining operations on small and large workpieces of various shapes, machining centers are now among the most important machine tools. Their selection depends on factors such 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 required.
- Vibration and chatter in machining are important considerations for workpiece dimensional accuracy, surface finish, and tool life. Stiffness and damping capacity of machine tools are important factors in controlling vibration and chatter. New materials are being developed continually for use in the construction of machine-tool structures.
- The economics of machining processes depends on factors such 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 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 construction
Pallet
Reconfigurable machines
Regenerative chatter
Self-excited vibration
Stiffness
Tool-exchange arm
Tool- and part-checking station
Touch probes
Turning center
Ultraprecision machining
Universal machining center
Work envelope

BIBLIOGRAPHY

Arnone, M., **High Performance Machining**, Hanser Gardner, 1998.

Boothroyd, G., and Knight, W.A., **Fundamentals of Machining and Machine Tools**, 3rd ed., Marcel Dekker, 2005.

Brown, J., **Advanced Machining Technology Handbook**, McGraw-Hill, 1998.

Erdel, B., **High-Speed Machining**, Society of Manufacturing Engineers, 2003.

Joshi, P.H., **Machine Tools Handbook**, McGraw-Hill, 2008.

Krar, S.F., and Check, A.F., **Technology of Machine Tools**, 6th ed., Glencoe/Macmillan McGraw-Hill, 2009.

Lopez, L.N., and Lamikiz, A. (eds.), **Machine Tools for High Performance Machining**, Springer, 2009.

Rivin, E.I., **Stiffness and Damping in Mechanical Design**, Marcel Dekker, 1999.

Youssef, H.A., and El-Hofy, H., **Machining Technology: Machine Tools and Operations**, CRC Press, 2008.

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. Explain the importance of foundations for installing machine tools.

25.6. What types of materials are machine-tool bases typically made from? Why?

25.7. What is meant by the "modular" construction of machine tools?

25.8. What factors contribute to costs in machining operations?

25.9. What is the high-efficiency machining range? Why is it so called?

QUALITATIVE PROBLEMS

25.10. Explain the technical and economic factors that led to the development of machining centers.

25.11. Spindle speeds in machining centers vary over a wide range. Explain why this is so, giving specific applications.

25.12. Explain the importance of stiffness and damping of machine tools. Describe how they are implemented.

25.13. Are there machining operations described in Chapters 23 and 24 that cannot be performed in machining and turning centers? Explain, with specific examples.

25.14. How important is the control of cutting-fluid temperature in operations performed in machining centers? Explain.

25.15. Review Fig. 25.10 on modular machining centers, and describe some workpieces and operations that would be suitable on such machines.

25.16. Describe the adverse effects of vibration and chatter in machining operations.

25.17. Describe some specific situations in which thermal distortion of machine-tool components would be important.

25.18. Explain the differences in the functions of a turret and of a spindle in turning centers.

25.19. Explain how the pallet arrangements shown in Figs. 25.4a and b would be operated in using these machines on a shop floor.

25.20. 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.

25.21. 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.

25.22. Describe types and sizes of workpieces that would not be suitable for machining on a machining center. Give specific examples.

25.23. 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. 25.17 important? Explain.

25.24. Explain the advantages and disadvantages of machine-tool frames made of gray-iron castings.

25.25. What are the advantages and disadvantages of (a) welded-steel frames, (b) bolted steel frames, and (c) adhesively bonded components of machine tools? Explain.

25.26. What would be the advantages and limitations of using concrete or polymer–concrete in machine tools?

25.27. Explain how you would go about reducing each of the cost factors in machining operations. What difficulties would you encounter in doing so?

QUANTITATIVE PROBLEMS

25.28. A machining-center spindle and tool extend 250 mm from their machine-tool frame. Calculate the temperature change that can be tolerated in order to maintain a tolerance of 0.0025 mm in machining. Assume that the spindle is made of steel.

25.29. 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.

SYNTHESIS, DESIGN, AND PROJECTS

25.30. 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.

25.31. Review the technical literature and outline the trends in the design of modern machine tools. Explain why there are those trends.

25.32. Make a list of components of machine tools that could be made of ceramics, and explain why ceramics would be suitable.

25.33. 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.

25.34. 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.

25.35. 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.

25.36. 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.

25.37. 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.

Abrasive Machining and Finishing Operations

CHAPTER 26

- Abrasive machining is among the final steps in the production of parts and is an important series of processes because of its capability to impart high dimensional accuracy and surface finish.
- A wide variety of abrasive finishing processes is available, many based on the fundamental mechanism of abrasion.
- This chapter opens with a description of the grinding process, the mechanics of chip removal, and the roles of abrasives and bonds in grinding wheels.
- Some abrasive machining operations, including polishing, buffing, honing, and sanding, require a bonded or coated abrasive; other processes, such as ultrasonic machining, lapping, abrasive flow machining, and electrochemical machining and grinding, use loose abrasives that are not bonded.
- All of these processes are described in detail, and their capabilities and design considerations are presented.
- The chapter ends 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.

EXAMPLES:

26.1 Introduction

There are many situations in manufacturing where the processes described thus far cannot produce the required dimensional accuracy or surface finish for a part, or the workpiece material is too hard or too brittle to process. Consider, for example, the accuracy and fine surface finish required on ball and roller bearings, pistons, valves, cylinders, cams, gears, dies, and numerous precision components used in instrumentation. One of the most common methods for producing such demanding characteristics on parts is abrasive machining.

An **abrasive** is a small, hard particle having sharp edges and an irregular shape, unlike the cutting tools described earlier. Abrasives are capable of removing small amounts of material from a surface through a cutting process that produces tiny

FIGURE 26.1 A variety of bonded abrasives used in abrasive-machining processes. *Source*: Courtesy of Norton Company.

chips. Most of us are familiar with using *grinding wheels* (bonded abrasives), as shown in Fig. 26.1, to sharpen knives and tools, as well as using *sandpaper* or *emery cloth* to smoothen surfaces and remove sharp corners. As described in this chapter, abrasives also are used to hone, lap, buff, and polish workpieces. With the use of computer-controlled machines, abrasive processes are now capable of producing (a) a wide variety of workpiece geometries, as can be seen in Fig. 26.2 and (b) very fine dimensional accuracy and surface finishes, as shown in Figs. 23.13, 23.14, and 27.4. For example, dimensional tolerances can be less than 1 μm, and surface roughnesses can be as fine as 0.025 μm.

Because they are hard, abrasives also are used in *finishing processes* for heat-treated metals and alloys and for very hard parts in applications such 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, and (d) cleaning surfaces with jets of air or water containing abrasive particles.

In this chapter, the characteristics of abrasives are presented along with their use in various material-removal processes. As with cutting operations, the mechanics of abrasive-machining operations are described first. This knowledge will then assist in establishing the interrelationships between workpiece material and process variables and the dimensional accuracy, surface finish, and surface integrity of the parts produced.

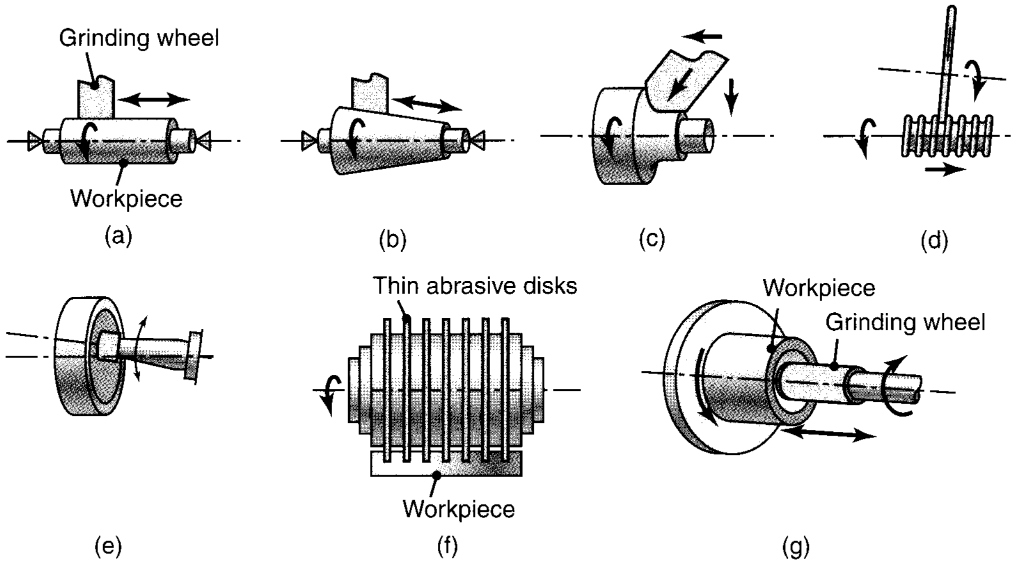


FIGURE 26.2 The types of workpieces and operations typical of grinding: (a) cylindrical surfaces, (b) conical surfaces, (c) fillets on a shaft, (d) helical profiles, (e) concave shape, (f) cutting off or slotting with thin wheels, and (g) internal grinding.

26.2 Abrasives and Bonded Abrasives

Abrasives that are used most commonly in abrasive-machining operations are as follows:

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. Because cubic boron nitride and diamond are the two hardest materials known, they are referred to as *superabrasives.*

In addition to hardness, an important characteristic of abrasives is **friability**—defined as the ability of abrasive grains to fracture (break down) into smaller pieces. This property gives abrasives their *self-sharpening* characteristics, which are 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. For example, aluminum oxide has lower friability than silicon carbide and, correspondingly, a lower tendency to fragment.

The *shape* and *size* of the abrasive grain also affect its friability. For example, blocky grains (which are analogous to a negative rake angle in single-point cutting tools, as shown in Fig. 21.3) are less friable than less blocky or plate-like grains. Furthermore, because the probability of defects diminishes as the grain size becomes smaller (due to the *size effect*), smaller grains are stronger and less friable than larger ones.

Abrasive Types. The abrasives commonly found in nature are *emery*, *corundum* (alumina), *quartz*, *garnet*, and *diamond*. Because these natural abrasives generally contain impurities and possess nonuniform properties, their performance is 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 as *dark* (less friable), *white* (very friable), and *single crystal.*
- **Seeded gel** was first introduced in 1987 and is the purest form of *unfused aluminum oxide*. It also is known as *ceramic aluminum oxide*. It has a grain size on the order of 0.2 μm, which is much smaller than other types of commonly used abrasive grains. These grains are sintered to form larger sizes. Because they

TABLE 26.1

Ranges of Knoop Hardness for Various Materials and Abrasives

Common glass	350–500	Titanium nitride	2000
Flint, quartz	800–1100	Titanium carbide	1800–3200
Zirconium oxide	1000	Silicon carbide	2100–3000
Hardened steels	700–1300	Boron carbide	2800
Tungsten carbide	1800–2400	Cubic boron nitride	4000–5000
Aluminum oxide	2000–3000	Diamond	7000–8000

are harder than fused alumina and have relatively high friability, seeded gels maintain their sharpness and are used especially for difficult-to-grind materials.
- **Silicon carbide** was first discovered in 1891 and is made with silica sand and petroleum coke. Silicon carbides are divided into *black* (less friable) and *green* (more friable) and generally have higher friability than aluminum oxides. Hence, they have a greater tendency to fracture and 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 that were described in Chapters 21 and 22. Also, abrasives have sharp edges, allowing the removal of very small quantities of material from the workpiece surface. Consequently, a very fine surface finish and dimensional accuracy can be obtained using abrasives as tools.

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 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 you can readily observe by noting the grit number printed on the back of the abrasive paper or cloth.

Abrasive-workpiece-material Compatibility. As in selecting cutting-tool materials for machining particular workpiece materials, the affinity of an abrasive grain to the workpiece material is an important consideration. The less the reactivity of the two materials, the less wear and dulling of the grains occur during grinding, either of which would make the operation less efficient and cause damage to the workpiece surface (see Section 26.3.1 for details). Recall that, because of its chemical affinity, diamond 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, cemented carbides, and some hardened steels.

26.2.1 Grinding Wheels

Because each abrasive grain typically removes only a very small amount of material at a time, high rates of material removal can be achieved only if a large number of these grains act together. This is 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. 26.3, 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 in order to provide clearance for the chips being produced and to provide cooling;

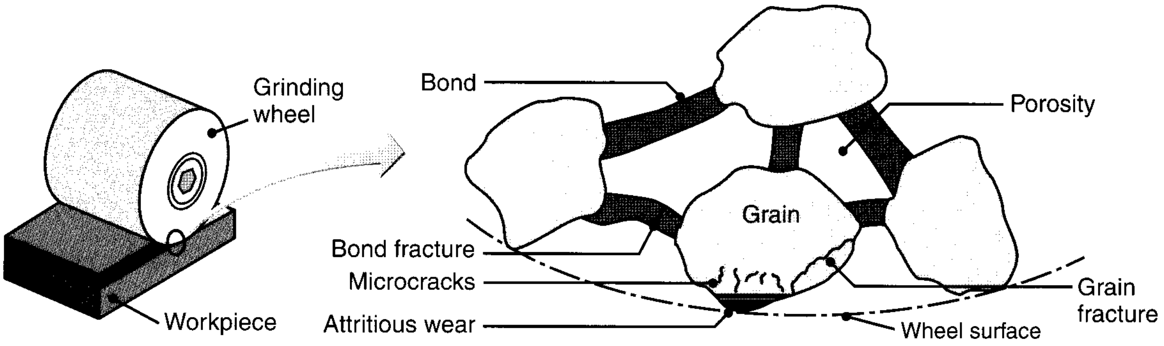
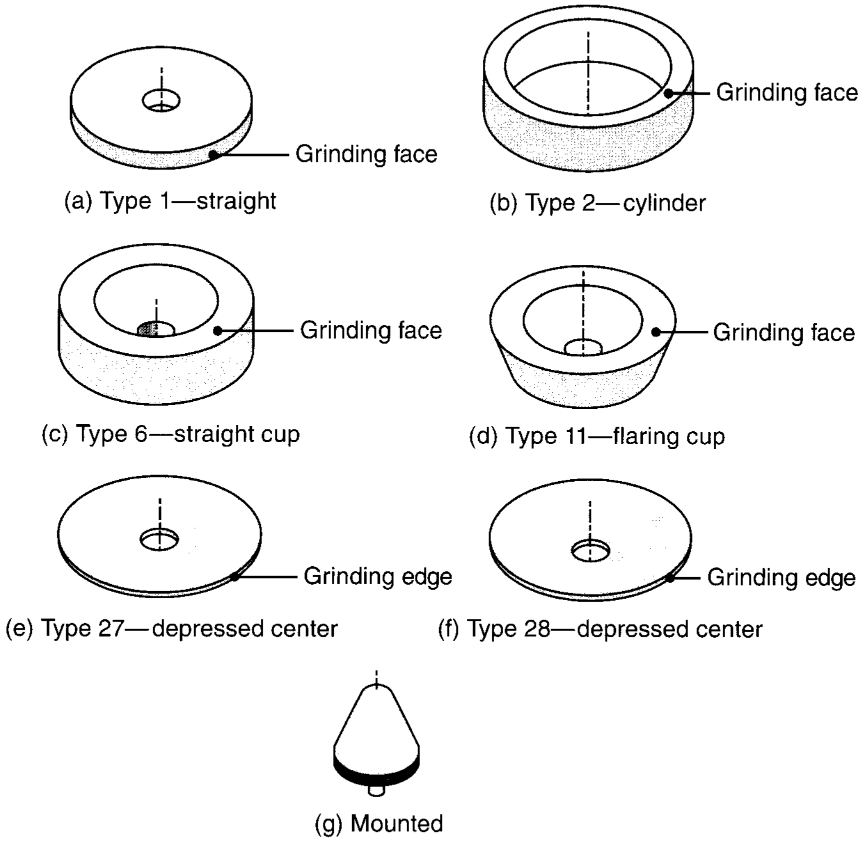


FIGURE 26.3 Schematic illustration of a physical model of a grinding wheel, showing its structure and its wear and fracture patterns.



(a) Type 1—straight
(b) Type 2—cylinder
(c) Type 6—straight cup
(d) Type 11—flaring cup
(e) Type 27—depressed center
(f) Type 28—depressed center
(g) Mounted

FIGURE 26.4 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.

otherwise, the chips would severely interfere with the grinding operation. Porosity can be observed by looking closely at the surface of any grinding wheel. As you can appreciate, it is impossible to utilize a grinding wheel with no porosity—one that is fully dense and solid; there simply is no room for any chips to form.

A very wide variety of types and sizes of abrasive wheels is made today. Some of the more commonly used types of grinding wheels for conventional abrasives are shown in Fig. 26.4. Superabrasive wheels are shown in Fig. 26.5.

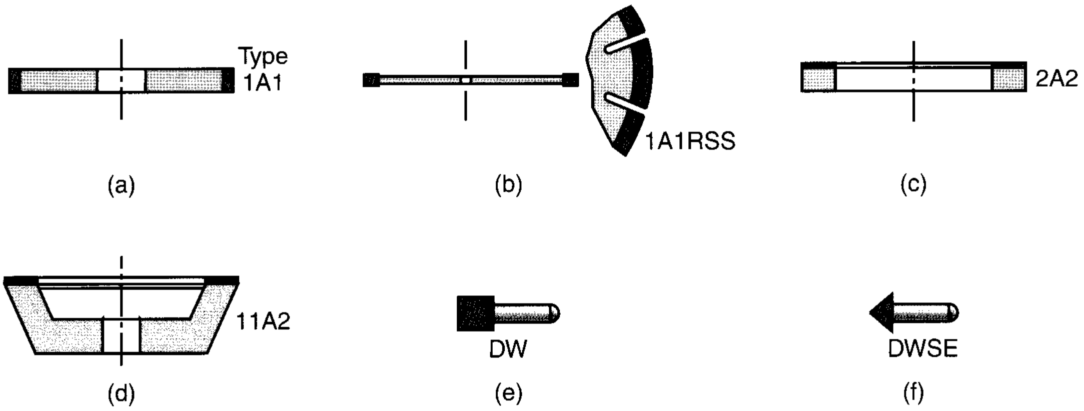


FIGURE 26.5 Examples of superabrasive wheel configurations. The annular regions (rims) are superabrasive grinding surfaces, and the wheel itself (core) generally is made of metal or composites. The bonding materials for the superabrasives are (a), (d), and (e) resinoid, metal, or vitrified; (b) metal; (c) vitrified; and (f) resinoid.

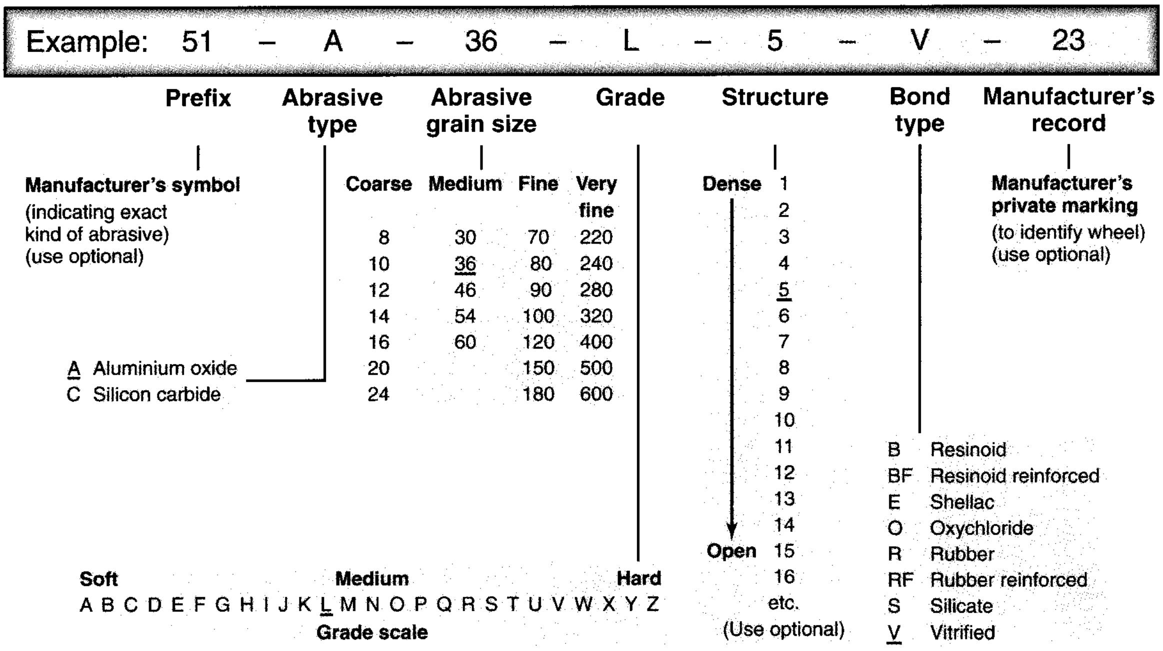


FIGURE 26.6 Standard marking system for aluminum-oxide and silicon-carbide bonded abrasives.

Note that, due to their high cost, only a small volume of these wheels consists of superabrasives.

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 26.6 shows the marking system for aluminum-oxide and silicon-carbide bonded abrasives, and Fig. 26.7 shows the marking system for diamond and cubic boron nitride bonded abrasives.

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 \$10 for conven-

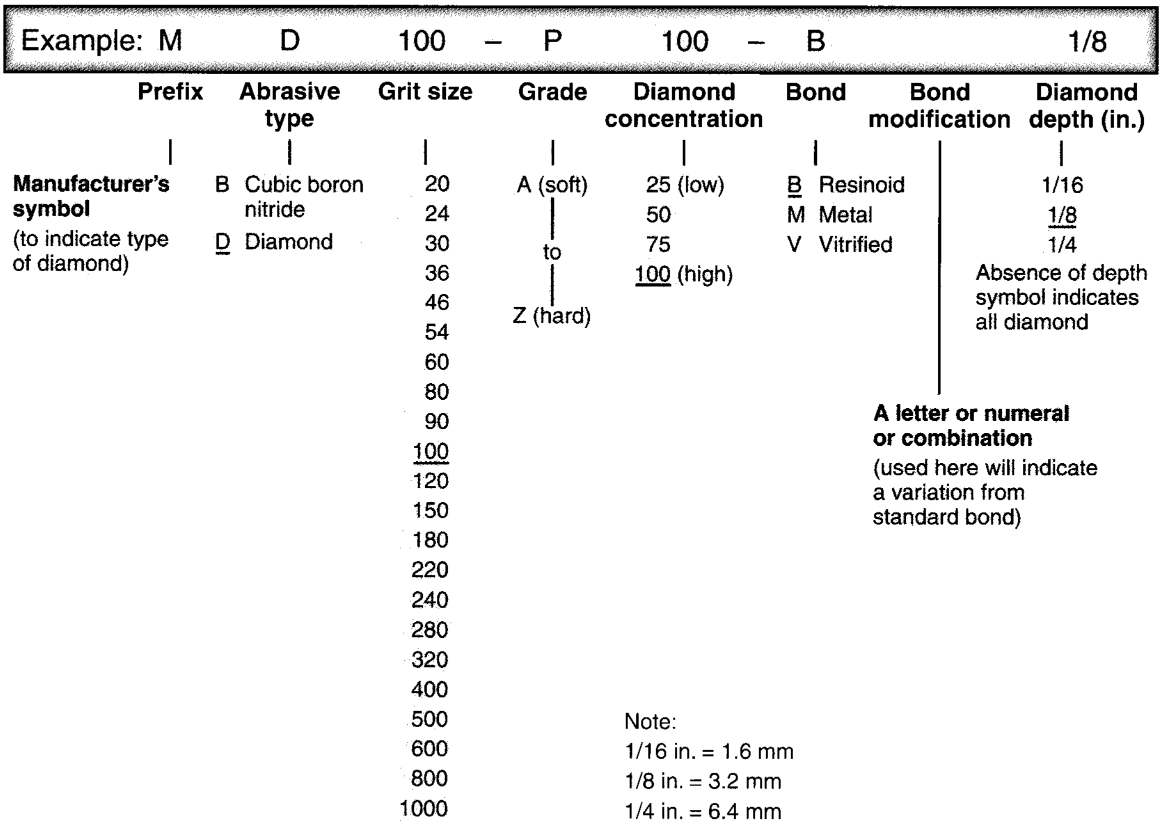


FIGURE 26.7 Standard marking system for cubic boron nitride and diamond bonded abrasives.

tional 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 wheel costs are \$500 for conventional abrasives, \$5,000 to \$8,000 for diamond, and almost \$20,000 for cubic boron nitride.

26.2.2 Bond Types

The common types of bonds used in bonded abrasives are as follows:

Vitrified. Essentially a glass, a *vitrified bond* (also called a *ceramic bond*) is the most common and widely used bond 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, which are similar to powder-metallurgy parts (Chapter 17), are 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 thermal cracking), finished to size, inspected for quality and dimensional accuracy, and tested for defects.

Wheels with vitrified bonds are strong, stiff, porous, and resistant to oils, acids, and water. However, they are brittle and lack resistance to mechanical and thermal shock. To improve their strength during use, vitrified wheels also are made with steel-backing plates or cups for better structural support of the bonded abrasive. The color of the grinding wheel can be modified by adding various elements during its manufacture; in this way, wheels can be color coded for use with specific workpiece materials, such as ferrous, nonferrous, ceramic, and so on.

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 manufacturing technique for producing 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.

Because the elastic modulus of thermosetting resins is lower than that of glasses, resinoid wheels are more flexible than vitrified wheels. A more recent development is the use of *polyimide* (Section 7.7) as a substitute for the phenolic in resinoid wheels; it is tougher and more resistant to higher temperatures. In addition to pressing, *injection molding* is another process used to manufacture grinding wheels (see Sections 17.3 and 19.3).

Reinforced Wheels. These wheels typically consist of one or more layers of *fiberglass mats* of various mesh sizes. The fiberglass in this laminate structure provides reinforcement in resinoid wheels by way of retarding the disintegration of the wheel should it break for some reason during use, rather than improving its strength. Large-diameter resinoid wheels can be supported additionally with one or more internal rings made of round steel bars inserted during the molding of the wheel.

Thermoplastic. In addition to thermsetting resins, thermoplastic bonds 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. 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 can be made in this manner and are used like saws for cutting-off operations (*cutoff blades*).

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, as shown in Fig. 26.5. 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 requirements such 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 with a particular desired shape. Layered wheels are lower in cost and are used for small production quantities.

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 bond in the wheel. Because strength and hardness are directly related (see 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 a soft wheel. The *structure* of a bonded abrasive is a measure of its *porosity* (i.e., the spacing between the grains, as shown in Fig. 26.3).

The structure of bonded abrasives ranges from *dense* to *open*, as shown in Fig. 26.6. As stated earlier, some porosity is essential to provide clearance for the grinding chips; otherwise, they would interfere with the grinding operation.

26.3 The Grinding Process

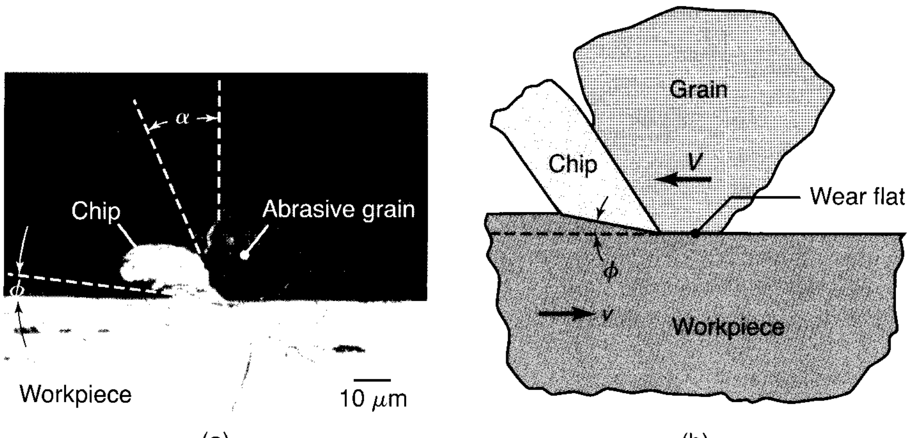
Grinding is a chip-removal process that uses an individual abrasive grain as the cutting tool (Fig. 26.8a). The major differences between the action of an abrasive grain and that of a single-point cutting tool can be summarized as follows:

- The individual abrasive grains have *irregular shapes* and are spaced *randomly* along the periphery of the wheel (Fig. 26.9).
- 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.)
- The radial positions of the grains over the peripheral surface of a wheel vary; thus, not all grains are active during grinding.
- Surface speeds (i.e., cutting speeds) in grinding are very high, typically 20 to 30 m/s, and may be as high as 150 m/s in high-speed grinding using specially designed and manufactured wheels.

The grinding process and its parameters can be observed best in the *surface-grinding* operation shown schematically in Fig. 26.10. A straight grinding wheel (Fig. 26.4a) with a diameter D removes a layer of metal at a depth d (**wheel depth of cut**). An individual grain on the periphery of the wheel moves at a tangential velocity V, while the workpiece moves at a velocity v. Each abrasive grain removes a small chip, which has an *undeformed thickness* (**grain depth of cut**), t, and an *undeformed length*, l.

From geometric relationships, it can be shown that the undeformed chip length, l, in surface grinding (Fig. 26.10) is approximated by the equation

$$l = \sqrt{Dd} \tag{26.1}$$



(a) (b)

FIGURE 26.8 (a) Grinding chip being produced by a single abrasive grain. Note the large negative rake angle of the grain. The inscribed circle is 0.065 mm in diameter. (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.

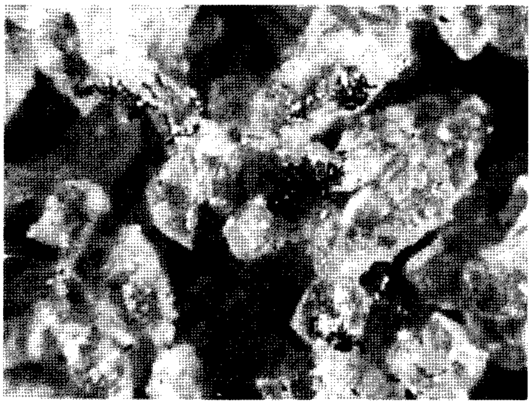


FIGURE 26.9 The surface of a grinding wheel (A46-J8V), showing abrasive grains, wheel porosity, wear flats on grains, and metal chips from the workpiece adhering to the grains. Note the random distribution and shape of the abrasive grains. Magnification: 50×. *Source*: S. Kalpakjian.

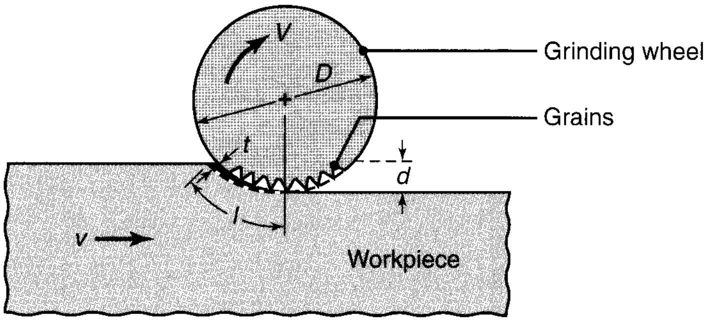


FIGURE 26.10 Schematic illustration of the surface-grinding process, showing various process variables. The figure depicts conventional (up) grinding.

and the undeformed chip thickness, t, by the equation

$$t = \sqrt{\left(\frac{4v}{VCr}\right)\sqrt{\left(\frac{d}{D}\right)}}, \tag{26.2}$$

where C is the number of cutting points per unit area of the periphery of the wheel; generally, C is estimated to be 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 has an estimated value typically between 10 and 20.

As an example, l and t can be calculated for the following process parameters: Let $D = 200$ 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 $r = 15$ results in

$$t = \sqrt{\frac{(4)(30)}{(1800)(2)(15)}\sqrt{\frac{0.05}{200}}} = 0.006 \text{ mm}$$

Because of plastic deformation, the actual chip will be shorter and thicker than the values calculated. (See Fig. 26.8.) It can be seen that grinding chips are much smaller than those typically obtained in metal-cutting operations, as described in Chapter 21.

Grinding Forces. A knowledge of grinding forces is essential for

- Estimating power requirements.
- Designing grinding machines and work-holding fixtures and devices.
- Determining the deflections that the workpiece, as well as the grinding machine itself, may undergo. Note that, unless accounted for, deflections adversely affect dimensional accuracy and are especially critical in precision and ultra-precision grinding.

TABLE 26.2

Approximate Specific-energy Requirements for Surface Grinding

Workpiece material	Hardness	Specific energy $\mathrm{W \cdot s/mm^3}$
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

If it is assumed 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** (which is tangential to the wheel) is proportional to the process variables:

$$\text{Grain force} \propto \left(\frac{v}{V}\sqrt{\frac{d}{D}}\right)(\text{UTS}). \qquad (26.3)$$

Forces in grinding are usually much smaller than those in the machining operations described in Chapters 23 and 24 because of the small dimensions involved. Grinding forces should be kept low in order to avoid distortion and to maintain the high dimensional accuracy of the workpiece.

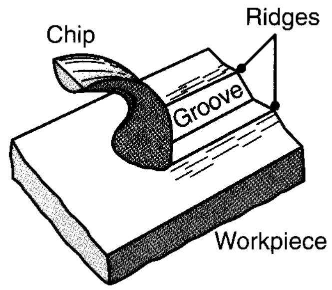


FIGURE 26.11 Chip formation and plowing of the workpiece surface by an abrasive grain.

Specific Energy. The *energy* dissipated in producing a grinding chip consists of the energy required for the following actions:

- Chip formation
- Plowing, as shown by the ridges formed in Fig. 26.11
- Friction, caused by rubbing of the grain along the workpiece surface.

Note that after some use, the grains along the periphery of the wheel develop a **wear flat** (Fig. 26.8b), a phenomenon that is similar to *flank wear* in cutting tools (Fig. 21.15). The wear flat continuously rubs along the ground surface, dissipates energy (because of friction), and makes the grinding operation less efficient.

Specific-energy requirements in grinding are defined as the energy per unit volume of material ground from the workpiece surface and are shown in Table 26.2. Note that these specific-energy levels are much higher than those in the machining operations given in Table 21.2. This difference has been attributed to factors such as the presence of a wear flat, high negative rake angles of the grains (which require more energy), and a possible contribution of the size effect (the smaller the chip, the higher the energy required to produce it). Also, it has been observed that with effective lubrication, specific energies in grinding can be reduced by a factor of four or more.

As shown in Example 26.1, the *grinding force* and the *thrust force* in grinding can be calculated from the specific-energy data.

EXAMPLE 26.1 Forces in Surface Grinding

A surface-grinding operation is being performed on low-carbon steel with a wheel of diameter $D = 250$ mm that is rotating at $N = 4000$ rpm and a width of cut of $w = 25$ mm. The depth of cut is $d = 0.05$ mm and the feed rate of the workpiece, v, is 1.5 m/min. Calculate the cutting force (the force tangential to the wheel), F_c, and the thrust force (the force normal to the workpiece surface), F_n.

Solution The material-removal rate (MRR) is determined as

$$\text{MRR} = dwv = (0.05)(25)(1500) = 1875 \text{ mm}^3$$

The power consumed is given by

$$\text{Power} = (u)(\text{MRR}),$$

where u is the specific energy, as obtained from Table 26.2. (See also Section 21.3.) For low-carbon steel, it is estimated to be 40 W • s/mm^3. Thus,

$$\text{Power} = 40 \frac{\text{W} \cdot \text{s}}{\text{mm}^3}\left(\frac{1 \text{ min}}{60 \text{ s}}\right) 1875 \frac{\text{mm}^3}{\text{min}}$$

By noting that 1W = 1N•m/s = 60 N•m/min,

$$\text{Power} = (1.25 \times 10^3)(60) = 75{,}000 \text{ N·m/min}$$

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 minute ($\omega = 2\pi N$). It follows that

$$75{,}000 = (F_c)\left(\frac{250}{2}\right)(2\pi)(4000)\left(\frac{1 \text{ m}}{1000 \text{ mm}}\right)$$

and therefore, $F_c = 24$ 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,

$$F_n = (1.3)(24) = 31 \text{ N}.$$

Temperature. The temperature rise in grinding is an important consideration because

- It can adversely affect the surface properties of the workpiece, including metallurgical changes.
- The temperature rise can cause residual stresses on 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:

$$\Delta T \propto D^{1/4} d^{3/4} \left(\frac{V}{v}\right)^{1/2}. \qquad (26.4)$$

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.

Peak temperatures during grinding can reach 1600°C. However, the time involved in producing a chip is extremely short (microseconds), so 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; see Section 25.5), only a fraction of the heat produced 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.

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 the sparks depend on the composition of the metal being ground. Charts are available that help identify the type of metal being ground from the appearance of its sparks. If the heat generated due to exothermic reaction is sufficiently high, chips can melt, acquire a spherical shape (because of surface tension), and solidify as metal particles.

Tempering. An excessive temperature rise in grinding can cause *tempering* and *softening* of the workpiece surface. Process variables must be selected carefully in order to avoid excessive temperature rise. The use of grinding fluids is an effective means of controlling temperature.

Burning. Excessive temperature during grinding may burn the workpiece surface. A *burn* is characterized by a bluish color on ground steel surfaces—an indication that high temperature caused oxidation. It can be detected by etching and metallurgical techniques. A burn may not be objectionable in itself, unless surface layers have undergone *phase transformations* (Chapter 4). For example, martensite forming in higher carbon steels from rapid cooling is called a *metallurgical burn*. This condition will adversely influence the surface properties of ground parts, reducing surface ductility and toughness.

Heat Checking. High temperatures in grinding may cause the workpiece surface to develop cracks; this condition is known as *heat checking*. The cracks usually are perpendicular to the grinding direction. Under severe conditions, however, parallel cracks also may appear. As expected, such a surface lacks toughness and has low fatigue and corrosion resistance. (Note that heat checking also occurs in dies during die casting, as described in Section 11.4.5.)

Residual Stresses. Temperature gradients within the workpiece during grinding are primarily responsible for *residual stresses*. Grinding fluids and their method of application, as well as process parameters such as depth of cut and speeds, significantly influence the magnitude and type of residual stresses developed (tension or compression). Because of the adverse effect of tensile residual stresses on fatigue strength, process variables should be selected carefully. 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** wheels) also may be used.

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. Grinding-wheel wear is caused by three different mechanisms: attritious grain wear, grain fracture, and bond fracture.

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. 26.8b). Wear involves both physical and chemical reactions, and is caused by the interaction of the grain with the workpiece material. The reactions are complex and 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 (grain and workpiece) are *chemically inert* with respect to each other, much like what has been observed with cutting tools. The more inert the materials, the lower the tendency for reaction and adhesion to occur between the grain and the workpiece. For example, because aluminum oxide is relatively inert with respect to iron, its rate of attritious wear when it is used to grind steels is much lower than that of silicon carbide and diamond. By contrast, silicon carbide can dissolve in iron, so it is not suitable for grinding steels. Cubic boron nitride has a higher inertness with respect to steels; hence, it is suitable for use as an abrasive.

Therefore, the selection of the type of abrasive for low attritious wear is based on the *reactivity* of the grain with the workpiece and on their relative mechanical properties, such as hardness and toughness. The environment and the type of grinding fluid used also have an influence on grain–workpiece interactions.

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, the grain should fracture or fragment at a moderate rate, so that new sharp cutting edges are produced continuously during grinding. This is equivalent to breaking a dull piece of chalk or a stone into two or more pieces in order to expose new sharp edges. Section 26.2 described the *friability* of abrasives (the extent to which they are self-sharpening) as an important factor in effective grinding.

The selection of grain type and size for a particular application also depends on the attritious-wear rate. A grain–workpiece material combination with high attritious wear and low grain friability dulls grains and develops a large wear flat. Grinding then becomes inefficient, and surface damage (such as burning) is likely to occur.

Bond Fracture. The strength of the bond (grade) is a significant parameter in grinding. If the bond is too strong, dull grains cannot be dislodged. This prevents other sharp grains along the circumference of the grinding wheel from contacting the workpiece to remove chips, and the grinding process becomes inefficient. On the other hand, if the bond is too weak, the grains are dislodged easily, and the wear rate of the wheel increases. In this case, maintaining dimensional accuracy becomes difficult.

In general, softer bonds are recommended for harder materials and for reducing residual stresses and thermal damage to the workpiece. Hard-grade wheels are used for softer materials and for removing large amounts of material at high rates.

26.3.2 Grinding Ratio

Grinding-wheel wear is generally correlated with the amount of workpiece material ground by a parameter called the *grinding ratio,* G, defined as

$$G = \frac{\text{Volume of material removed}}{\text{Volume of wheel wear}}. \tag{26.5}$$

In practice, grinding ratios vary widely, ranging from 2 to 200 and even higher, depending on the type of wheel, workpiece material, grinding fluid, and process parameters (such as depth of cut and speeds of the wheel and workpiece). It has been shown that effective grinding fluids can increase the grinding ratio by a factor of 10 or more, thus greatly improving wheel life.

During grinding, a particular wheel may **act soft** (i.e., exhibit high wear) or **act hard** (exhibit low wear), regardless of the wheel grade. Note, for example, that an ordinary pencil *acts* soft when writing on rough paper and 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 greater the tendency for the grains to fracture or be dislodged from the wheel surface, the higher the wheel wear, and the lower the grinding ratio. From Eq. (26.3), note that the grain force increases with the strength of the workpiece material, work speed, and depth of cut and decreases with increasing wheel speed and wheel diameter. Thus, a grinding wheel acts soft when v and d increase or when V and D decrease.

Note also that attempting to obtain a high grinding ratio in practice (so as to extend wheel life) isn't always desirable, because high ratios may indicate grain dulling and thus possible surface damage of 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

A surface-grinding operation is being carried out with the wheel running at a constant spindle speed. Will the wheel act soft or hard as the wheel wears down over time? Assume that the depth of cut, d, remains constant and the wheel is dressed periodically (see Section 26.3.3).

Solution Referring to Eq. (26.3), note that the parameters which change over time in this operation are the wheel diameter, D, and its surface speed, V. As D becomes smaller, the relative grain force increases; thus, the wheel acts softer. To accommodate the changes due to the reduction of wheel diameter 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

- *Conditioning* worn grains on the surface of a grinding wheel by producing *sharp new edges* on grains so that they cut more effectively.
- *Truing*, which is producing a true circle on a wheel that 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. 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. 26.9) become filled or clogged with chips from the workpiece. Loading can occur in grinding soft materials or from improper selection of wheels or process parameters. A loaded wheel cuts inefficiently and generates much frictional heat, which results in surface damage and loss of dimensional accuracy of the workpiece.

The techniques used to dress grinding wheels are as follows:

- A specially shaped *diamond-point tool* or *diamond cluster* is moved across the width of the grinding face of a rotating wheel and removes a small 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, respectively. In practice, however, the wear of the diamond can be significant with harder wheels, in which case a diamond disk or cup wheel can be used.
- A set of *star-shaped steel disks* is pressed manually against the wheel. Material is removed from the wheel surface by crushing the grains. As a result, this

method produces a coarse surface on the wheel and is used only for rough grinding operations on bench or pedestal grinders.

- *Abrasive sticks* may be 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, as described in Chapter 27. These processes erode away very thin layers of the metal bond and thus expose new diamond cutting edges.
- Dressing for form grinding involves *crush dressing* or *crush forming*. The process consists of pressing a metal roll on the surface of the grinding wheel, which typically is a vitrified wheel. The roll (which usually is made of high-speed steel, tungsten carbide, or boron carbide) has a machined or ground profile (form) on its periphery. Thus, it reproduces a replica of this profile on the surface of the grinding wheel being dressed. (See Section 26.4.)

Dressing techniques and their frequency are important for quality, because they affect grinding forces and workpiece surface finish. Modern computer-controlled grinders are equipped with automatic dressing features, which dress the wheel continually 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 usually is monitored precisely by using piezoelectric or acoustic-emission sensors (Section 37.7). Furthermore, features such as 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 would be 2 to 10 μm. Consequently, 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. 26.4a is cylindrical; therefore, it produces a flat surface. However, the surface can be shaped into various forms by dressing it (Fig. 26.12a). Although templates have been used for this purpose, modern grinders commonly 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. 26.12b) and produces very accurate surfaces. Note in Fig. 26.12b 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, like the terms *machinability* (Section 21.7) or *forgeability* (Section 14.5) is difficult to define precisely. Grindability is a general indicator of how easy it is to grind a material and includes considerations such as the quality of the surface produced, surface finish, surface integrity, wheel wear, cycle time, and overall economics of the operation. As in machinability, grindability of a material can be enhanced greatly by proper selection of process parameters (Table 26.3), grinding wheels, and grinding fluids, as well as by using the appropriate machine characteristics, fixturing methods, and work-holding devices.

Grinding practices are now well established for a wide variety of metallic and nonmetallic materials, including newly developed aerospace materials and composites. Specific recommendations for selecting wheels and appropriate process parameters for metals can be found in various handbooks, manufacturers' literature, and references in the bibliography of this chapter.

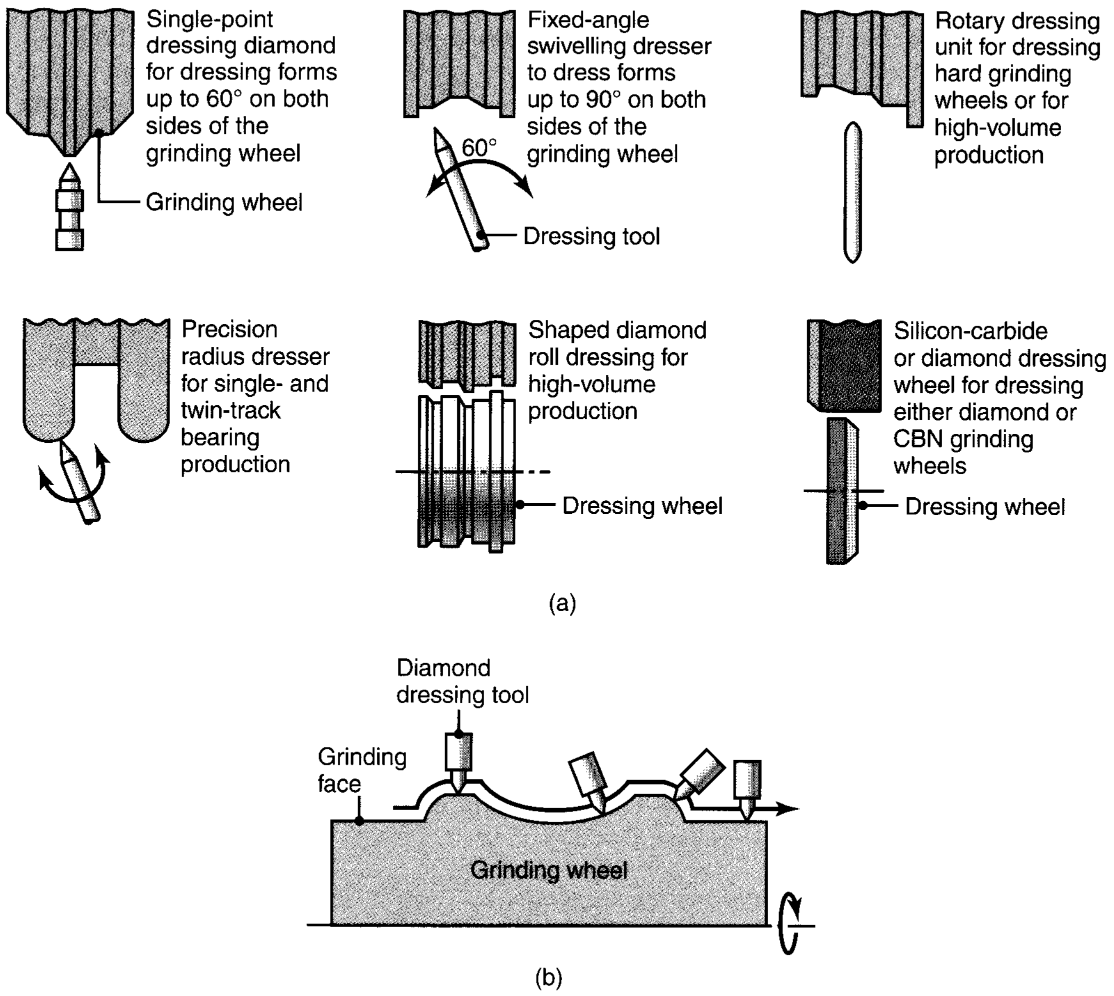


FIGURE 26.12 (a) Forms of grinding-wheel dressing. (b) Shaping the grinding face of a wheel by dressing it by computer control. Note that the diamond dressing tool is normal to the surface at the point of contact with the wheel. *Source*: Courtesy of Okuma Machinery Works, Ltd.

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	—	—

Wheel selection involves not only the shape of the wheel and the shape of the part to be produced, but the characteristics of the workpiece material as well. On the basis of the discussion thus far, it can be seen that the physical and mechanical properties of the workpiece material are important in the selection of a type of abrasive and a bond.

Ductile-regime Grinding. Ceramics can be ground with relative ease using diamond wheels, as well as carefully selected process parameters. It has been shown, for example, that with light passes and machines with high stiffness and damping capacity, it is possible to produce continuous chips and good surface integrity by grinding brittle materials (Fig. 26.11) in a process known as *ductile-regime grinding.* However, because ceramic chips are typically 1 to 10 μm in size, they are more difficult to remove from grinding fluids than metal chips and require fine filters and special techniques.

26.4 Grinding Operations and Machines

The selection of a grinding process and machine 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 features such as automatic workpiece loading and unloading, part clamping, dressing, and wheel shaping. Grinders also 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 detected readily during the dressing cycle.

The relative movement of the 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). Surface grinders make up the largest percentage of grinders used in industry, followed by bench grinders (usually with two wheels at each end of the spindle), cylindrical grinders, and tool and cutter grinders—the least common being internal grinders.

Surface Grinding. *Surface grinding* is one of the most common operations (Fig. 26.13), generally involving the grinding of flat surfaces. Typically, the workpiece is secured on a *magnetic chuck* attached to the worktable of the *grinder*

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
Ultrasonic machining	Holes and cavities with various shapes; suitable for hard and brittle materials; medium labor skill	—

*Larger capacities are available for special applications.

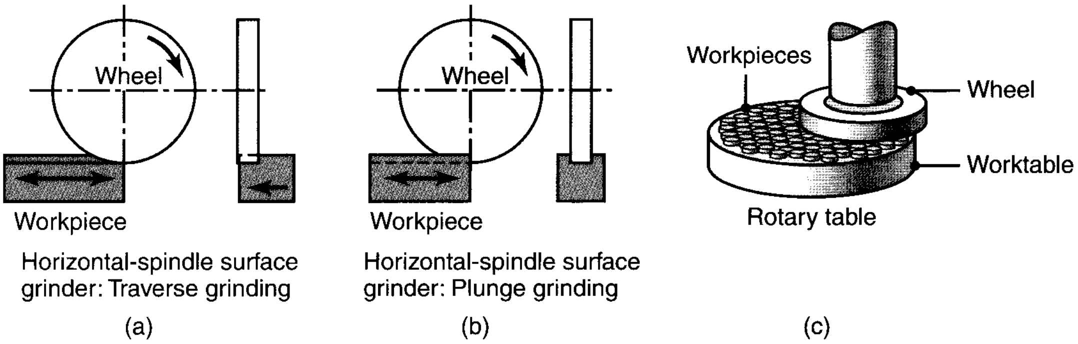


FIGURE 26.13 Schematic illustrations of various surface-grinding operations. (a) Traverse grinding with a horizontal-spindle surface grinder. (b) Plunge grinding with a horizontal-spindle surface grinder, producing a groove in the workpiece. (c) A vertical-spindle rotary-table grinder (also known as the *Blanchard* type).

(Fig. 26.14); nonmagnetic materials are held by vises, vacuum chucks, or some other fixture. A straight wheel is mounted on the horizontal spindle of the surface grinder. Traverse grinding occurs as the table reciprocates longitudinally and is fed laterally (in the direction of the spindle axis) after each stroke. In *plunge grinding*, the wheel is moved radially *into* the workpiece, as it is when grinding a groove (Fig. 26.13b).

In addition to the surface grinder shown in Fig. 26.14, other types include *vertical spindles* and *rotary tables* (referred to as the *Blanchard* type, Fig. 26.13c). These configurations allow a number of 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. 26.15).

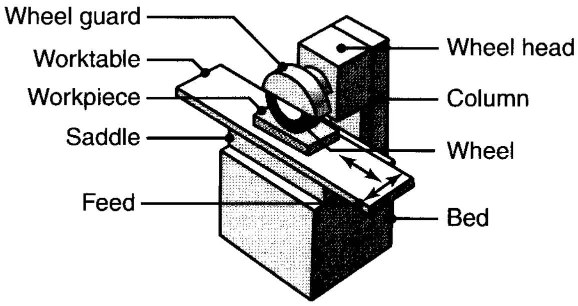


FIGURE 26.14 Schematic illustration of a horizontal-spindle surface grinder.

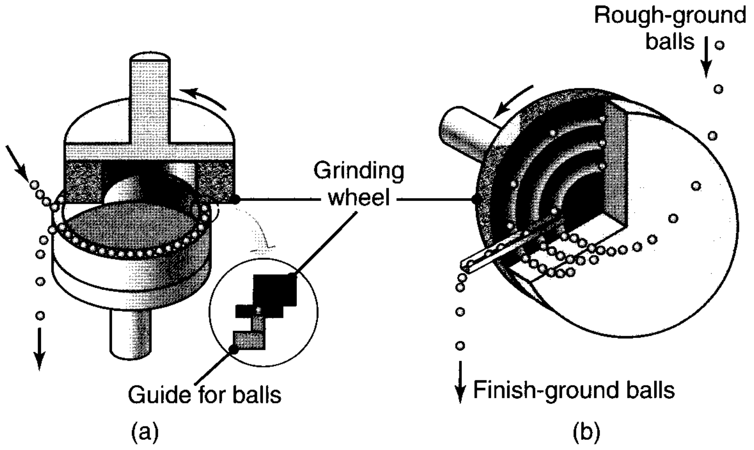


FIGURE 26.15 (a) Rough grinding of steel balls on a vertical-spindle grinder. The balls are guided by a special rotary fixture. (b) Finish grinding of balls in a multiple-groove fixture. The balls are ground to within 0.013 mm of their final size.

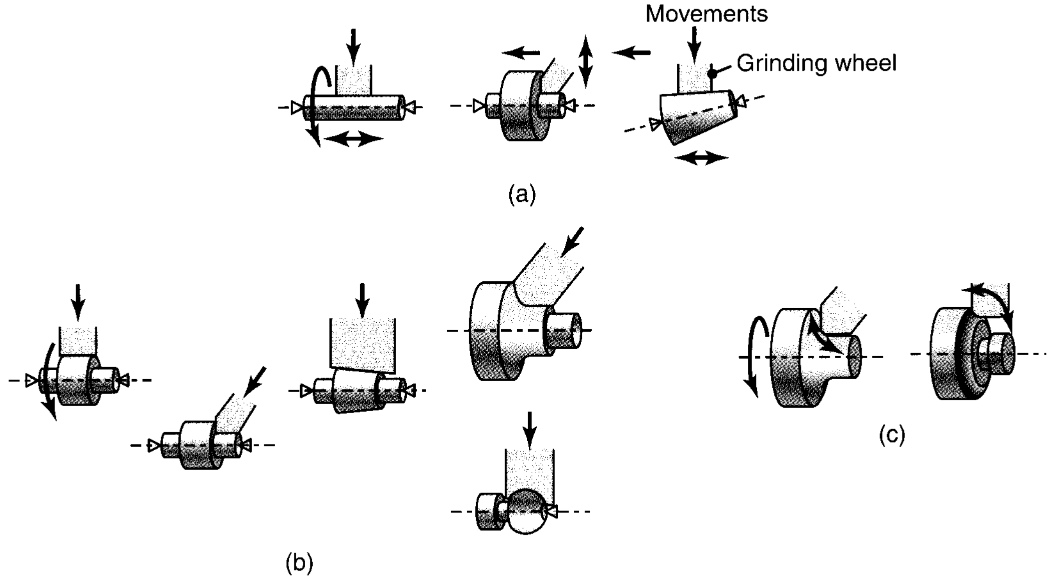


FIGURE 26.16 Examples of various cylindrical-grinding operations: (a) traverse grinding, (b) plunge grinding, and (c) profile grinding. *Source*: Courtesy of Okuma Machinery Works, Ltd.

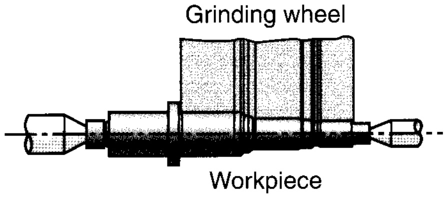


FIGURE 26.17 Plunge grinding of a workpiece on a cylindrical grinder with the wheel dressed to a stepped shape.

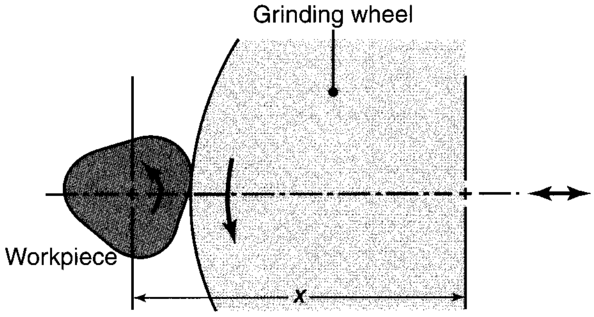


FIGURE 26.18 Schematic illustration of grinding a noncylindrical part on a cylindrical grinder with computer controls to produce the shape. The part rotation and the distance x between centers are varied and synchronized to grind the particular workpiece shape.

Cylindrical Grinding. In *cylindrical grinding* (also called *center-type grinding*, Fig. 26.16; see also Fig. 26.2), 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 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 machines 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 workpiece are parallel. The wheel and workpiece are each driven by separate motors and at different speeds. Long workpieces with two or more diameters can be ground on cylindrical grinders. As *form grinding* and *plunge grinding*, cylindrical grinding also can produce shapes in which the wheel is dressed to the workpiece form to be ground (Fig. 26.17).

Cylindrical grinders are identified by the maximum diameter and length of the workpiece that can be ground—as are engine lathes. In *universal grinders*, both the workpiece and the wheel axes can be moved and swiveled around a horizontal plane, thus permitting the grinding of tapers and other shapes.

With computer control, *noncylindrical* parts such as cams can be ground on rotating workpieces. As illustrated in Fig. 26.18, the workpiece spindle speed is synchronized such that the radial distance, x, between the workpiece and the wheel axes is varied continuously to produce a particular shape, such as the one shown.

Thread grinding is done on cylindrical grinders using specially dressed wheels matching the shape of the threads, as shown in Fig. 26.19. (See also *centerless grinding.*) Although expensive, threads produced by grinding are the most accurate of any manufacturing process and have a very fine surface finish; typical applications requiring such threads include ballscrew mechanisms used for precise movement of machine components. The workpiece and wheel movements are synchronized to produce the pitch of the thread, usually in about six passes.

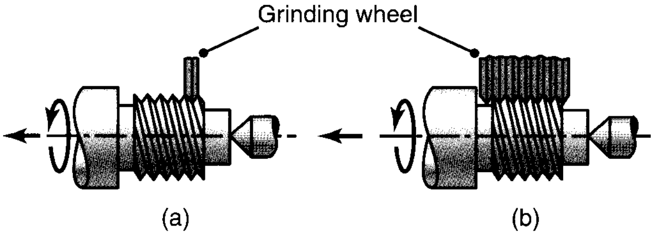


FIGURE 26.19 Thread grinding by (a) traverse and (b) plunge grinding.

EXAMPLE 26.3 Cycle Patterns in Cylindrical Grinding

As in most grinding operations, the grinding wheel in cylindrical grinding typically makes several passes along a path in order to produce the final geometry on the workpiece. Figure 26.20 illustrates the cycle patterns for producing various shapes on a multifunctional, computer-controlled precision grinder. The downward arrowheads with numbers indicate the beginning of the grinding cycle.

The determination of the optimum and most economical pattern for minimum cycle time depends on the amount of material to be removed, the shape of the part, and the process parameters chosen. These patterns are generated automatically by the software in the computer controls of the grinder.

Source: Courtesy of Toyoda Machinery.

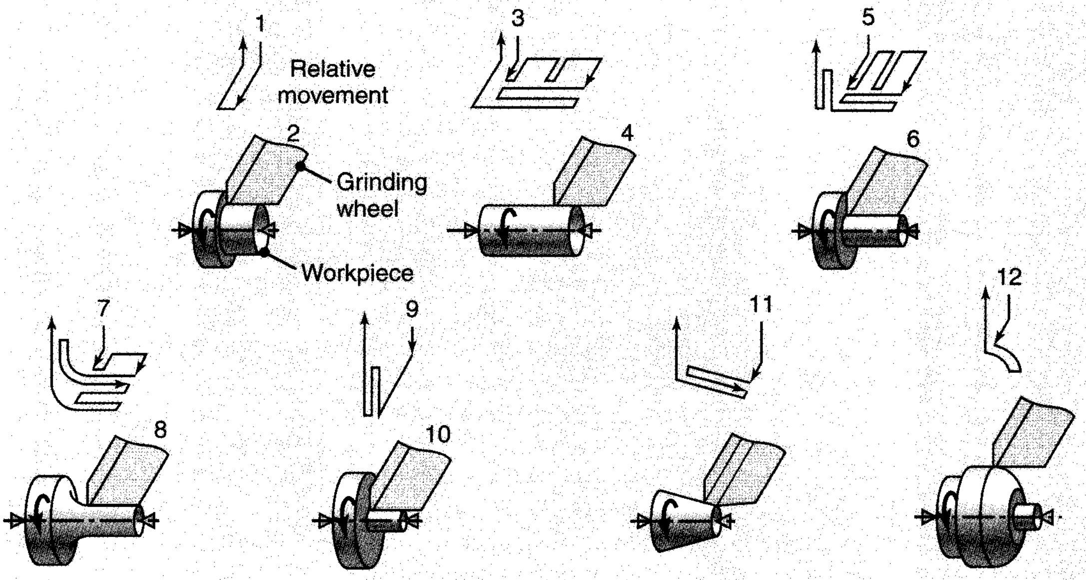
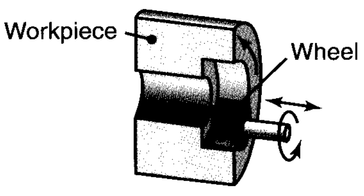
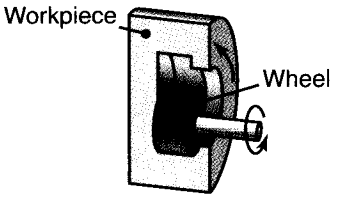


FIGURE 26.20 Cycle patterns for a CNC precision grinder.

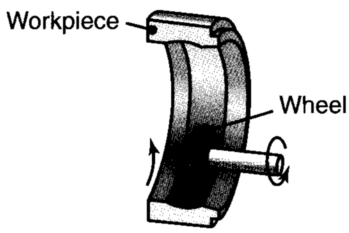
Internal Grinding. In *internal grinding* (Fig. 26.21), a small wheel is used to grind the inside diameter of the part, such as in bushings and bearing races. The workpiece is held in a rotating chuck and the wheel rotates at 30,000 rpm or higher. Internal profiles also can be ground with profile-dressed wheels that move radially into the workpiece. The headstock of internal grinders can be swiveled on a horizontal plane to grind tapered holes.



(a) Traverse grinding



(b) Plunge grinding



(c) Profile grinding

FIGURE 26.21 Schematic illustrations of internal grinding operations: (a) traverse grinding, (b) plunge grinding, and (c) profile grinding.

Centerless Grinding. *Centerless grinding* is a high-production process for continuously grinding cylindrical surfaces in which the workpiece is supported not by centers (hence the term "centerless") or chucks, but by a *blade* (Figs. 26.22a and b). Typical parts made by centerless grinding are roller bearings, piston pins, engine valves, camshafts, and similar components. Parts with diameters as small as 0.1 mm can be ground. Centerless grinders (Fig. 26.22d) are capable of wheel surface speeds on the order of 10,000 m/min, typically using cubic boron nitride abrasive wheels.

In *through-feed grinding*, the workpiece is supported on a work-rest blade and is ground between two wheels. Grinding is done by the larger wheel, while the smaller wheel regulates the axial movement of the workpiece. The *regulating wheel* (which is rubber bonded) is tilted and runs at a speed of only about one-twentieth of the grinding-wheel speed.

Parts with variable diameters (such as bolts, valve tappets, and multiple-diameter shafts) can be ground by centerless grinding in a process called *infeed* or *plunge grinding* (Fig. 26.22b). This operation is similar to plunge or form grinding on cylindrical grinders. Tapered pieces are centerless ground by *end-feed grinding*. High-production-rate thread grinding also can be done 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. 26.22c).

Creep-feed Grinding. Grinding traditionally has been associated with small rates of material removal (Table 26.3) and fine finishing operations. However, grinding can also be used for large-scale metal-removal operations similar to, and competing with, milling, broaching, and planing. In *creep-feed grinding*, the wheel depth of cut, *d*, is as much as 6 mm and the workpiece speed is low (Fig. 26.23). To keep workpiece temperatures low and improve surface finish, the wheels are softer grade resin bonded and have an open structure.

The machines 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 grinding fluids. Grinders are equipped with features for continuously dressing the wheel, using a diamond roll as the dressing tool.

Creep-feed grinding can be competitive with other machining processes and economical for specific applications, such as grinding shaped punches, key seats, twist-drill flutes, the roots of turbine blades (Fig. 26.23c), and various complex superalloy parts. Because the wheel is dressed to the shape of the workpiece to be produced, the workpiece does not have to be previously shaped by milling, shaping, or broaching. Near-net-shape castings and forgings are thus 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. This operation can be economical in specific applications and can compete favorably with machining processes, particularly milling, turning, and broaching. In this operation, surface finish is of secondary importance and the grinding wheel (or belt) is utilized to its fullest while minimizing cost per piece. The dimensional tolerances are on the same order as those obtained

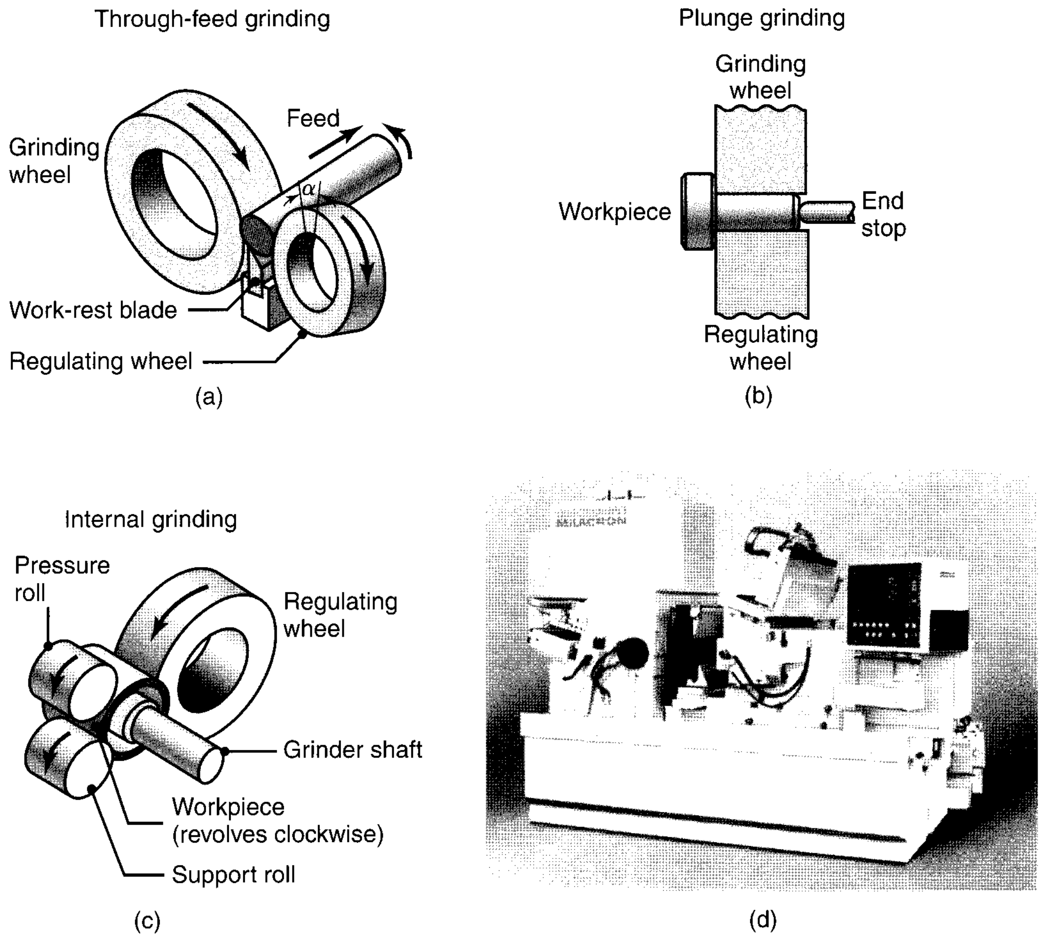


FIGURE 26.22 Schematic illustrations of centerless grinding operations: (a) through-feed grinding, (b) plunge grinding, (c) and internal grinding; (d) a computer numerical-control cylindrical-grinding machine. *Source*: Courtesy of Cincinnati Milacron, Inc.

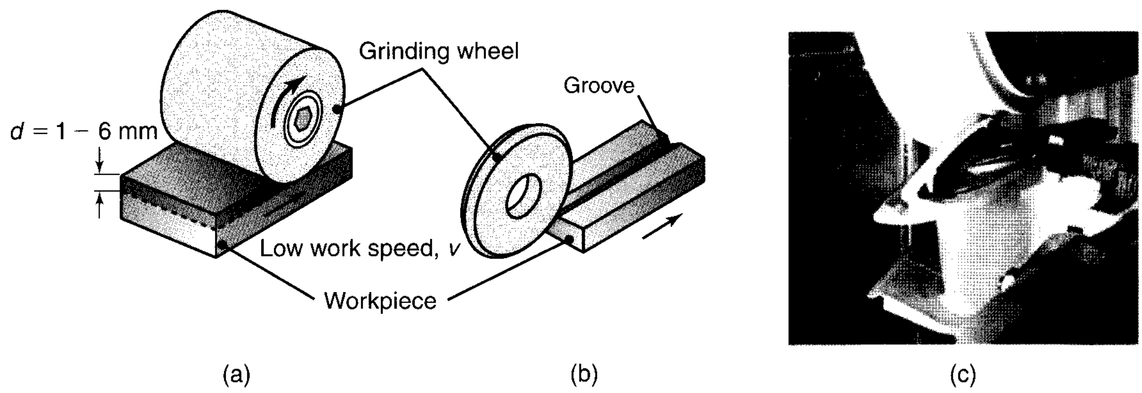


FIGURE 26.23 (a) Schematic illustration of the creep-feed grinding process. Note the large wheel depth of cut, d. (b) A shaped groove produced on a flat surface by creep-feed grinding in one pass. Groove depth is typically on the order of a few mm. (c) An example of creep-feed grinding with a shaped wheel. This operation also can be performed by some of the processes described in Chapter 27. *Source*: Courtesy of Blohm, Inc.

by most machining processes. Heavy stock removal by grinding is also performed on welds, castings, and forgings to smoothen weld beads and remove flash.

EXAMPLE 26.4 Grinding versus Hard Turning

In Section 25.6, *hard turning* was described. An example is the machining of heat-treated steels (usually above 45 HRC) using a single-point polycrystalline cubic boron nitride cutting tool. In view of the discussions presented thus far, it is evident that grinding and hard turning can be competitive in specific applications. Consequently, there has been considerable debate regarding the respective merits of the two processes.

Hard turning continues to be increasingly competitive with grinding, and dimensional tolerances and surface finish are beginning to approach those obtained with grinding. As seen by comparing Tables 21.2 and 26.2, turning requires much less energy than grinding. Also, thermal and other types of damage to the workpiece surface is less likely to occur, cutting fluids may not be necessary, and the machine tools are less expensive. In addition, finishing the part while it is still chucked in the lathe eliminates the need for material handling and setting the part in the grinder.

However, work-holding devices for large and slender workpieces during hard turning can present significant problems because the cutting forces used are higher than grinding forces. Furthermore, tool wear and its control can be a significant problem compared with the automatic dressing of grinding wheels. It is evident that the competitive position of hard turning versus grinding has to be evaluated individually for each application in terms of product surface finish, integrity, quality, and overall economics.

Other Grinding Operations. Various operations that are carried out on special-purpose grinders are summarized as follows:

- *Universal tool and cutter grinders* are used for grinding single-point or multipoint tools and cutters, including drills. They are equipped with special work-holding devices for accurate positioning of the tools to be ground. A variety of CNC tool grinders is available, making the operation simpler and faster and with consistent results. However, the cost of these grinders ranges from about \$150,000 to \$400,000.
- *Tool-post grinders* are self-contained units and usually are attached to the tool post of a lathe (as in Fig. 23.2). The workpiece is mounted on the headstock and is ground by moving the tool post. These grinders are versatile, but the lathe components should be protected from abrasive debris.
- *Swing-frame grinders* are 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 driven pneumatically, electrically, or with a flexible shaft connected to an electric motor or a gasoline engine. They are used for operations such as grinding off weld beads and *cutting-off* operations using thin abrasive disks.
- *Bench and pedestal grinders* are used for the routine offhand grinding of tools and small parts. They usually are equipped with two 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, as described in 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.

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.)

- Improves the efficiency of the operation by reducing wheel wear and loading and by lowering power consumption.

Grinding fluids are typically 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 (which is a mixture of fluid and air). Because of the high surface speeds involved, an airstream (*air blanket*) around the periphery of the wheel usually prevents 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 in which the grinding fluid is applied 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. This can cause the workpiece to expand, making it difficult to control dimensional tolerances. To maintain a more uniform workpiece temperature, a common method is to use refrigerating systems (chillers) through which the grinding fluid is circulated continuously and maintained at a roughly constant temperature. As described in Section 22.12, the biological and ecological aspects of disposing of, treating, and recycling, 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 operations because it adversely affects surface finish and wheel performance. Studying **chatter marks** on ground surfaces often can identify their source, which may include (a) the bearings and spindles of the grinding machine, (b) nonuniformities in the grinding wheel (as manufactured), (c) uneven wheel wear, (d) poor dressing techniques, (e) using 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.

The important factors in controlling chatter are the stiffness of the grinding machine, the stiffness of work-holding devices, and damping. General guidelines have been established to reduce the tendency for chatter in grinding, especially (a) using soft-grade wheels, (b) dressing the wheel frequently, (c) changing dressing techniques, (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 and cause serious injury. Certain procedures must therefore be followed carefully in their handling, storage, and use, and 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 high temperature or humidity. They should be inspected visually for cracks and damage prior to installing them on grinders. Vitrified wheels should be tested prior to use by *ringing* them (i.e., 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.

Damage to a grinding wheel can reduce its **bursting speed** severely. 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—that is, its bond, grade, and structure. In diamond and cBN wheels (Fig. 26.5), which are operated at high surface speeds, the type of core material used in the wheel affects the bursting speed. Metal cores, for example, have the highest bursting speed—on the order of about 250 m/s.

26.5 Design Considerations for Grinding

Design considerations for grinding operations are similar to those for machining that were described in various sections in Chapters 23 and 24. In addition, specific attention should be given to the following:

- Parts to be ground should be designed so that they can be mounted securely, either in chucks, magnetic tables, or suitable fixtures and work-holding devices. Thin, straight, or tubular workpieces may distort during grinding, thus 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, and 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 by prior machining.
- In centerless grinding, short pieces may be difficult to grind accurately because the blade does not support them. In through-feed grinding, only the largest diameter on the parts can be ground.
- Designs requiring accurate form grinding should be kept simple 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, designs should require that a minimum amount of material be removed by grinding—except 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. (This guideline is applicable to all manufacturing processes and operations.)

26.6 Ultrasonic Machining

In *ultrasonic machining* (UM), material is removed from a surface by microchipping and erosion with loose, fine abrasive grains in a water slurry (Fig. 26.24a). The tip of the tool (called a *sonotrode*) vibrates at a frequency of 20 kHz and a low

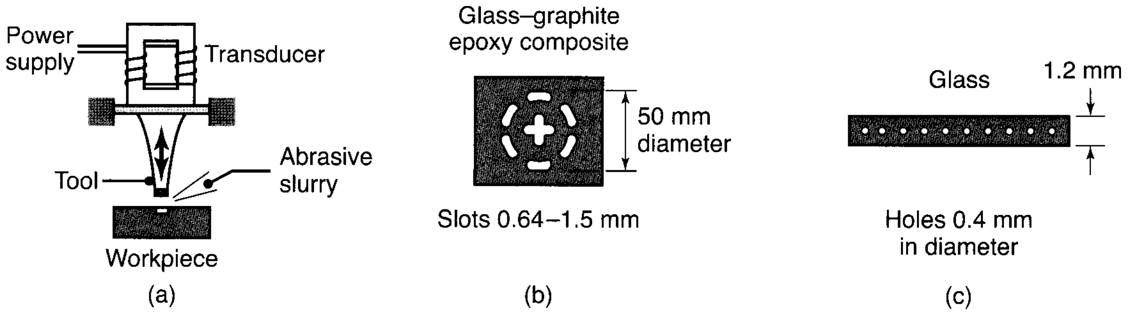


FIGURE 26.24 (a) Schematic illustration of the ultrasonic machining process. (b) and (c) Types of parts made by this process. Note the small size of the holes produced.

amplitude of 0.0125 to 0.075 mm. This vibration imparts a high velocity to abrasive grains between the tool and the workpiece. The stress produced by the impact of abrasive particles on the workpiece surface is high because (a) the time of contact between the particle and the surface is very short (10 to 100 μs) and (b) the area of contact is very small. In brittle materials, these impact stresses are sufficiently high to cause microchipping and erosion of the workpiece surface.

Ultrasonic machining is best suited for materials that are hard and brittle, such as ceramics, carbides, precious stones, and hardened steels. Two applications of ultrasonic machining are shown in Figs. 26.24b and c. A special tool is required for each shape to be produced; hence it is also called a *form tool.* The tip of the tool (which is attached to a transducer through the toolholder) usually is made of mild steel.

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 for finishing operations. The grains are carried in a water slurry with concentrations of 20 to 60% by volume; the slurry also carries the debris away from the cutting zone.

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 a face-milling operation (Fig. 24.5), but with the inserts being replaced by abrasives. The chips produced are washed away by a coolant that is pumped through the core of the rotating tool. Rotary ultrasonic machining (RUM) is particularly effective in producing deep holes and high material-removal rates in ceramic parts.

Design Considerations for Ultrasonic Machining. The basic design guidelines for UM include the following:

- Avoid sharp profiles, corners, and radii, because these are eroded by the abrasive slurry.
- Realize that holes produced will have some taper.
- Note that, because of the tendency for the chipping of brittle materials at the exit end of holes machined, the bottom of the parts should have a backup plate.

26.7 Finishing Operations

In addition to those described thus far, several processes that utilize fine abrasive grains are used on workpieces as the final finishing operation. However, these operations can significantly affect production time and product cost. Thus, they should be specified with due consideration to their costs and benefits.

Coated Abrasives. Common examples of *coated abrasives* are sandpaper and emery cloth. The majority of coated abrasives are made of aluminum oxide, with silicon carbide and zirconia alumina making up the rest. Coated abrasives usually have a much more open structure than the abrasives on grinding wheels, and they have grains that are more pointed and aggressive. The grains are deposited electrostatically on flexible backing materials, such as paper, cotton, rayon polyester, polynylon, and various blends thereof. As shown in Fig. 26.25, the bonding material (matrix) typically is resin, which first is applied to the backing (*make coat*); then the grains are bonded with a second layer (*size coat*). The grains have their long axes perpendicular to the plane of the backing, thus improving their cuttting action.

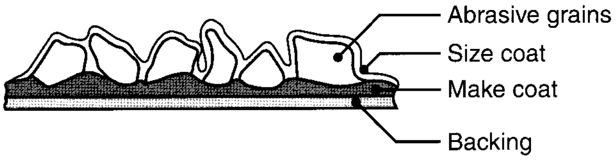


FIGURE 26.25 Schematic illustration of the structure of a coated abrasive. Sandpaper (developed in the 16th century) and emery cloth are common examples of coated abrasives.

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 in woodworking. The precision of the surface finish obtained depends primarily on the grain size.

Belt Grinding. Coated abrasives also are used as *belts* for high-rate material removal with good surface finish. *Belt grinding* has become an important production process, in some cases replacing conventional grinding operations. Belts with grit numbers ranging from 16 to 1500 are available. Belt speeds are in the range from 700 to 1,800 m/min. Machines for abrasive-belt operations require proper belt support and have 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 abrasives in the shape of tiny aluminum-oxide pyramids are placed in a predetermined orderly arrangement on the belt surface. Used on stainless steels and superalloys, they perform more consistently than conventional coated abrasives and the temperatures involved are lower. Typical applications for finishing with coated abrasives include belt grinding of surgical implants, golf clubs, firearms, turbine blades, and medical and dental instruments.

EXAMPLE 26.5 Belt Grinding of Turbine Nozzle Vanes

The turbine nozzle vane shown in Fig. 26.26 was investment cast (Section 11.8) 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 1,800 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.

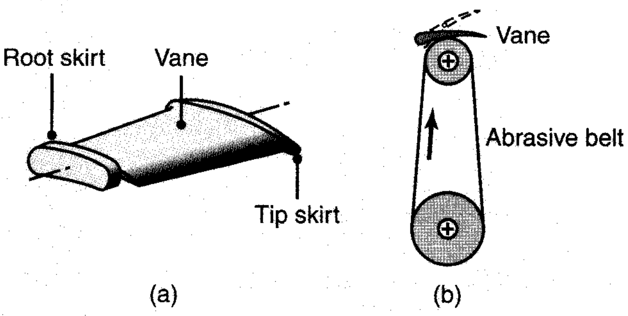


FIGURE 26.26 Turbine nozzle considered in Example 26.5.

Wire Brushing. In this process, also called *power brushing*, the workpiece is held against a circular wire brush that rotates at speeds ranging from 1750 rpm for large wheels to 3500 rpm for small wheels. As they rub, the tips of the wire produce longitudinal scratches on the workpiece surface. Wire brushing is used to produce a fine or controlled surface texture. Performed under the proper conditions, wire brushing also may be considered as a light material-removal process. In addition to metal wires, polymeric wires (such as nylon) embedded with abrasives can be used effectively. (See also *diamond wire saws*, Section 24.5.)

Honing. *Honing* is an operation that is used primarily to improve the surface finish of holes produced by processes such as boring, drilling, and internal grinding. The honing tool consists of a set of aluminum-oxide or silicon-carbide bonded abrasive sticks, usually called *stones* (Fig. 26.27). They are mounted on a mandrel that rotates in the hole at surface speeds of 45 to 90 m/min, applying a radial force. The tool has a reciprocating axial motion, which produces a crosshatched pattern on the surface of the hole. The stones can be adjusted radially for different hole sizes. Oil- or water-based honing fluids generally are used to help flush away the debris. Honing is also done on external cylindrical or flat surfaces and to manually remove sharp edges on cutting tools and inserts. (See Fig. 22.5.)

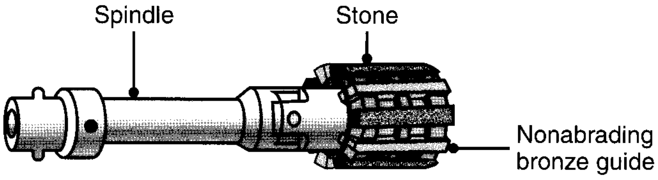


FIGURE 26.27 Schematic illustration of a honing tool used to improve the surface finish of bored or ground holes.

The quality of the surface finish produced by honing can be controlled by the type and size of abrasive used, the pressure applied, and speed. A fluid is used to remove chips and to keep temperatures low. 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 a short stroke. The motion of the stone is controlled so that the grains do not travel along the same path on the surface of the workpiece. Examples of external superfinishing of a round part are shown in Fig. 26.28.

Lapping. This is an operation used for finishing flat, cylindrical, or curved surfaces. The *lap* (Fig. 26.29a) is relatively soft and porous and usually is made of cast iron, copper, leather, or cloth. The abrasive 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

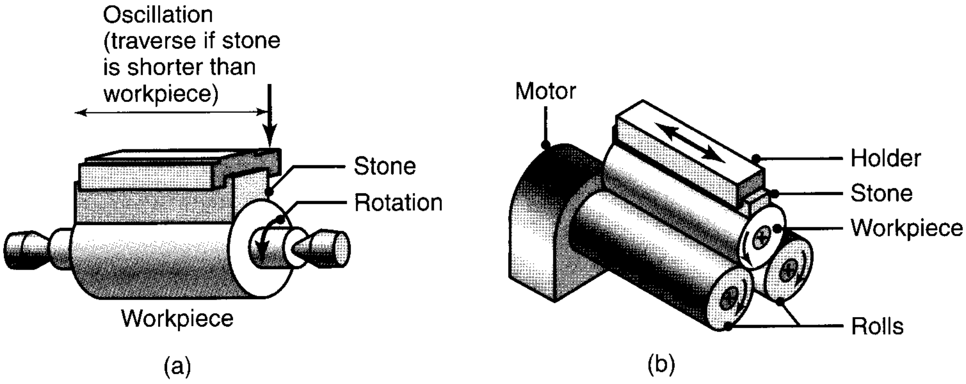


FIGURE 26.28 Schematic illustrations of the superfinishing process for a cylindrical part. (a) Cylindrical microhoning. (b) Centerless microhoning.

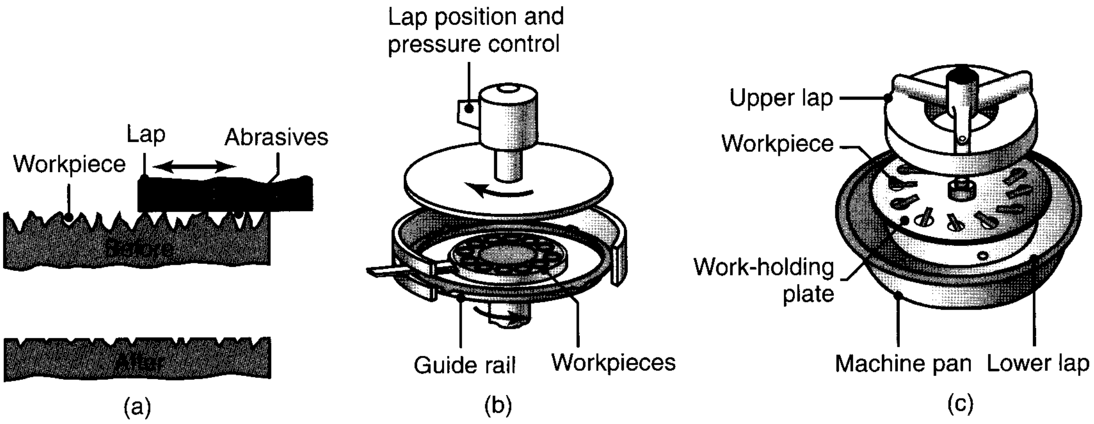


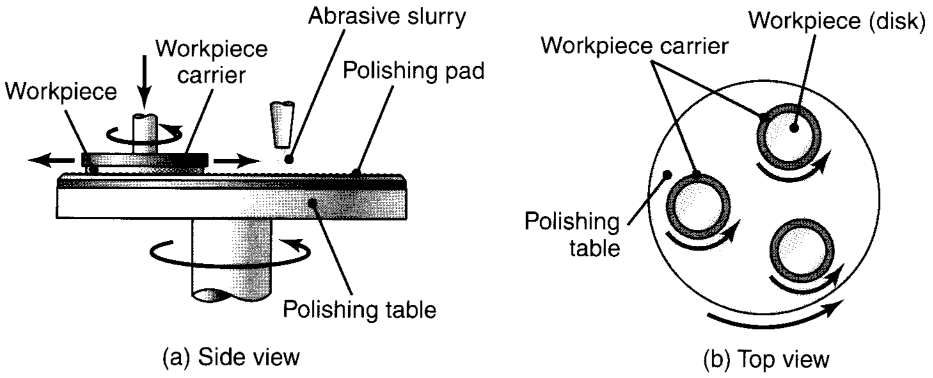
FIGURE 26.29 (a) Schematic illustration of the lapping process. (b) Production lapping on flat surfaces. (c) Production lapping on cylindrical surfaces.

on hypoid gears for rear axles. Depending on the type and hardness of the workpiece material, 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 pieces is done on machines similar to those shown in Figs. 26.29b and c.

Polishing. *Polishing* is a process that produces a smooth, lustrous surface finish. The basic mechanism involved in the polishing process is the softening and smearing of surface layers by frictional heating developed during polishing and some very fine scale abrasive removal from the workpiece surface. The shiny appearance of polished surfaces results from a smearing action.

Polishing is done with disks or belts made of fabric, leather, or felt that are typically coated with fine powders of aluminum oxide or diamond. In *double-sided polishing*, pairs of pads are fixed on the faces of platens that rotate horizontally and in opposite directions. Parts with irregular shapes, sharp corners, deep recesses, and sharp projections can be difficult to polish.



(a) Side view (b) Top view

FIGURE 26.30 Schematic illustration of the chemical–mechanical polishing process. This process is used widely in the manufacture of 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.

Chemical-mechanical Polishing. *Chemical–mechanical polishing* (CMP) is extremely important in the semiconductor industry. The process, shown in Fig. 26.30, uses a suspension of abrasive particles in a water-based solution with a chemistry selected to give controlled corrosion. The process thus removes material from the workpiece through combined abrasion and corrosion effects. The results are a surface of exceptionally fine finish and a workpiece that is especially flat. For this reason, the process is often referred to as **chemical–mechanical planarization** (see Section 28.10).

A major application of this process is the polishing of silicon wafers (Section 28.4), in which case the primary function of CMP is to polish a wafer at the microlevel. Therefore, to remove material evenly across the whole wafer, the wafer is held face down on a rotating carrier and is pressed against a polishing pad attached to a rotating disk, as shown in Fig. 26.30. The carrier and pad angular velocities are selected so that wear is uniform across the wafer surface; the angular velocities that achieve this surface result in a constant relative velocity between the carrier and pad on the axis connecting their centers. The pad contains grooves intended to uniformly supply slurry to all wafers. Also, pad rotation ensures that a linear lay does not develop (see Section 33.3).

Specific abrasive and solution-chemistry combinations have been developed for the polishing of copper, silicon, silicon dioxide, aluminum, tungsten, and other metals. For example, for silicon dioxide or silicon polishing, 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 (to be described in Section 34.9). Because there is no mechanical contact with the workpiece, this process is suitable particularly for polishing irregular shapes. The electrolyte attacks projections and peaks on the workpiece surface at a higher rate than the rest of the surface, producing a smooth surface. Electropolishing is also used for deburring operations.

Polishing in Magnetic Fields. In this technique, abrasive slurries are supported with magnetic fields. There are two basic methods:

a. In the **magnetic-float polishing** of ceramic balls, illustrated schematically in Fig. 26.31a, a magnetic fluid (containing abrasive grains and extremely fine

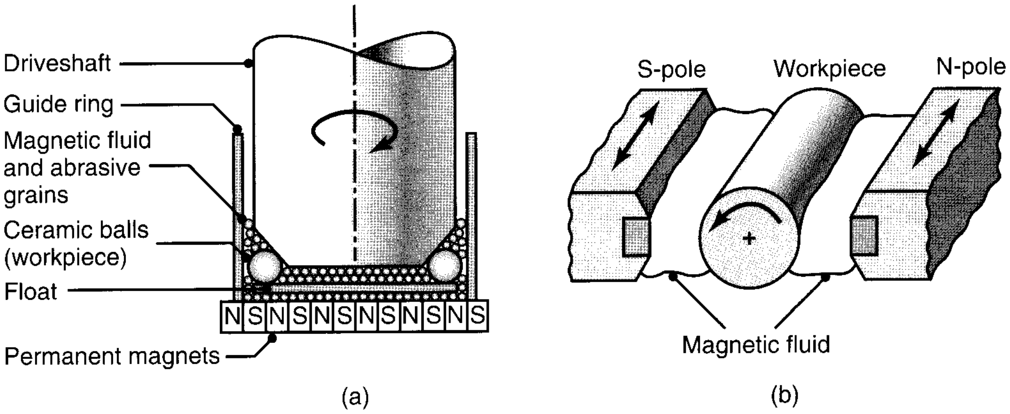


FIGURE 26.31 Schematic illustration of polishing of balls and rollers by magnetic fields. (a) Magnetic-float polishing of ceramic balls. (b) Magnetic-field-assisted polishing of rollers. *Source*: After R. Komanduri, M. Doc, and M. Fox.

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 (which is made of a nonmagnetic material) all are 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 controllable; 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.

b. **Magnetic-field-assisted polishing** of ceramic rollers is illustrated in Fig. 26.31b. A ceramic or steel roller (as the workpiece) is clamped and rotated on a spindle. The magnetic poles are oscillated, introducing a vibratory motion to the magnetic–abrasive conglomerate. This action polishes the cylindrical roller surface. Bearing steels of 63 HRC have been mirror finished in 30 seconds with this process.

Buffing. *Buffing* is similar to polishing (with the exception that an even finer surface finish is obtained) with very fine abrasives on soft disks that typically are made of cloth or hide. The abrasive is supplied externally from a stick of abrasive compound.

26.8 Deburring Operations

Burrs are thin ridges, usually triangular in shape, that develop along the edges of a workpiece from operations such as machining, shearing sheet metals (as in Figs. 16.2 and 16.3), and trimming forgings and castings. Burrs can be detected by simple means, such as with a finger, toothpick, or cotton swab. Visual inspection of burrs includes the use of magnifiers and microscopes. Although efforts at defining *burr* are being made, there are no widely accepted standards for such a definition, partly because there a variety of burrs can develop on parts.

On the one hand, burrs have several disadvantages: (a) They may interfere with the mechanical assembly of parts and can cause jamming and misalignment of parts, as well as short circuits in electrical components. (b) Because they are usually sharp, they can be a safety hazard to personnel. (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 on thin drilled or tapped components (such as the tiny parts in mechanical watches and mechanisms) can provide additional thickness and thus improve the holding torque of screws.

Several **deburring processes** are available. Their cost-effectiveness depends on factors such as the extent of deburring required, part complexity and 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 with files and scrapers. It is estimated that manual deburring can contribute up to 10% of the cost of manufacturing the part.
2. Mechanical deburring by machining pieces such as cylindrical parts on a rotating spindle.
3. Wire brushing or using rotary nylon brushes consisting of filaments embedded with abrasive grits.
4. Using 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 machining 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 workpieces. In this batch-type operation, 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 either *vibrated* or *tumbled* by various mechanical 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 purposes such as decreasing and adding corrosion resistance to the part being deburred. When chemically active fluids and abrasives are used, this is a form of chemical–mechanical polishing.

Shot Blasting. Also called **grit blasting**, this process involves abrasive particles (usually sand) that are propelled by a high-velocity jet of air, or by a rotating wheel, onto the surface of the workpiece. Shot blasting is particularly useful in deburring metallic and nonmetallic materials, and in stripping, cleaning, and removing surface oxides. The surfaces produced have a matte finish, but surface damage can result if the process parameters are not properly controlled. **Microabrasive blasting** consists of small-scale polishing and etching, using very fine abrasives, on bench-type units.

Abrasive-flow Machining. This process involves the use of abrasive grains, such as silicon carbide or diamond, that are mixed in a putty-like matrix and then forced back and forth through the openings and passageways in the workpiece. The movement of the abrasive matrix under pressure erodes away both burrs and sharp corners and polishes 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. Pressures applied range 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 to be deburred. The deburring of a turbine impeller by abrasive-flow machining is illustrated in Fig. 26.32.

In **microabrasive-flow machining**, a process mechanics similar to that used in ordinary abrasive-flow machining is used, but with much smaller abrasive media

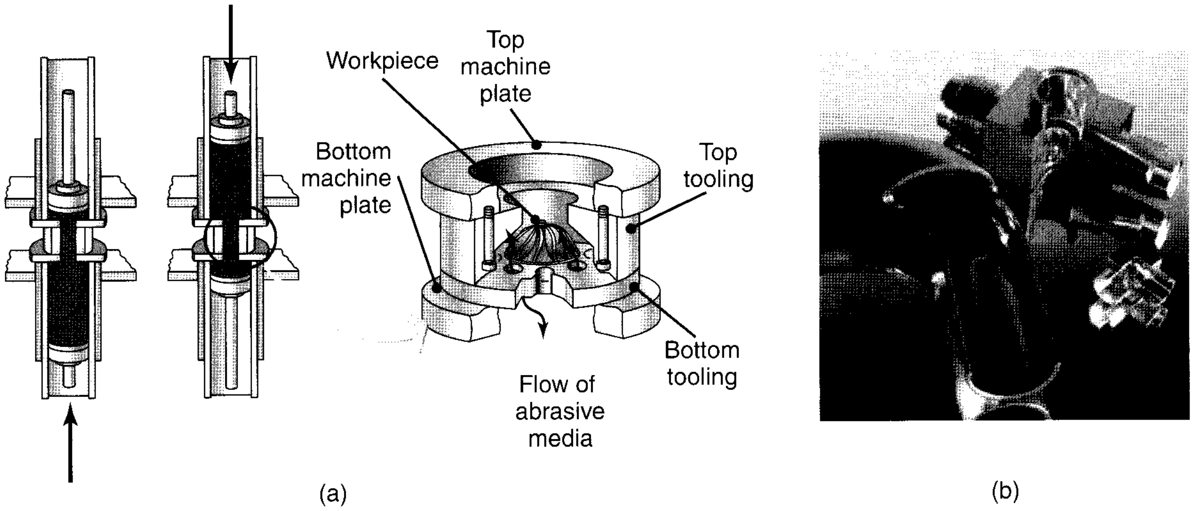


FIGURE 26.32 (a) Schematic illustration of abrasive-flow machining to de-burr a turbine impeller. The arrows indicate movement of the abrasive media. Note the special fixture, which is usually different for each part design. (b) Valve fittings treated by abrasive-flow machining to eliminate burrs and improve surface quality. *Source*: Courtesy of Kennametal Extrude Hone.

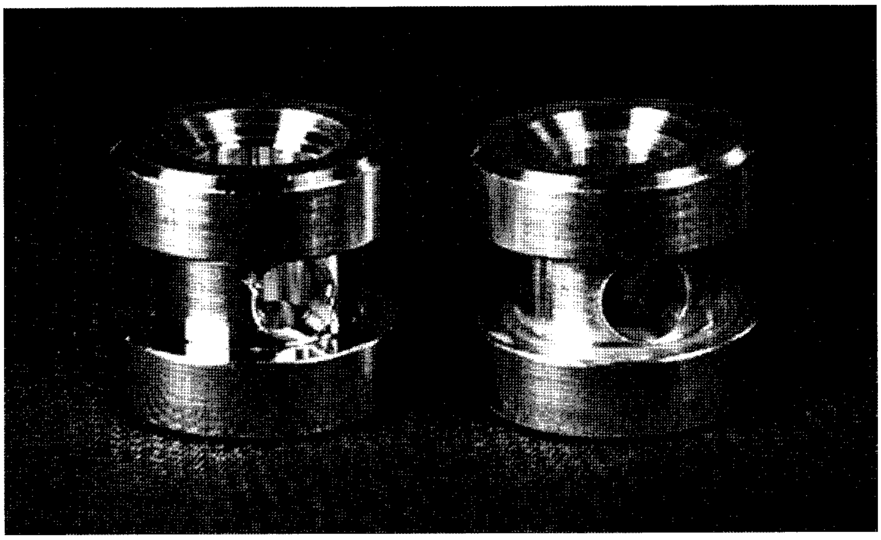


FIGURE 26.33 An example of thermal energy deburring. *Source*: Courtesy of Kennametal Extrude Hone.

and less viscous carriers. This 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 fine nozzles, where a burr or rough surface finish would adversely affect flow quality.

Thermal Energy Deburring. This process consists of placing the part in a chamber that 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 melt (see Fig. 26.33), while the temperature of the part reaches only about 150°C. However, there are drawbacks to the process: (a) Larger burrs tend to form beads after melting, (b) thin and slender parts may distort, and (c) thermal energy deburring does not polish or buff the workpiece surfaces, as do several 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 results in more consistent and repeatable deburring. An example is the robotic deburring (see also Section 37.6.3) of a die-cast outboard motor housing shown in Fig. 26.34. In another application, the manual deburring of a double-helical gear for a helicopter gearbox was deburred in 150 minutes, whereas robotic deburring required 15 minutes.

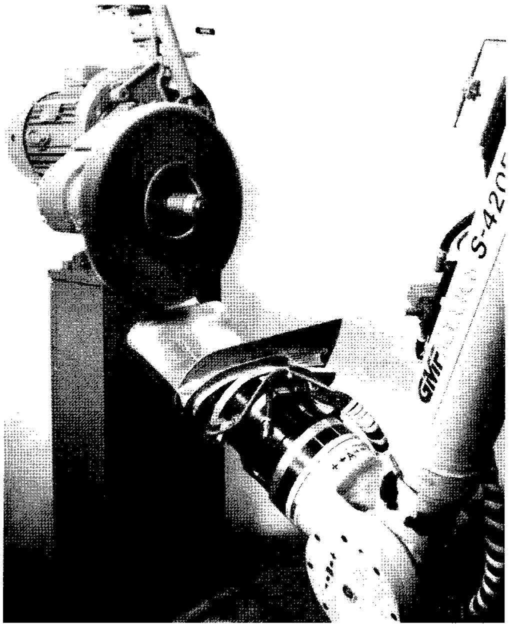


FIGURE 26.34 A deburring operation on a robot-held die-cast part for an outboard motor housing. The operation uses a grinding wheel. Abrasive belts (Fig. 26.26) or flexible abrasive radial-wheel brushes also can be used for such operations. *Source*: Courtesy of Acme Manufacturing Company.

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 high enough dimensional accuracy or good-quality surface finishes. Abrasive processes may be used both as a finishing process and as a large-scale material-removal operation. For example, creep-feed grinding is an economical alternative to machining operations (such as milling), even though wheel wear is high. Also, grinding and hard turning have become competitive for some specific applications.

Much progress has been made in automating the equipment involved in these operations, including the use of computer controls, process optimization, and robotic handling of parts. Consequently, labor costs and production times have been reduced, even though such machinery generally requires significant capital investment.

Because they are additional operations, the processes described in this chapter can significantly affect product cost, especially since these processes are relatively slow. Furthermore, as the surface-finish requirement increases, more operations are necessary; hence, the cost increases further, as can be seen in Fig. 26.35. Note how rapidly the cost increases as surface finish is improved by additional processes, such as grinding and honing.

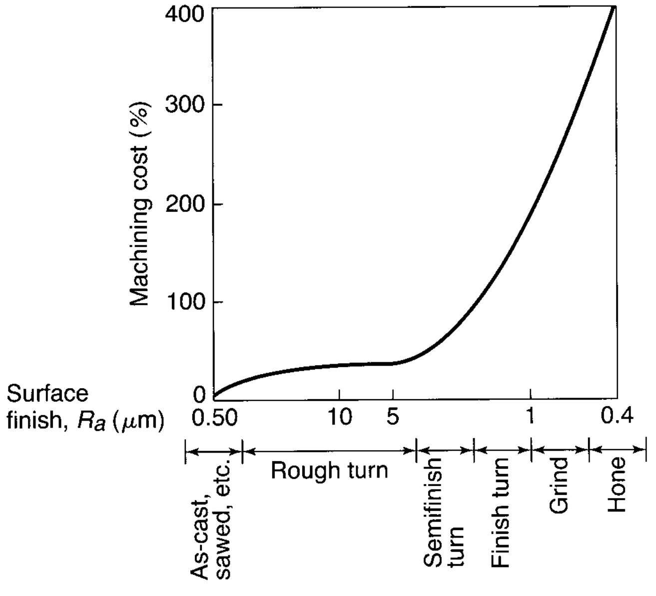


FIGURE 26.35 Increase in the cost of machining and finishing a part as a function of the surface finish required. This is the main reason that the surface finish specified on parts should not be any finer than is necessary for the part to function properly.

The total cost of abrasive operations depends on various factors, such as part size, shape, surface finish, and dimensional accuracy required, as well as machinery, tooling, and labor involved. Whereas for grinding, machinery costs can be high, the costs for other finishing processes are rather low. Grinding-wheel costs are generally low compared with other aspects of the overall grinding operation. As discussed previously, however, the cost of wheels can run into 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. 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 particular product, the conceptual and design stages should involve an analysis of the level of surface finish and the dimensional accuracy required and of whether they can be relaxed. Furthermore, all processes that precede finishing operations should be analyzed for their capability to produce more acceptable surface characteristics. As described throughout the book, this can be accomplished through the proper selection of tools, process parameters, metalworking fluids, and the characteristics of the machine tools and the work-holding devices involved.

SUMMARY

- Abrasive machining often is necessary and economical when workpiece hardness is high, the materials are brittle, and surface finish and dimensional tolerance requirements are high.
- 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 consist of a combination of abrasive grains and bonding agents. Important characteristics of wheels are abrasive grain and bond types, grade, and hardness. Wheels also may be reinforced to maintain their integrity if and when a crack develops during their use.
- Grinding-wheel wear is an important consideration in the surface quality and integrity of the ground product. Dressing and truing of wheels are important operations and are done by various techniques.
- A variety of abrasive-machining processes and machinery is available for surface, external, and internal grinding. Abrasive machining also is used for large-scale material-removal processes (such as creep-feed grinding), making it competitive with machining 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, however, 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 are also 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
Electropolishing
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 nitride
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

Brown, J., **Advanced Machining Technology Handbook**, McGraw-Hill, 1998.

Gillespie, L.K., **Deburring and Edge Finishing Handbook**, Society of Manufacturing Engineers/American Society of Mechanical Engineers, 2000.

———, **Hand Deburring—Increasing Shop Productivity**, Society of Manufacturing Engineers, 2003.

Hwa, L.S., **Chemical Mechanical Polishing in Silicon Processing**, Academic Press, 1999.

Krar, S., **Grinding Technology**, 2nd ed., Delmar Publishers, 1995.

Krar, S., and Ratterman, E., **Superabrasives: Grinding and Machining with CBN and Diamond**, McGraw-Hill, 1990.

Machinery's Handbook, Industrial Press, revised periodically.

Malkin, S., and Guo, C., **Grinding Technology: Theory and Applications of Machining with Abrasives**, 2nd ed., Industrial Press, 2008.

Marinescu, I.D. (ed.), **Handbook of Ceramics Grinding & Polishing**, Noyes, 1998.

Marinescu, I.D., Hitchiner, M., Uhlmann, E., and Rowe, W.B. (eds.), **Handbook of Machining with Grinding Wheels**, CRC Press, 2006.

Marinescu, I.D., Uhlmann, E., and Doi, T. (eds.), **Handbook of Lapping and Polishing**, CRC Press, 2006.

Maroney, M.L., **A Guide to Metal and Plastic Finishing**, Industrial Press, 1991.

Nachtman, E.S., and Kalpakjian, S., **Lubricants and Lubrication in Metalworking Operations**, Marcel Dekker, 1985.

Oliver, M.R., **Chemical Mechanical Planarization of Semiconductor Materials**, Springer, 2004.

Salmon, S.C., **Modern Grinding Process Technology**, McGraw-Hill, 1992.

Schneider, A.F., **Mechanical Deburring and Surface Finishing Technology**, Marcel Dekker, 1990.

Shaw, M.C., **Principles of Abrasive Processing**, Oxford, 1996.

Sluhan, C. (ed.), **Cutting and Grinding Fluids: Selection and Application**, Society of Manufacturing Engineers, 1992.

Szymanski, A., and Borkowski, J., **Technology of Abrasives and Abrasive Tools**, Ellis Horwood, 1992.

Tool and Manufacturing Engineers Handbook, 4th ed., Vol. 1: **Machining**, Society of Manufacturing Engineers, 1983.

Webster, J.A., Marinescu, I.D., and Trevor, T.D., **Abrasive Processes**, Marcel Dekker, 1999.

REVIEW QUESTIONS

26.1. What is an abrasive? What are superabrasives?

26.2. How is the size of an abrasive grain related to its number?

26.3. Why are most abrasives made synthetically?

26.4. Describe the structure of a grinding wheel and its features.

26.5. Explain the characteristics of each type of bond used in bonded abrasives.

26.6. What causes grinding sparks in grinding? Is it useful to observe them? Explain.

26.7. Define *metallurgical burn.*

26.8. Define (a) friability, (b) wear flat, (c) grinding ratio, (d) truing, and (e) dressing.

26.9. What is creep-feed grinding and what are its advantages?

26.10. How is centerless grinding different from cylindrical grinding?

26.11. What are the differences between coated and bonded abrasives?

26.12. What is the purpose of the slurry in chemical mechanical polishing?

QUALITATIVE PROBLEMS

26.13. Explain why grinding operations may be necessary for components that have previously been machined.

26.14. Why is there such a wide variety of types, shapes, and sizes of grinding wheels?

26.15. Explain the reasons for the large difference between the specific energies involved in machining (Table 21.2) and in grinding (Table 26.2).

26.16. The grinding ratio, G, depends on the type of grinding wheel, workpiece hardness, wheel depth of cut, wheel and workpiece speeds, and the type of grinding fluid. Explain.

26.17. What are the consequences of allowing the temperature to rise during grinding? Explain.

26.18. Explain why speeds are much higher in grinding than in machining operations.

26.19. It was stated that ultrasonic machining is best suited for hard and brittle materials. Explain.

26.20. Explain why parts with irregular shapes, sharp corners, deep recesses, and sharp projections can be difficult to polish.

26.21. List the finishing operations commonly used in manufacturing operations. Why are they necessary? Explain why they should be minimized.

26.22. Referring to the preceding chapters on processing of materials, list the operations in which burrs can develop on workpieces.

26.23. Explain the reasons that so many deburring operations have been developed over the years.

26.24. What precautions should you take when grinding with high precision? Comment on the machine, process parameters, grinding wheel, and grinding fluids.

26.25. Describe the factors involved in a grinding wheel acting "soft" or acting "hard."

26.26. What factors could contribute to chatter in grinding? Explain.

26.27. Generally, it is recommended that, in grinding hardened steels, the grinding be wheel of a relatively soft grade. Explain.

26.28. In Fig. 26.4, 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.

26.29. Describe the effects of a wear flat on the overall grinding operation.

26.30. What difficulties, if any, could you encounter in grinding thermoplastics? Thermosets? Ceramics?

26.31. Observe the cycle patterns shown in Fig. 26.20 and comment on why they follow those particular patterns.

26.32. Which of the processes described in this chapter are suitable particularly for workpieces made of (a) ceramics, (b) thermoplastics, and (c) thermosets? Why?

26.33. Grinding can produce a very fine surface finish on a workpiece. Is this finish necessarily an indication of the quality of a part? Explain.

26.34. Jewelry applications require the grinding of diamonds into desired shapes. How is this done, since diamond is the hardest material known?

26.35. List and explain factors that contribute to poor surface finish in the processes described in this chapter.

QUANTITATIVE PROBLEMS

▶**26.36.** Calculate the chip dimensions in surface grinding for the following process variables: $D = 250$ mm, $d = 0.025$ mm, $v = 30$ m/min, $V = 1500$ m/min, $C = 1$ per mm^2, and $r = 20$.

▶**26.37.** If the strength of the workpiece material is doubled, 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?

▶**26.38.** Assume that a surface-grinding operation is being carried out under the following conditions: $D = 200$ mm, $d = 0.1$ mm, $v = 40$ m/min, and $V = 2{,}000$ m/min. These conditions are then changed to the following: $D = 150$ mm, $d = 0.1$ mm, $v = 30$ m/min, and $V = 2{,}500$ m/min. How different is the temperature rise from the rise that occurs with the initial conditions?

26.39. 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.

▶**26.40.** 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?

26.41. 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.

26.42. 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.

▶**26.43.** It is known that, in grinding, heat checking occurs when grinding is done with a spindle speed of 5,000 rpm, a wheel diameter of 200 mm, and a depth of cut of 0.04 mm for a feed rate of 15 m/min. For this reason, the standard operating procedure is to keep the spindle speed at 3,500 rpm. If a new, 250-mm-diameter wheel is used, what spindle speed can be used before heat checking occurs? What spindle speed should be used to keep the same grinding temperatures as those encountered with the existing operating conditions?

▶**26.44.** A grinding operation takes place with a 10-in. grinding wheel with a spindle speed of 4,000 rpm. The workpiece feed rate is 15 m/min and the depth of cut is 0.05 mm. Contact thermometers record an approximate maximum temperature of 980°C. If the workpiece is steel, what is the temperature if the speed is increased to 5,000 rpm? What if the speed is 10,000 rpm?

26.45. Derive an expression for the angular velocity of the wafer shown in Fig. 26.30b as a function of the radius and angular velocity of the pad in chemical–mechanical polishing.

SYNTHESIS, DESIGN, AND PROJECTS

26.46. With appropriate sketches, describe the principles of various fixturing methods and devices that can be used for the processes described in this chapter.

26.47. 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.

26.48. 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.

26.49. 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.

26.50. 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.

26.51. 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. 26.6, comment on your observations, including the most common types of wheels available in the store.

26.52. 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.

26.53. 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.

26.54. Obtain pieces of sandpaper and emery cloth of different coarseness. Using a magnifier or a microscope, observe their surface features and compare them with Fig. 26.25.

26.55. 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.

Advanced Machining Processes

CHAPTER 27

EXAMPLE:

CASE STUDIES:

- It is often necessary to machine or finish products that are made of especially hard or strong materials. In that case, the conventional machining and grinding strategies described thus far become impractical.
- This chapter describes advanced machining processes that are based on nonmechanical means of material removal. The chapter begins by examining chemical machining and blanking and photochemical blanking processes, in which material is removed through the corrosive action of 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.
- High-energy beams in processes such as laser and electron beam machining have been used extensively, and these are also described, as are water-jet and abrasive-jet machining operations with unique applications.
- The chapter ends with an examination of trends in hybrid machining and the economics of advanced machining operations.

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 involved material removal by mechanical means: chip formation, abrasion, or microchipping. However, there are situations in which 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 (see Fig. 2.15).

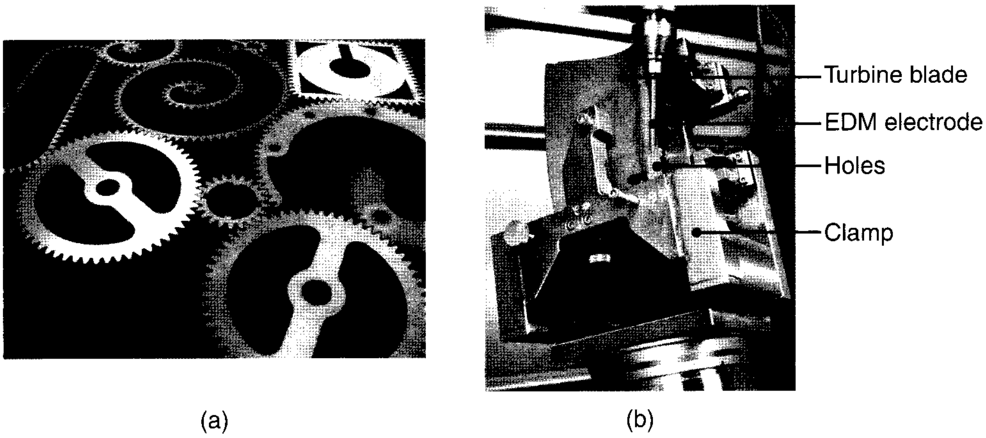


FIGURE 27.1 Examples of parts made by advanced machining processes. (a) Samples of parts produced by water-jet cutting. (b) Turbine blade, produced by plunge electrical-discharge machining, in a fixture to produce the holes also by electrical-discharge machining. *Source*: (a) Courtesy of OMAX Corporation; (b) Courtesy of HI-TEK Mfg., Inc.

- The workpiece material is too **brittle** to be machined without damage to the workpiece. This is typically the case with highly heat treated alloys, glass, ceramics, and powder-metallurgy parts.
- The workpiece is too **flexible** or too slender to withstand forces in machining or grinding, or the parts are difficult to clamp in fixtures and work-holding devices.
- The **shape** of the part is *complex* (Fig. 27.1), including such features as internal and external profiles or holes with high length-to-diameter ratios in very hard materials.
- Special **surface finish** and **dimensional tolerance** requirements exist that cannot be obtained by other manufacturing processes or are uneconomical through alternative processes.
- The **temperature rise** during processing and **residual stresses** developed in the workpiece are not desirable or acceptable.

The preceding difficulties led to the development of chemical, electrical, laser, and high-energy beams as energy sources for removing material from metallic or nonmetallic workpieces, as outlined in Table 27.1. These advanced methods, which in the past have been called *nontraditional* or *unconventional machining*, began to be introduced in the 1940s. As shall be seen, such processes remove material not by producing chips as in machining and grinding, but by means such as chemical dissolution, etching, melting, evaporation, and hydrodynamic action—sometimes 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 machining methods. This chapter describes those processes, including their characteristics, typical applications, limitations, product quality, dimensional accuracy, surface finish, and economics.

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 DC; 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
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

27.2 Chemical Machining

Chemical machining (CM) was developed from the observation that chemicals attack and etch most metals, stones, and some ceramics, thereby removing small amounts of material from the surface. The CM process is carried out by chemical dissolution using **reagents** or **etchants**, such as acids and alkaline solutions. Chemical machining is the oldest of the advanced machining processes and has been used in engraving metals and hard stones, in deburring, and in the production of printed-circuit boards and microelectronic devices (see Chapters 28 and 29).

Chemical Milling. In *chemical milling*, shallow cavities are produced on plates, sheets, forgings, and extrusions, generally for the overall reduction of weight (Fig. 27.2). The process has been used on a wide variety of metals, with depths of

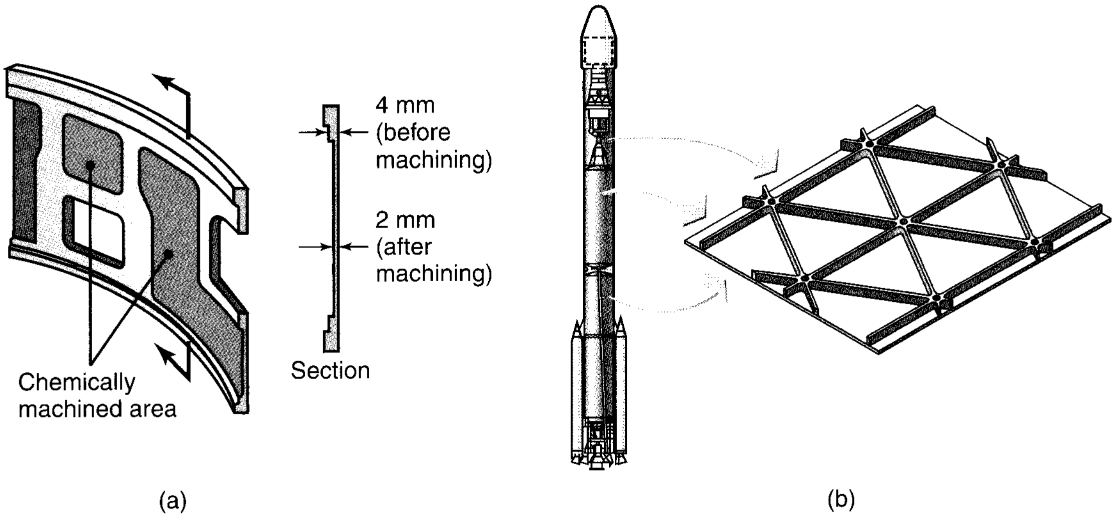


FIGURE 27.2 (a) Missile skin-panel section contoured by chemical milling to improve the stiffness-to-weight ratio of the part. (b) Weight reduction of space-launch vehicles by the chemical milling of aluminum-alloy plates. The plates are chemically milled after they have been formed into shape by a process such as roll forming or stretch forming. The design of the chemically machined rib patterns can be modified readily at minimal cost.

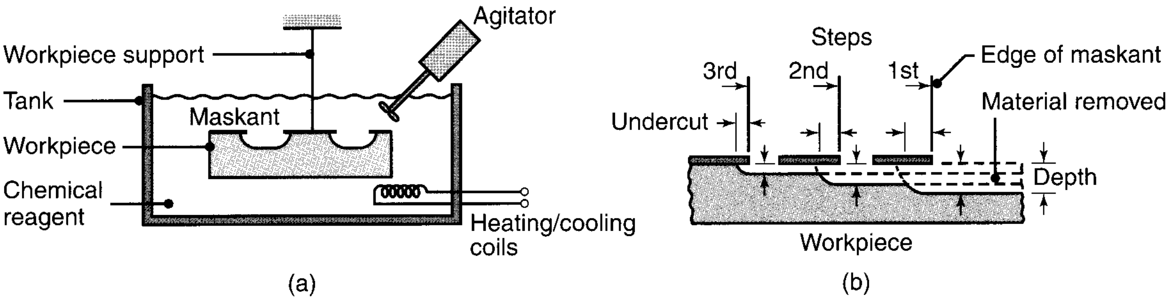


FIGURE 27.3 (a) Schematic illustration of the chemical-machining process. Note that no forces or machine tools are involved in this process. (b) Stages in producing a profiled cavity by chemical machining; note the undercut.

removal as large as 12 mm. Selective attack by the chemical reagent on different areas of the workpiece surfaces is controlled by removable layers of material (called **masking**, Fig. 27.3a) or by partial immersion in the reagent.

The procedure for chemical milling consists of the following:

1. If the part to be machined has residual stresses from prior processing, the stresses first should be relieved in order to prevent warping after chemical milling.
2. The surfaces are degreased and cleaned thoroughly to ensure both good adhesion of the masking material and uniform material removal. Scale from heat treatment also should be removed.
3. The masking material is applied. Masking with tapes or paints (**maskants**) is a common practice, although elastomers (rubber and neoprene) and plastics

(polyvinyl chloride, polyethylene, and polystyrene) also are used. The maskant material should not react with the chemical reagent.

4. If required, the maskant covering various regions that require etching is peeled off by the scribe-and-peel technique.
5. The exposed surfaces are machined chemically with 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) during chemical milling is important in order to remove a uniform depth of material.
6. After machining, the parts should be washed thoroughly to prevent further reactions with, or exposure to, any etchant residues.
7. The rest of the masking material is removed and the part is cleaned and inspected. Note that the masking material is unaffected by the reagent, but usually is dissolved easily by a different type of solvent.
8. Additional finishing operations may be performed on chemically milled parts.
9. This sequence of operations can be repeated to produce stepped cavities and various contours (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 3.7×15 m. The process is also used to fabricate microelectronic devices and often is referred to as **wet etching** for these products (see Section 28.8.1). The ranges of surface finish and tolerance obtained by chemical machining and other machining processes are shown in Fig. 27.4.

Some surface damage may result from chemical milling because of *preferential etching* and *intergranular attack*, which adversely affect surface properties. The chemical milling of welded and brazed structures may result in uneven material removal. The chemical milling of castings may result in uneven surfaces caused by porosity in, and nonuniformity of, the material.

Chemical Blanking. *Chemical blanking* is similar to the blanking of sheet metals in that it is used to produce features which penetrate through the thickness of the material (as shown in Fig. 16.4), with the exception that the material is removed by chemical dissolution rather than by shearing. Typical applications for chemical blanking are the burr-free etching of printed-circuit boards, decorative panels, and thin sheet-metal stampings, as well as the production of complex or small shapes.

Photochemical Blanking. *Photochemical blanking* (also called *photoetching*) is a modification of chemical milling. Material is removed (usually from a flat, thin sheet) by photographic techniques. Complex, burr-free shapes can be blanked (Fig. 27.5) on metals as thin as 0.0025 mm. Sometimes called **photochemical machining**, this process is also used for etching.

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×. A photographic negative is then made and reduced to the size of the finished part. The reduced negative of the design is called **artwork**. The original (enlarged) drawing allows inherent design errors to be reduced by the amount of reduction (such as 100×) for the final artwork image.
2. The sheet blank is coated with a photosensitive material (**photoresist**) by dipping, spraying, spin casting, or roller coating and is then dried in an oven. This coating is often called the *emulsion*.

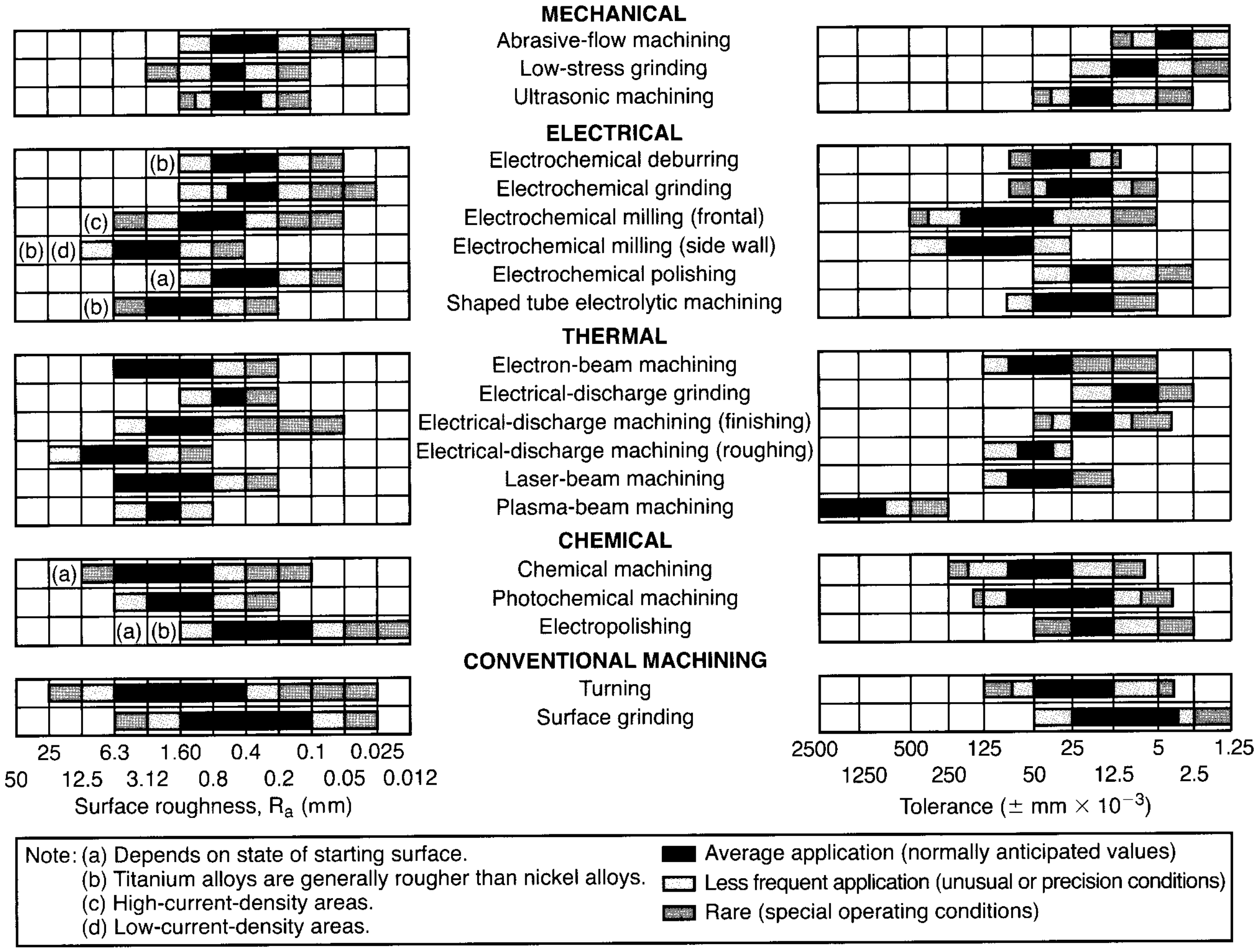


FIGURE 27.4 Surface roughness and tolerances obtained in various machining processes. Note the wide range within each process (see also Fig. 23.13). *Source*: *Machining Data Handbook*, 3rd ed. Copyright © 1980. Used by permission of Metcut Research Associates, Inc.

3. The negative is placed over the coated blank and exposed to ultraviolet light, which hardens the exposed areas.
4. The blank is developed, dissolving the unexposed areas.
5. The blank is then immersed into a bath of reagent (as in chemical milling) or sprayed with the reagent, which etches away the exposed areas.
6. The masking material is removed, and the part is washed thoroughly to remove all chemical residues.

Typical applications for photochemical blanking are fine screens, printed-circuit cards, electric-motor laminations, flat springs, and assorted components for miniaturized systems. Although skilled labor is required, tooling costs are low, the process can be automated, and it is economical for medium- to high-production volume.

Photochemical blanking is capable of making very small parts when traditional blanking dies (Section 16.2) are difficult to produce. The process is also effective for blanking fragile workpieces and materials.

The 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.

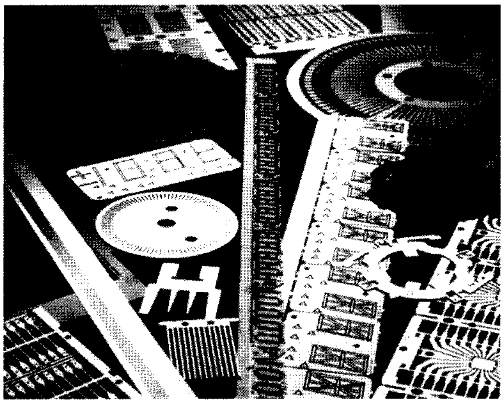


FIGURE 27.5 Various parts made by chemical blanking. Note the fine detail. *Source*: Courtesy of Buckbee-Mears, St. Paul, Minnesota.

Design Considerations for Chemical Machining. Design guidelines for chemical machining are as follows:

- Because the etchant attacks all exposed surfaces continuously, designs involving sharp corners, deep and narrow cavities, severe tapers, folded seams, or porous workpiece materials should be avoided.
- Because the etchant attacks the material in both vertical and horizontal directions, *undercuts* may develop (as shown by the areas under the edges of the maskant in Fig. 27.3). Typically, tolerances of ±10% of the material thickness can be maintained in chemical blanking.
- In order to improve the production rate, the bulk of the workpiece should be shaped by other processes (such as machining) prior to chemical machining.
- Dimensional variations can occur because of size changes in deposited mask pattern (artwork) due to humidity and temperature. These variations can be minimized by properly selecting artwork media and by controlling both the environment in which the artwork is generated and the production area in the plant.
- Many product designs are now made with computer-aided design systems (Chapter 38). However, product drawings must be translated into a protocol that is compatible with the equipment that photochemically generates artwork.

27.3 Electrochemical Machining

Electrochemical machining (ECM) is basically the reverse of electroplating (Section 34.9). An **electrolyte** acts as current carrier (Fig. 27.6), and the high rate of electrolyte movement 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 female mating image of the tool shape.

The shaped tool, either a solid or tubular form, 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. It is pumped through the passages in the tool at rates of 10 to 16 m/s. A DC power 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.

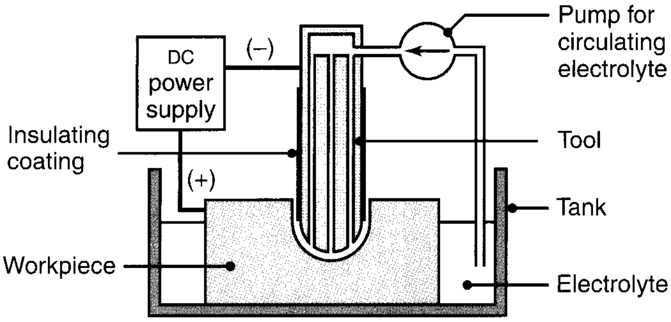


FIGURE 27.6 Schematic illustration of the electrochemical machining process.

The *material-removal rate* (MRR) in electrochemical machining for a current efficiency of 100% may be estimated from the equation

$$\text{MRR} = CI \tag{27.1}$$

where MRR is in mm^3/min, I is the current in amperes, and C is a material constant, in $\text{mm}^3/\text{A-min}$. For pure metals, C depends on the valence: The higher the valence, the lower is its 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. 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.

Process Capabilities. The basic concept of electrochemical machining was patented in 1929 and developed rapidly during the 1950–1960s, whereupon it became an important manufacturing process, although not as widely used as other processes described in this chapter. 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. 27.7), as well as in the automotive (engines castings and gears) and medical industries.

Electrochemical machining is also used for machining and finishing forging-die cavities (*die sinking*) and to produce small holes. Versions of this process are used for turning, facing, milling, slotting, drilling, trepanning, and profiling, as well as in the production of continuous metal strips and webs. More recent applications of ECM include *micromachining* (Chapters 28 and 29) for the electronics industry.

A modification of ECM is *shaped-tube electrolytic machining* (STEM), used for drilling small-diameter deep holes, as in turbine blades. The tool is a titanium tube, coated with an electrically insulating resin. Holes as small as 0.5 mm can be drilled, at depth-to-diameter ratios of as high as 300:1.

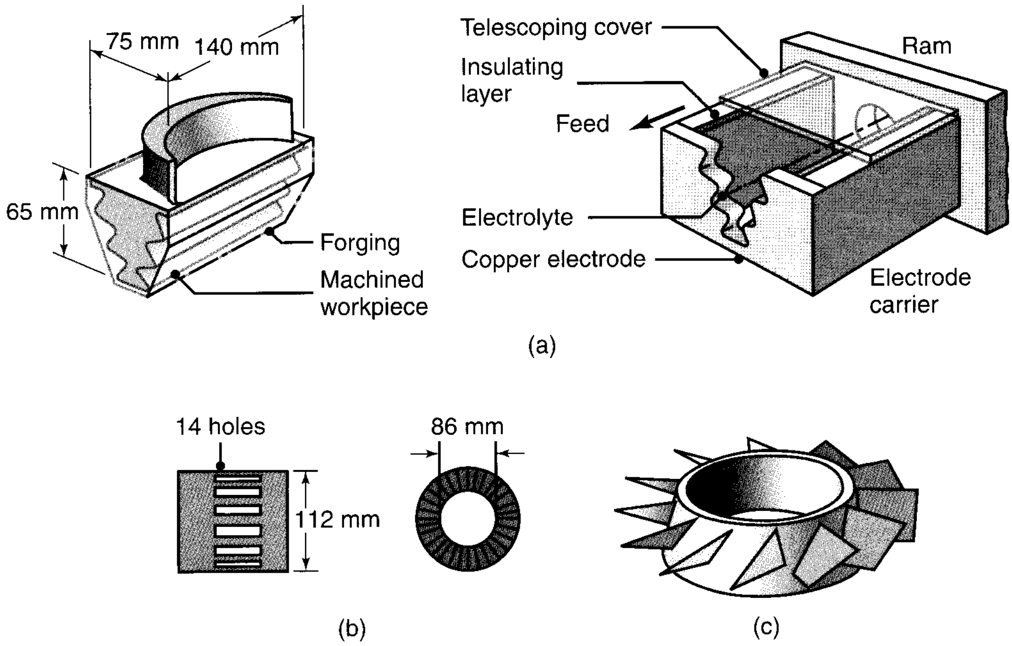


FIGURE 27.7 Typical parts made by electrochemical machining. (a) Turbine blade made of a nickel alloy of 360 HB. Note the shape of the electrode on the right. (b) Thin slots on a 4340-steel roller-bearing cage. (c) Integral airfoils on a compressor disk.

The ECM process leaves a burr-free, bright surface; in fact, it can also be used as a deburring process. It does not cause any thermal damage to the part, and the absence of tool forces prevents part distortion. 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 the ECM process should be compared carefully with those of components made by other material-removal methods.

Electrochemical-machining systems now are available as *numerically controlled machining centers* with the capability of high production rates, high flexibility, and the maintenance of close dimensional tolerances. The ECM process also can be combined with electrical-discharge machining (EDM) on the same machine (*hybrid machining*).

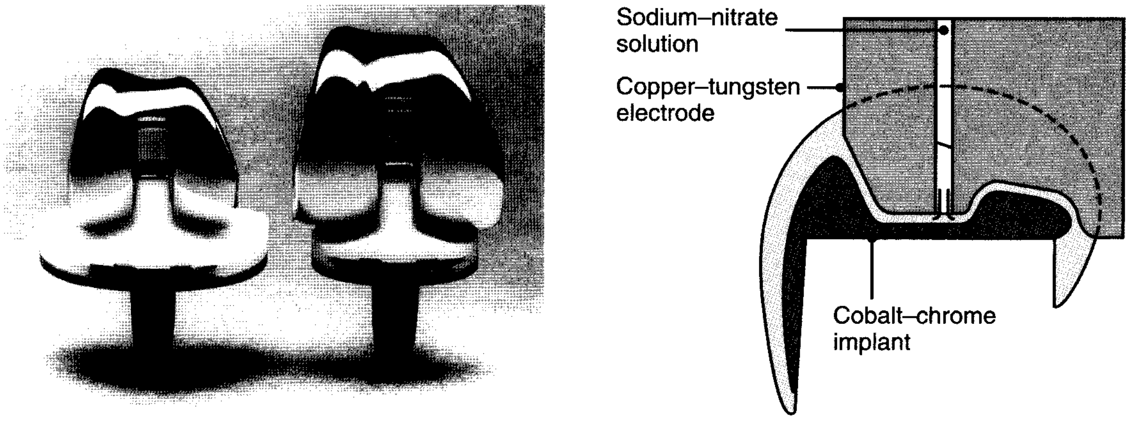
Design Considerations for Electrochemical Machining. The following are general design guidelines for electrochemical machining:

- Because of the tendency for the electrolyte to erode away sharp profiles, electrochemical machining is not suited for producing sharp square corners or flat bottoms.
- Controlling the electrolyte flow may be difficult, so irregular cavities may not be produced to the desired shape with acceptable dimensional accuracy.
- Designs should make provision for a small taper for holes and cavities to be machined.

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, as shown in Fig. 27.8a. The 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 been concerned particularly with the contact surface in the cavity of the



(a)

(b)

FIGURE 27.8 (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 Biomet, Inc.

metal implant that mates with a protrusion on the polyethylene insert. As the knee articulates during normal motion, the polyethylene slides against the metal part, becoming a potentially serious wear site. This geometry is necessary to ensure lateral stability of the knee (that is, to prevent the knee from buckling sideways).

In order to produce a smooth surface, the 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 quality. The interior surfaces of this part are extremely difficult to access for grinding, and the cobalt–chrome alloy is difficult to grind. Consequently, advanced machining processes (particularly electrochemical machining) were ideal candidates for this operation.

As shown in Fig. 27.8b, the current procedure consists of placing the metal implant in a fixture and bringing a tungsten electrode of the desired final contour in close proximity to the implant. The electrolyte is a sodium nitrate and water mixture and is pumped through the tool, filling the gap between the tool and 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 volume can be controlled to maximize surface quality. When the flow rate is too low, defects appear on the machined surface as localized dimples; if the rate is too high, machining times become longer. Typical machining times for this part are four to six minutes.

Source: Courtesy of T. Hershberger and R. Redman, Biomet, Inc., Warsaw, Indiana.

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 (die sinking). Investigations have shown that PECM improves fatigue life, and the process has been proposed as a possible method for eliminating the recast layer left on die and mold surfaces by electrical-discharge machining. The tolerances obtained typically are in the range from 20 to 100 μm.

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. However, it is difficult to maintain precise alignment when moving the workpiece from the EDM to the PECM. If misaligned significantly, all polishing will occur where the gap is smallest, and passivation will occur where the gap is largest. Also, the process leaves metal residues suspended in the aqueous solution, which is harmful to the environment if discarded without treatment.

The ECM process can be effective for micromachining as well. The complete absence of tool wear implies that the process can be used for precision-electronic components, although the erosion problem due to stray current has to be overcome. ECM machines now have increasing flexibility by the implementation of numerical controls.

27.4 Electrochemical Grinding

Electrochemical grinding (ECG) combines electrochemical machining with conventional grinding. The equipment used is similar to a conventional grinder, except that the wheel is a rotating cathode embedded with abrasive particles (Fig. 27.9a). The wheel is metal bonded with diamond or aluminum-oxide abrasives and rotates at a surface speed from 1,200 to 2,000 m/min.

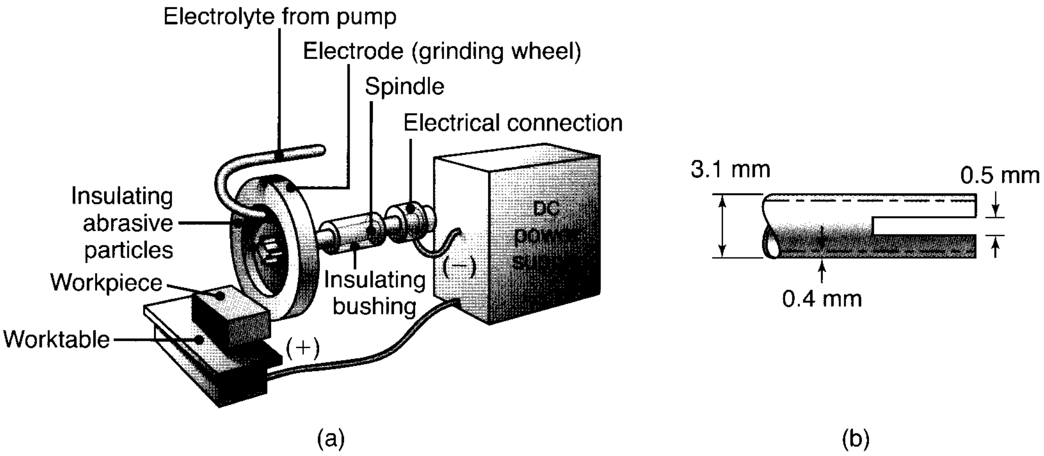


FIGURE 27.9 (a) Schematic illustration of the electrochemical-grinding process. (b) Thin slot produced on a round nickel-alloy (Inconel) tube by this process.

The abrasives have two functions: (1) to serve as insulators between the wheel and the workpiece and (2) to 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. 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. Therefore, wheel wear is very low, and the workpiece 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. 27.9b), but it is not adaptable to cavity-sinking operations. EGC can be applied successfully to carbides and high-strength alloys. It offers a distinct advantage over traditional diamond-wheel grinding in processing very hard materials where wheel wear can be high. ECG machines are available with numerical controls, which improve dimensional accuracy and repeatability and increase productivity.

Electrochemical honing combines the fine abrasive action of honing 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.

Design Considerations for Electrochemical Grinding. In addition to those design considerations already listed for electrochemical machining, ECG requires two more:

- Designs should avoid sharp inside radii.
- If flat surfaces are to be produced, the electrochemically ground surface 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. We know that when two current-conducting wires are allowed to touch each other, an arc is produced. If we look closely at the point of contact between the two wires, we note that a small portion of the metal has been eroded away, leaving a small crater.

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 has become 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 workpiece, connected to a DC power supply and placed in a **dielectric** (electrically nonconducting) fluid (Fig. 27.10a). When the potential difference between the tool and the workpiece 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 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. The melting point and the latent heat of melting are important physical properties that determine the volume of metal removed per discharge. As these quantities

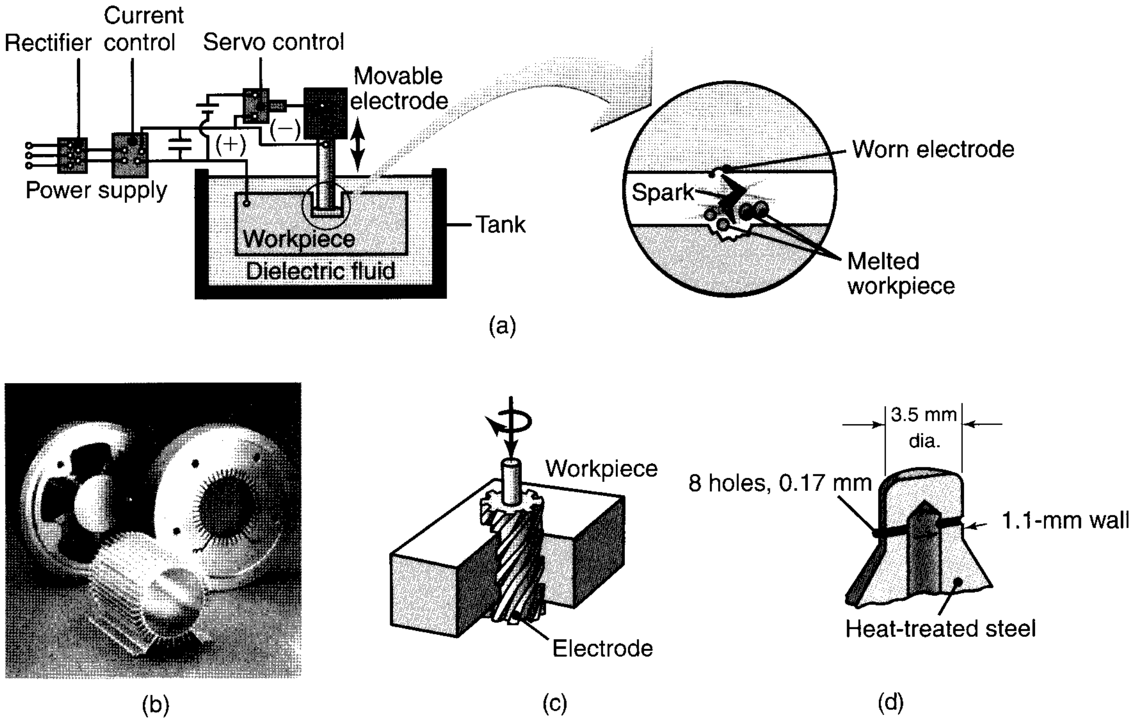


FIGURE 27.10 (a) Schematic illustration of the electrical-discharge machining process. This is one of the most widely used machining processes, particularly for die-sinking applications. (b) Examples of cavities produced by the electrical-discharge machining process, using shaped electrodes. The two 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.

increase, the rate of material removal decreases. The material-removal rate can be estimated from the approximate empirical formula

$$\mathrm{MRR} = 4 \times 10^4\, I T_w^{-1.23}, \tag{27.2}$$

where MRR is in $\mathrm{mm}^3/\mathrm{min}$, I is the current in amperes, and T_w is the melting point of the workpiece in °C.

The workpiece is fixtured within the 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. Thus, the downward feed of the tool is controlled by a servomechanism, which automatically maintains a constant gap.

Because the process doesn't involve mechanical energy, the hardness, strength, and toughness of the workpiece material do not necessarily influence the removal rate. The frequency of discharge or the energy per discharge, the voltage, and the current usually are varied to control the removal rate. The removal rate and surface roughness increase with (a) increasing current density and (b) 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.

The most common dielectric fluids are mineral oils, although kerosene and distilled and deionized water also are used in specialized applications. Clear, low-viscosity fluids also are available; although more expensive, these fluids make cleaning easier. The machines are equipped with a pump and filtering system for the dielectric fluid.

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 CNC machining techniques. Tungtsen-wire electrodes as small as 0.1 mm in diameter have been used to produce holes with depth-to-hole diameter ratios of up to 400:1.

The sparks in this process also erode away the electrode, changing its geometry and adversely affecting the shape produced and its dimensional accuracy. *Tool (electrode) wear* is thus an important factor. *Wear ratio* is defined as the ratio of the volume of workpiece material removed to the volume of tool wear. This ratio ranges from about 3:1 for metallic electrodes to as high as 100:1 for graphite electrodes.

It has been shown that tool wear is related to the melting points of the materials involved: The lower the melting point of the electrode, the higher is the wear rate. Also, the higher the current, the higher is the wear. Consequently, graphite electrodes have the highest wear resistance. Tool wear can be minimized by reversing the polarity and using copper tools—a process called **no-wear EDM.**

Process Capabilities. Electrical-discharge machining has numerous applications, such as the production of 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 deep, small-diameter holes with tungsten wire used as the electrode, narrow slots in parts, cooling holes in superalloy turbine blades, and various intricate shapes (see

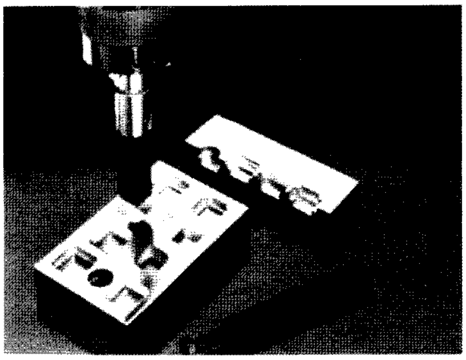


FIGURE 27.11 Stepped cavities produced with a square electrode by the EDM process. The workpiece moves in the two principal horizontal directions, and its motion is synchronized with the downward movement of the electrode to produce these cavities. Also shown is a round electrode capable of producing round or elliptical cavities. *Source*: Courtesy of AGIE USA, Ltd.

Figs. 27.10b and c). Stepped cavities can be produced by controlling the relative movements of the workpiece in relation to the electrode (Fig. 27.11).

Because of the molten and resolidified (recast) surface structure, high rates of material removal produce a very rough surface finish with poor surface integrity and low fatigue properties. Therefore, finishing cuts are made at low removal rates, or the recast layer is removed subsequently by finishing operations. It has been shown that 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 as follows:

- Parts should be designed so that the required electrodes can be shaped properly and 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 a high production rate, the bulk of material removal should be done by conventional processes (roughing out).

27.5.1 Wire EDM

A variation of EDM is *wire EDM* (Fig. 27.12), or *electrical-discharge wire cutting*. In this process, which is similar to contour cutting with a band saw (see Fig. 24.25), a slowly moving wire travels along a prescribed path, cutting the workpiece. This process is used to cut plates as thick as 300 mm and to make punches, tools, and dies from hard metals. It also can cut intricate components for the electronics industry. Figure 27.13a shows a thick plate being cut by this process, and a wire EDM machine is shown in Fig. 27.13b.

The wire is usually made of brass, copper, tungsten, or molybdenum; zinc- or brass-coated and multicoated 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.

The wire is usually used only once, as it is relatively inexpensive compared with the type of operation it performs. It 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 trend in the use of dielectric fluids is toward clear, low-viscosity fluids.

The cutting speed generally is given in terms of the cross-sectional area cut per unit time. Typical examples are 18,000 mm^2/hr for 50-mm thick

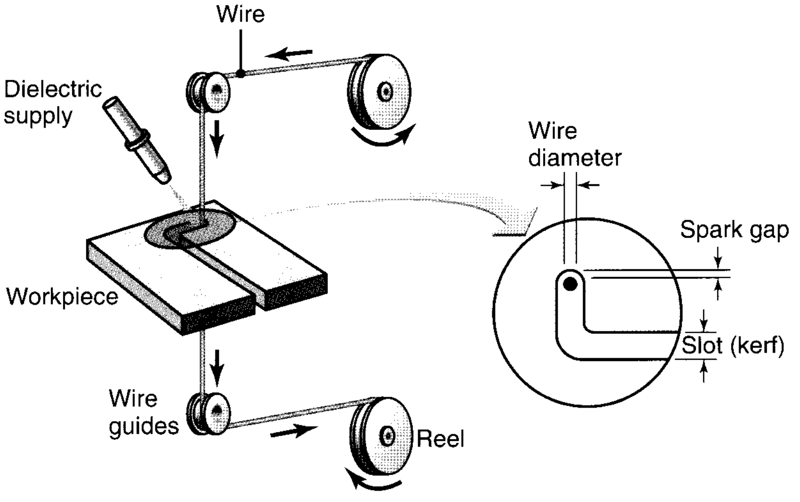


FIGURE 27.12 Schematic illustration of the wire EDM process. As many as 50 hours of machining can be performed with one reel of wire, which is then discarded.

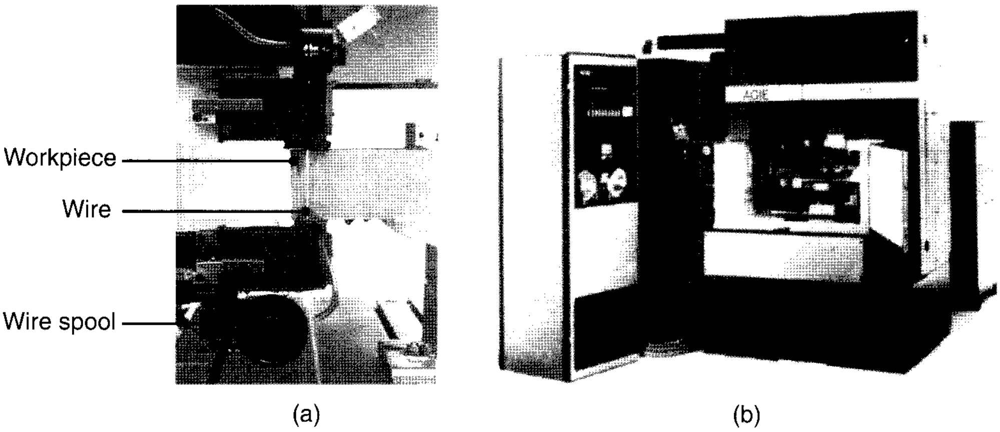


FIGURE 27.13 (a) Cutting a thick plate with wire EDM. (b) A computer-controlled wire EDM machine. *Source*: Courtesy of AGIE USA, Ltd.

D2 tool steel and 45,000 mm^2/hr for 150-mm thick aluminum. These removal rates indicate a linear cutting speed of 18,000/50 = 360 mm/hr = 6 mm/min and 45,000/150 = 300 mm/hr = 5 mm/min, respectively.

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 (Fig. 27.13b) and its angle with respect to the workpiece plane.
- Multiheads for cutting two parts at the same time.
- Features such as controls for preventing wire breakage.
- Automatic self-threading features 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. Many modern wire EDM machines allow the control of the feed and take-up ends of the wire to traverse independently in two principal directions, so that tapered parts can be made.

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 rotating wheel and the workpiece. This process is used primarily for grinding carbide tools and dies but can also be used with fragile parts, such as surgical needles, thin-walled tubes, and honeycomb structures. The ECDG process is faster than EDG, but power consumption is higher. The EDG process can be combined with electrochemical grinding. The process is then called **electrochemical-discharge grinding** (ECDG).

Material is removed by chemical action (with the electrical discharges from the graphite wheel breaking up the oxide film) and is washed away by the electrolyte flow. The material-removal rate in EDG 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 mm^3/A-min. For example, $K = 4$ for tungsten carbide and $K = 16$ for steel.

In **sawing** with EDM, a setup similar to a band or circular saw (but without any teeth) is used with the same electrical circuit for EDM. Narrow cuts can be made at high rates of metal removal. Because cutting forces are negligible, the process can be used on thin and slender components as well.

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 surface of the workpiece (Fig. 27.14a). 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):

CO_2 (*pulsed* or *continuous wave*)
Nd:YAG (neodymium: yttrium–aluminum–garnet)

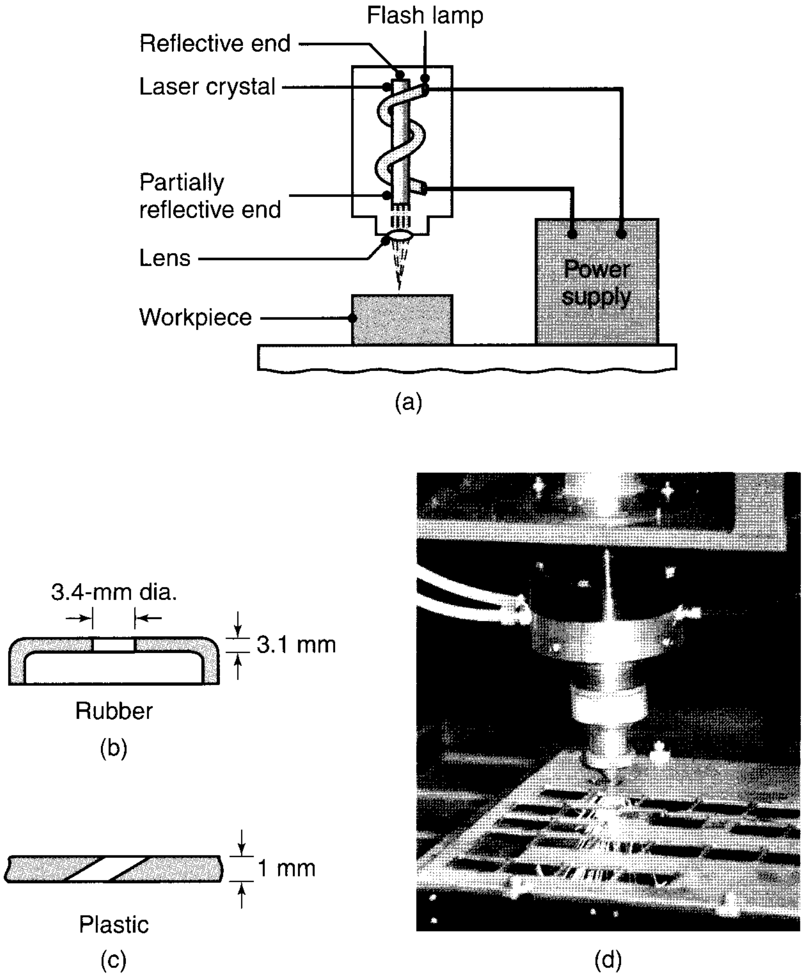


FIGURE 27.14 (a) Schematic illustration of the laser-beam machining process. (b) and (c) Examples of holes produced in nonmetallic parts by LBM. (d) Cutting sheet metal with a laser beam. *Source*: (d) Courtesy of Rofin-Sinar, Inc.

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

Note: P = pulsed, CW = continuous wave, Nd:YAG = neodymium:yttrium–aluminum–garnet.

Nd:glass, ruby
Diode lasers
Excimer lasers (from the words *exci*ted 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. 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 rough and has a *heat-affected zone* (as described in Section 30.9), which, in critical applications, may have to be removed or heat treated. Kerf width is an important consideration, as it is in other cutting processes, such as sawing, wire EDM, and electron-beam machining.

Laser beams may be used in combination with a gas stream (such as oxygen) to increase energy absorption (**laser-beam torch**) for cutting sheet metals. High-pressure, inert-gas (nitrogen or argon) assisted laser cutting is used for stainless steel and aluminum; it leaves an oxide-free edge that can improve weldability. Gas streams also have the important function of blowing away molten and vaporized material from the workpiece surface.

Process Capabilities. Laser-beam machining is used widely for drilling, trepanning, and cutting metals, nonmetallic materials, ceramics, and composite materials (Figs. 27.14b and c). The abrasive nature of composite materials and the cleanliness of the operation have 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 produced in various materials, although a more practical minimum is 0.025 mm. Steel plates as thick as 32 mm can be cut with laser beams.

Laser-beam machining is being used increasingly in the electronics and automotive industries. Bleeder holes for fuel-pump covers and lubrication holes in transmission hubs are being drilled with lasers, for example. The cooling holes in the first stage of producing vanes for the Boeing 747 jet engines also are produced by lasers. Significant cost savings have been achieved by laser-beam machining—a process that is competing with electrical-discharge machining.

Laser beams are also used for the following applications:

- **Welding**
- Small-scale and localized **heat treating** of metals and ceramics to modify their surface mechanical and tribological properties
- **Marking** of parts, such as letters, numbers, codes, etc. Marking also can be done (a) with ink, (b) with mechanical devices such as punches, pins, styluses, and scroll rolls, (c) by stamping, and (d) by etching. Although the equipment is more expensive than that used in other methods, marking and engraving with lasers has become increasingly 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, and low setup times, as well as the availability of multi-kW machines and two- and three-dimensional computer-controlled robotic laser-cutting systems—is an attractive feature. Therefore, laser cutting for sheets can compete successfully with the traditional punching processes described in Chapter 16. There are efforts now to combine the two processes for improved overall efficiency. (See Section 27.10 and Example 27.1.)

Extreme caution should be exercised with lasers. Even low-power lasers can cause damage to the retina of the eye if proper safety precautions are not observed.

Design Considerations for LBM. General design guidelines for laser-beam machining are as follows:

- Designs with sharp corners should be avoided, since they can be difficult to produce.
- Deep cuts will produce tapered walls.
- Reflectivity of the workpiece surface is an important consideration in laser-beam machining; because they reflect less, dull and unpolished surfaces are preferable.
- Any adverse effects on the properties of the machined materials caused by the high local temperatures and heat-affected zone should be investigated.

EXAMPLE 27.1 Combining Laser Cutting and Punching of Sheet Metal

As has been seen, 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 cutting are generally (a) smaller batches, (b) the flexibility of the operation, (c) a wide range of thicknesses, (d) prototyping capability, (e) materials and composites that otherwise might be cut with difficulty, and (f) complex geometries that can be programmed.

Advantages and drawbacks of punching include (a) required large lot sizes in order to economically justify purchasing tooling, (b) relatively simple parts, (c) a small range of part thicknesses, (d) fixed and limited punch geometries (even when using turrets), (e) rapid production, and (f) integration with subsequent processing after punching.

It is evident that 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 to their full extent, but 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 the machining centers described in Section 25.2. For example, turret-punch presses have been equipped with an integrated laser head; the machine can punch or laser cut, but it cannot do both simultaneously.

Several factors have to be taken into account in such a combination with respect to the characteristics of each operation: (1) the ranges of sizes, thicknesses, and shapes to be produced and how they are to be nested; (2) processing and setup times, including the loading, fixturing, and unloading of parts; (3) programming for cutting; and (4) the process capabilities of each method, including dynamic characteristics, vibrations, and shock from punching (and isolation) that may disturb adjustments and alignments of the laser components.

27.7 Electron-beam Machining

The energy source in *electron-beam machining* (EBM) is high-velocity electrons, which strike the workpiece surface and generate heat (Fig. 27.15). The machines utilize voltages in the range from 50 to 200 kV to accelerate the electrons to speeds of 50 to 80% of the speed of light. Applications of this process are similar to those of laser-beam machining, except that EBM requires a vacuum. Consequently, it is used much less than laser-beam machining.

Electron-beam machining can be used for very accurate cutting of a wide variety of metals. Surface finish is better and kerf width is narrower than in other thermal cutting processes. (See also Section 30.6 on *electron-beam welding*.) The interaction of the electron beam with the workpiece surface produces hazardous X-rays. Therefore, the equipment should be used only by highly trained personnel.

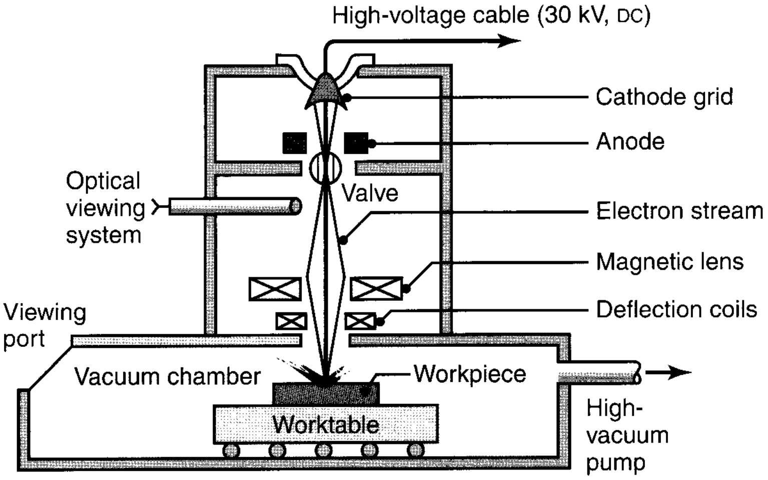


FIGURE 27.15 Schematic illustration of the electron-beam machining process. Unlike LBM, this process requires a vacuum, so the workpiece size is limited to the size of the vacuum chamber.

Design Considerations for EBM. The guidelines for EBM generally are similar to those for LBM; additional considerations are as follows:

- Because vacuum chambers have limited capacity, individual parts or batches should closely match the size of the vacuum chamber for a high production rate per cycle.
- If a part requires electron-beam machining on only a small portion of its volume, consideration should be given to manufacturing it as a number of smaller components and assembling them after electron-beam machining.

Plasma-arc Cutting. In plasma-arc cutting (PAC), *plasma beams* (ionized gas) are used to rapidly cut ferrous and nonferrous sheets and plates. (See also Section 30.3.) The temperatures generated are very high (9400°C in the torch for oxygen as a plasma gas). Consequently, the process is fast, the kerf width is small, and the surface finish is good. Parts as thick as 150 mm can be cut. Material-removal rates are much higher than those associated with the EDM and LBM processes, and parts can be machined with good reproducibility. Plasma-arc cutting is highly automated today, using programmable controllers.

27.8 Water-jet Machining

When we put our hand across a jet of water or air, we feel a considerable concentrated force acting on it. This force results from the momentum change of the stream and, in fact, is the principle on which the operation of water or gas turbines is based. In *water-jet machining* (WJM) (also called **hydrodynamic machining**), this force is utilized in cutting and deburring operations (Fig. 27.16). (See also *water-jet peening*, Section 34.2.)

The water jet acts like a saw and cuts a narrow groove in the material. A pressure level of about 400 MPa is generally used for efficient operation, although pressures as high as 1400 MPa can be generated. Jet-nozzle diameters range between 0.05 and 1 mm. A water-jet cutting machine and its operation are shown in Fig. 27.16b. A variety of materials can be cut, including plastics, fabrics, rubber, wood products, paper, leather, insulating materials, brick, and composite materials (Fig. 27.16c).

Depending on the materials, thicknesses can range up to 25 mm and higher. Vinyl and foam coverings for automobile dashboards (as well as some body panels) are being cut by multiple-axis, robot-guided water-jet machining equipment. Because it is an efficient and clean operation compared with other cutting processes, it is also used in the food-processing industry for cutting and slicing food products.

The advantages of this process are as follows:

- 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, the process is suitable for flexible materials.
- Little wetting of the workpiece takes place.
- The burr produced is minimal.
- It is an environmentally safe manufacturing process.

Abrasive Water-jet Machining. In *abrasive water-jet machining* (AWJM), the water jet contains abrasive particles (such as silicon carbide or aluminum oxide), which increase the material-removal rate above that of water-jet machining.

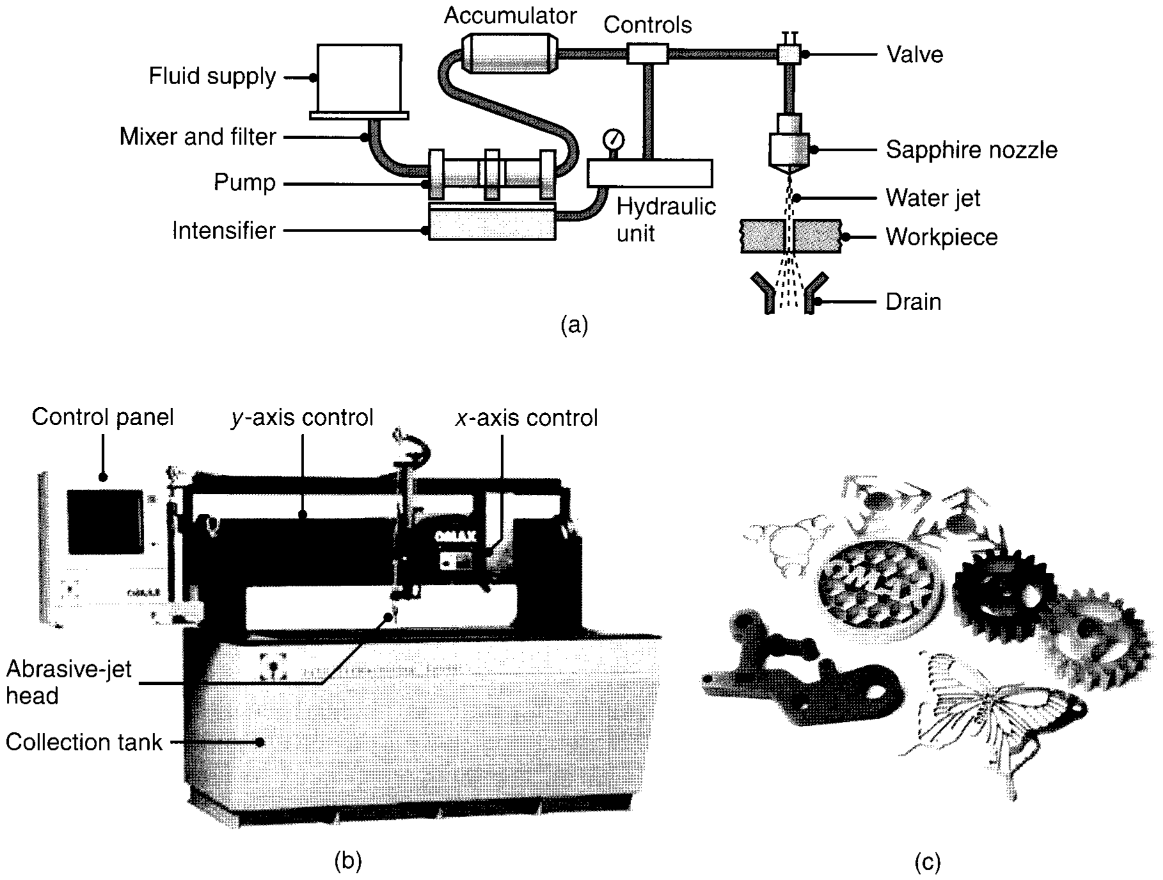


FIGURE 27.16 (a) Schematic illustration of the water-jet machining process. (b) A computer-controlled water-jet cutting machine. (c) Examples of various nonmetallic parts produced by the water-jet cutting process. *Source*: Courtesy of OMAX Corporation.

Metallic, nonmetallic, and advanced composite materials of various thicknesses can be cut in single or multilayers.

AWJM is suitable particularly 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 much lower for metals. Consequently, the process may not be acceptable for situations requiring high production rates.

The minimum hole size that can be produced satisfactorily to date is about 3 mm; maximum hole depth is on the order of 25 mm. With multiple-axis and robot-controlled machines, complex three-dimensional parts can be machined to finish dimensions. The optimum level of abrasives in the jet stream is controlled automatically in modern AWJM systems. Nozzle life has been improved by making nozzles from rubies, sapphires, and carbide-based composite materials (Fig. 27.16a).

27.9 Abrasive-jet Machining

In *abrasive-jet machining* (AJM), a high-velocity jet of dry air, nitrogen, or carbon dioxide containing abrasive particles is aimed at the workpiece surface under controlled conditions (Fig. 27.17). The impact of the particles develops a sufficiently concentrated force (see also Section 26.6) to perform operations such as (a) cutting

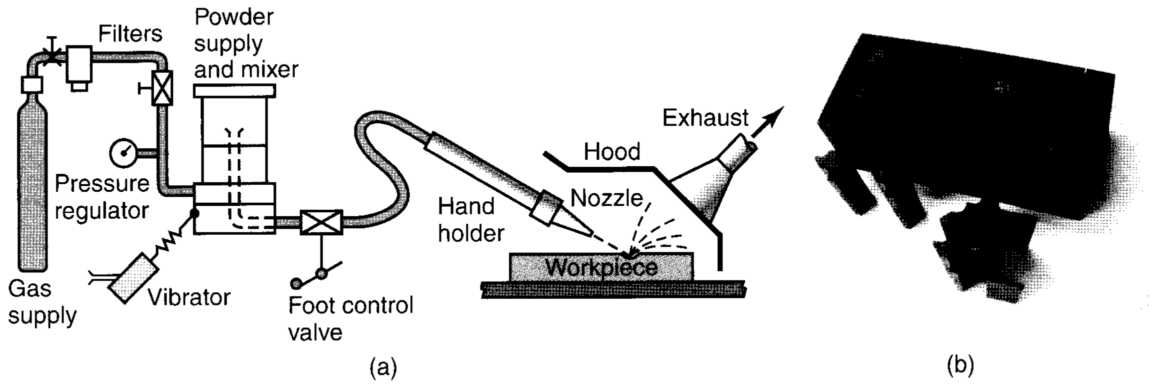


FIGURE 27.17 (a) Schematic illustration of the abrasive-jet machining process. (b) Examples of parts made by abrasive-jet machining, produced in 50-mm thick 304 stainless steel. *Source*: Courtesy of OMAX Corporation.

small holes, slots, or intricate patterns in very hard or brittle metallic and nonmetallic materials, (b) deburring or removing small flash from parts, (c) trimming and beveling, (d) removing oxides and other surface films, and (e) generally cleaning components with irregular surfaces.

The gas-supply pressure is on the order of 850 kPa, and the abrasive-jet velocity can be as high as 300 m/s and is controlled by a valve. The nozzles are usually made of tungsten carbide or sapphire, both of which have abrasive wear resistance. The abrasive 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 in metal parts tend to be tapered. There is some hazard involved in using this process, because of airborne particulates. The problem can be avoided by using the abrasive water-jet machining process.

27.10 Hybrid Machining Systems

A more recent development in manufacturing is the concept of *hybrid machining systems*. Two or more machining processes are combined into one system to take advantage of the capabilities of each process, increasing production speed and thus improving the efficiency of the operation. The system is able to handle a variety of materials, including metals, ceramics, polymers, and composites. Examples of such systems include combinations and integration of the following processes:

a. Abrasive machining and electrochemical machining
b. Abrasive machining and electrical discharge machining
c. Abrasive machining and electrochemical finishing
d. Water-jet cutting and wire EDM
e. High-speed milling, laser ablation, and blasting, as an example of *three* integrated processes.
f. Machining and blasting
g. Electrochemical and electrical discharge machining (ECDM), also called electrochemical spark machining (ECSM).

h. Machining and forming processes, such as laser cutting and punching of sheet metal, described in Example 27.1.

i. Combinations of various forming, machining, and joining processes.

The implementation of these concepts and the development of machinery and control systems present significant challenges. Important considerations include factors such as (a) the workpiece material and its manufacturing characteristics; see, for example, Table I.3 in the General Introduction, (b) compatibility of processing parameters, such as speeds, sizes, forces, energies, and temperature, among the two or more processes to be integrated, (c) cycle times of each individual operation involved, (d) possible adverse effects of the presence of various elements such as abrasives, chemicals, wear particles, chips, and contaminants on the overall operation, and (e) 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 useful particularly for difficult-to-machine materials and for parts with complex internal and external profiles. The economic production run for a particular process depends on the costs of tooling and equipment, the operating costs, the material-removal rate required, and the level of operator skill required, as well as on secondary and finishing operations that subsequently may be necessary.

In chemical machining (which is inherently a slow process), an important factor is the cost of reagents, maskants, and 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 in these processes, as can be seen in Table 27.1. The cost of tooling and equipment also varies considerably, as does the operator skill required. The high capital investment for machines (such as those for electrical and high-energy-beam machining, especially when equipped with robotic controls) has to be justified in terms of the production runs and the feasibility of manufacturing the same part by other means if at all possible.

CASE STUDY 27.2 Manufacture of Small Satellites

Satellites constructed 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 to reduce the size of satellites, none greater than the cost of putting the satellite into orbit. One of the main sources of weight in a satellite is its propulsion system, which is essential to changing its orbit or correcting for drift. Figure 27.18a shows the propulsion system for a microsatellite, incorporating a number of cold gas microthrusters, a propellant storage tank, filters, and temperature and pressure sensors.

Selected components of the propulsion system are shown in Fig. 27.18b. Note that the production of these miniature parts would be extremely difficult through conventional forming, casting, or machining technology. Furthermore, connecting the plumbing for all of the 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

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

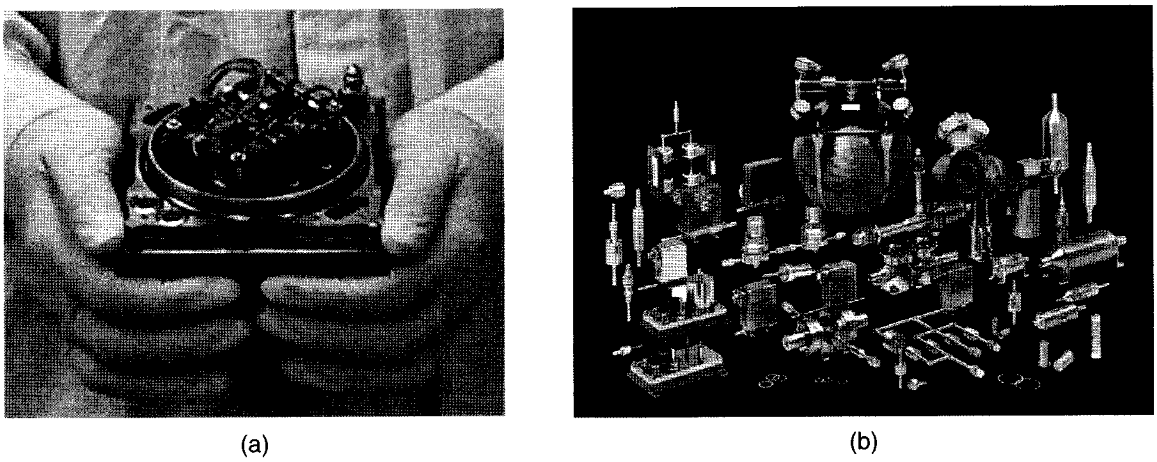

(a) (b)

FIGURE 27.18 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.

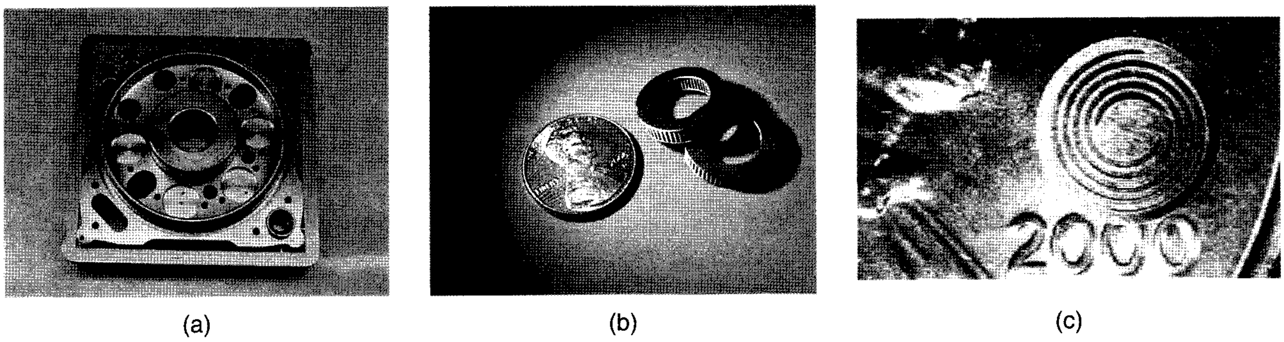

(a) (b) (c)

FIGURE 27.19 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.

are welded or mechanically fastened. Such a support is shown in Fig. 27.19, along with valve springs and filters that are produced through photochemical blanking combined with diffusion-bonding processes.

Figure 27.20 depicts the manufacturing sequence followed. Titanium is commonly used for propulsion system components, because it has a high strength-to-weight ratio, thus making possible lightweight designs.

(a) (b)

(c) (d)

(e) (f)

FIGURE 27.20 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.

A mask is prepared as described in 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 (see Section 31.7) to create internal features such as flow channels.

Such fully integrated systems have resulted in 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.

SUMMARY

- Advanced machining processes have unique capabilities and utilize chemical, electrochemical, electrical, and high-energy-beam sources of energy. The mechanical properties of the workpiece material are not significant, because these processes rely on mechanisms that do not involve the strength, hardness, ductility, or toughness of the material. Rather, they involve physical, chemical, and electrical properties.
- Chemical and electrical methods of machining are suitable particularly for hard materials and complex shapes. They do not produce forces (and therefore can be used for thin, slender, and flexible workpieces), significant temperatures, or residual stresses. However, the effects of these processes on surface integrity must be investigated, as they can damage surfaces considerably, thus reducing the fatigue life of the product.
- High-energy-beam machining processes basically utilize laser beams, electron beams, and plasma beams. They have important industrial applications, possess a high flexibility of operation with robotic controls, and are economically competitive with various other processes.
- Water-jet machining, abrasive water-jet machining, and abrasive-jet machining processes can be used for cutting as well as 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 by increasing speed and reducing, production times.

KEY TERMS

Abrasive-jet machining
Abrasive water-jet 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
Electron-beam machining
Etchant
Hybrid machining
Hydrodynamic machining
Laser
Laser-beam machining
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

ASM Handbook, Vol. 16: **Machining,** ASM International, 1989.

Brown, J., **Advanced Machining Technology Handbook,** McGraw-Hill, 1998.

Chryssolouris, G., and Sheng, P., **Laser Machining, Theory and Practice,** Springer, 1991.

Crafer, R.C., and Oakley, P.J., **Laser Processing in Manufacturing,** Chapman & Hall, 1993.

Dahotre, N.B., and Harimkar, S.P., **Laser Fabrication and Machining of Materials,** Springer, 2007.

El-Hofy, H.A.-G., **Advanced Machining Processes,** McGraw-Hill, 2005.

Gillespie, L., **Deburring and Edge Finishing Handbook,** Society of Manufacturing Engineers, 1999.

Grzesik, W., **Advanced Machining Processes of Metallic Materials: Theory, Modelling and Applications,** Elsevier, 2008.

Guitrau, E.B., **The EDM Handbook,** Hanser-Gardner, 1997.

Jain, V.K., and Pandey, P.C., **Theory and Practice of Electrochemical Machining,** Wiley, 1993.

Jameson, E.C., **Electrical Discharge Machining,** Society of Manufacturing Engineers, 2001.

Lange, K. (ed.), **Die Manufacture,** Chapter 32 in *Handbook of Metal Forming,* McGraw-Hill, 1985.

McGeough, J.A., **Advanced Methods of Machining,** Chapman & Hall, 1988.
Migliore, L., **Laser Materials Processing,** Marcel Dekker, 1996.
Momber, A.W., and Kovacevic, R., **Principles of Abrasive Water Jet Machining,** Springer, 1998.
Powell, J., CO_2 **Laser Cutting,** 2nd ed., Springer, 1998.
Sommer, C., **Non-Traditional Machining Handbook,** Advance Publishing, 1999.
Sommer, C., and Sommer, S., **Wire EDM Handbook,** Technical Advanced Publishing Co., 1997.
Steen, W.M., **Laser Material Processing,** Springer, 1991.
Taniguchi, N., **Energy-Beam Processing of Materials,** Oxford, 1989.

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 capabilities of wire EDM? Could this process be used to make tapered parts? Explain.

27.7. Describe the advantages of water-jet machining.

27.8. What is the difference between photochemical blanking and chemical blanking?

27.9. What type of workpiece is not suitable for laser-beam machining?

27.10. What is an undercut? Why must it be considered in chemical machining?

27.11. Explain the principle of hybrid machining.

QUALITATIVE PROBLEMS

27.12. Give technical and economic reasons that the processes described in this chapter might be preferred over those described in the preceding chapters.

27.13. Why is the preshaping or premachining of parts sometimes desirable in the processes described in this chapter?

27.14. Explain why the mechanical properties of workpiece materials are not significant in most of the processes described in this chapter.

27.15. Why has electrical-discharge machining become so widely used in industry?

27.16. Describe the types of parts that are suitable for wire EDM.

27.17. Which of the advanced machining processes would cause thermal damage? What is the consequence of such damage to workpieces?

27.18. Which of the processes described in this chapter require a vacuum? Explain why?

27.19. Describe your thoughts regarding the laser-beam machining of nonmetallic materials. Give several possible applications, including their advantages compared with other processes.

27.20. Are deburring operations still necessary for some parts made by advanced machining processes? Explain and give several specific examples.

27.21. List and explain factors that contribute to a poor surface finish in the processes described in this chapter.

27.22. What is the purpose of the abrasives in electrochemical grinding?

27.23. Which of the processes described in this chapter are suitable for producing very small and deep holes? Explain.

27.24. Is kerf width important in wire EDM? Explain.

27.25. Comment on your observations regarding Fig. 27.4.

27.26. Why may different advanced machining processes affect the fatigue strength of materials to different degrees?

QUANTITATIVE PROBLEMS

27.27. A 200-mm-deep hole that is 30 mm in diameter is being produced by electrochemical machining. A high production rate is more important than machined surface quality. Estimate the maximum current and the time required to perform this operation.

27.28. If the operation in Problem 27.27 were performed on an electrical-discharge machine, what would be the estimated machining time?

27.29. A cutting-off operation is being performed with a laser beam. The workpiece being cut is 12 mm thick and 380 mm long. If the kerf is 2.4 mm wide, estimate the time required to perform this operation.

27.30. 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

27.31. Explain why it is difficult to produce sharp profiles and corners with some of the processes described in this chapter.

27.32. 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.

27.33. Would the processes described in this chapter be difficult to perform on various nonmetallic or rubber-like materials? Explain your thoughts, commenting on the influence of various physical and mechanical properties of workpiece materials, part geometries, etc.

27.34. 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.

27.35. 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.

27.36. 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.

27.37. 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.

27.38. How would you manufacture a large-diameter, conical, round metal disk with a thickness that decreases from the center outward? Make appropriate sketches.

27.39. Describe the similarities and differences among the various design guidelines for the processes described in this chapter.

27.40. Describe any workpiece size limitations in advanced machining processes. Give examples.

27.41. Suggest several design applications for the types of parts shown in Fig. 27.5. (See also Fig. 27.16c.)

27.42. 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.

27.43. 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.

27.44. 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.

27.45. *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.

27.46. 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.

27.47. 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.

27.48. 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.

27.49. 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.

Micromanufacturing and Fabrication of Microelectronic Devices

PART V

In order to appreciate the importance of the topics covered in the two chapters in this part of the book, consider the manufacture of a simple spur gear made of metal. If the gear is 100 mm in diameter, it can be produced by traditional means, such as starting with a cast or forged blank and machining it, as described in various chapters. A gear that is 2 mm in diameter, however, can be difficult to produce to the desired dimensional accuracy. If sufficiently thin, the gear could be made from sheet metal by very fine blanking or chemical etching, or by electroforming. If the gear is only a few micrometers in size, it can be produced with techniques involving optical lithography, wet and dry chemical etching, and related processes described in the next two chapters. A gear that is only a few nanometers in diameter would be extremely difficult to produce; indeed, such a gear would, at most, have only a few tens of atoms across its surface. 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.

For most of the engineering profession's history, engineers have emphasized the design and manufacture of relatively large components. Conventional manufacturing processes, described in Chapters 11 through 27, typically produce parts that are larger than a millimeter or so and can be described as visible to the naked eye. The size of such parts generally are referred to as **macroscale**, the word "macro" being derived from the Greek *makros*, meaning "long". The processing of such parts is known as **macromanufacturing.** Numerous examples can be given, ranging from products found in a hardware store, to castings and forgings used in machinery, and to products as large as automobiles, aircraft, and ships. 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. Note that all of the examples and case studies given thus far in this book have been examples of macromanufacturing.

The gear shown in Fig. V.1 is the size of a few tens of micrometers across and fits into the realm of micromanufacture. Micromanufacturing, which by definition refers to manufacturing on a microscopic scale (that is, not visible to the naked eye), has been developed mostly for electronic devices of all kinds, including computer processors and memory chips, sensors, and magnetic storage devices. For the most part, this type of manufacturing relies heavily on lithography approaches, wet and dry etching, and coating operations. In addition, the micromanufacturing of semiconductors exploits the unique ability of silicon to form oxides.

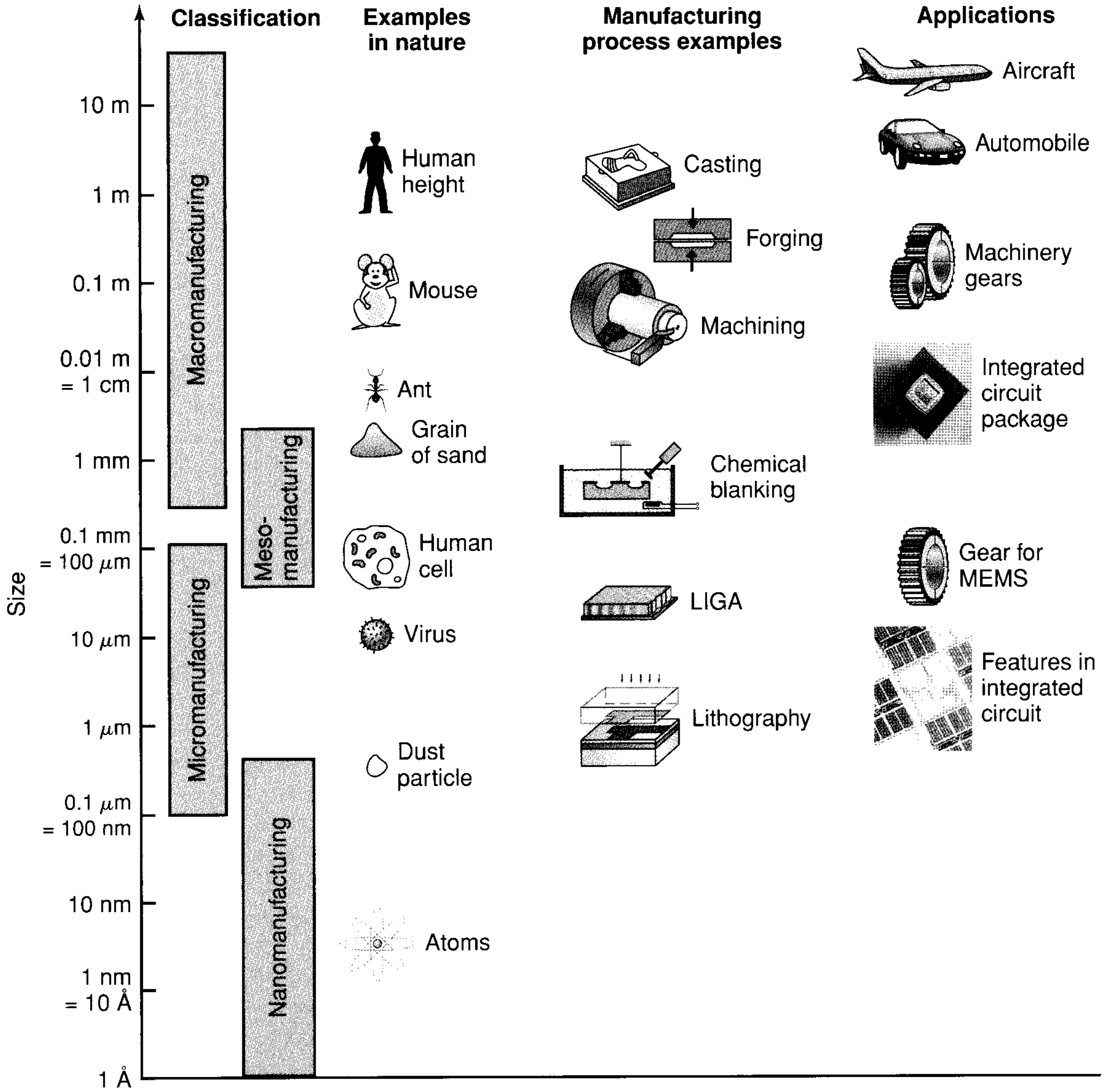


FIGURE V.1 Illustration of the regimes of macro-, meso-, micro-, and nanomanufacturing, the range of common sizes of parts, and the capabilities of manufacturing processes in producing those parts.

Examples of products that rely upon micromanufacturing approaches are a wide variety of sensors and probes (see Fig. V.2), ink-jet printing heads, microactuators and associated devices, magnetic hard-drive heads, and microelectronic devices such as computer processors and memory chips. Micromanufacturing methods allow the production of a wide variety of features at these length scales, but most experience is with electronic devices. Microscale mechanical devices are still a relatively new technology, but one that has developed with surprising speed.

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; medical devices such as stents and valves; and mechanical watches, with components exactly the same as the gear shown in Fig. V.1.

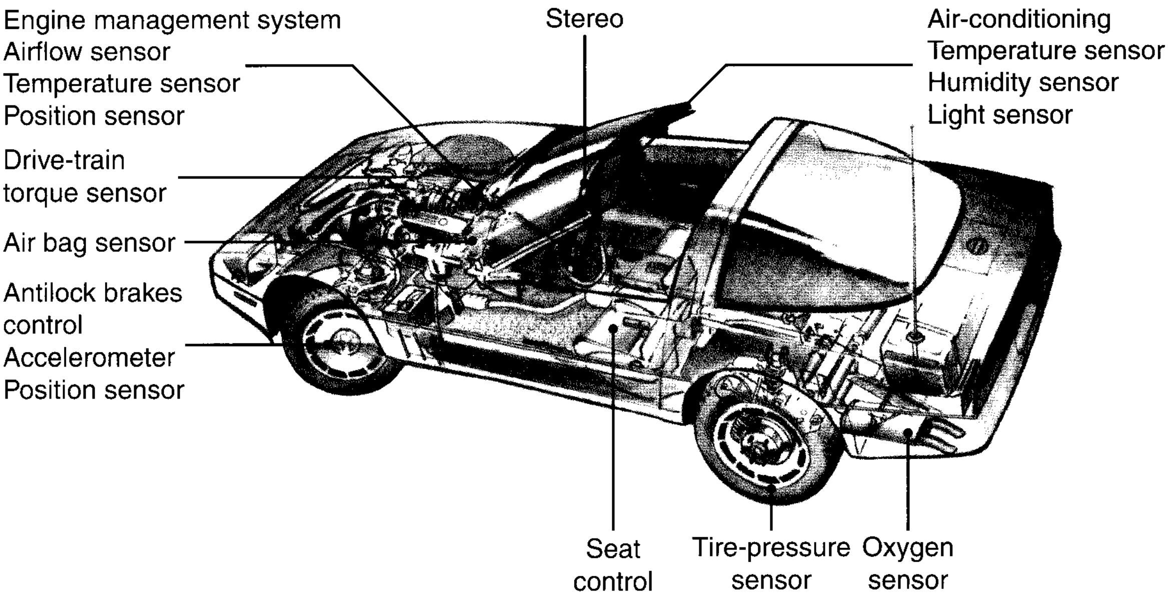


FIGURE V.2 Microelectronic and microelectromechanical devices and parts used in a typical automobile.

In **nanomanufacturing**, parts are produced at nanometer length scales, that is, one billionth of a meter and typically 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; molecularly engineered medicines and other forms of biomanufacturing are the only commercial examples. However, it is now recognized that many physical and biological processes act at the nanoscale and that this approach holds much promise for future innovations.

This part of the book emphasizes micro- and nanomanufacturing. Although these subdisciplines within the broad range of manufacturing engineering are now over five decades old, they have developed rapidly in the past two decades or so. Products made by micro- or nanomanufacturing have become very pervasive in modern society. Computers, communications, video, and control hardware of all types depend upon these manufacturing and material approaches.

In Chapter 28, we describe the manufacture of silicon wafers and microelectronic devices, which include a wide variety of computer processors, memory devices, and integrated circuits. Communication, entertainment, control, transportation, engineering design and manufacturing, and medicine have been changed greatly by the ready availability of metal-oxide-semiconductor (MOS) devices, usually based on single-crystal silicon. Microelectronics is the best known and commercially important example of micromanufacturing, with some aspects of the applications exemplifying nanomanufacturing. The techniques used in packaging and assembling integrated circuits onto printed circuit boards also are presented.

The 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** (MEMSs). While the historical origins of MEMS manufacture stem from the same processes used for microelectronic systems and from identical processes and production sequences that are still used, several unique approaches have been developed for the manufacture of microscale electromechanical devices and systems.

CHAPTER 28 Fabrication of Microelectronic Devices

EXAMPLES:

- This chapter presents the science and the technologies involved in the production of integrated circuits, a product that has fundamentally changed our society.
- The chapter begins by examining silicon, the preferred material for most integrated circuits, and its unique properties that make it attractive.
- Beginning with a cast ingot, the machining operations required to produce a wafer are described.
- Next, the production of patterns on wafers is discussed, including the processes of lithography, wet and dry etching, and doping.
- Metallization and testing are then described, as are the approaches for obtaining electrical connections from integrated circuits to circuit boards.
- The chapter concludes with a description of the different packages used for integrated circuits.

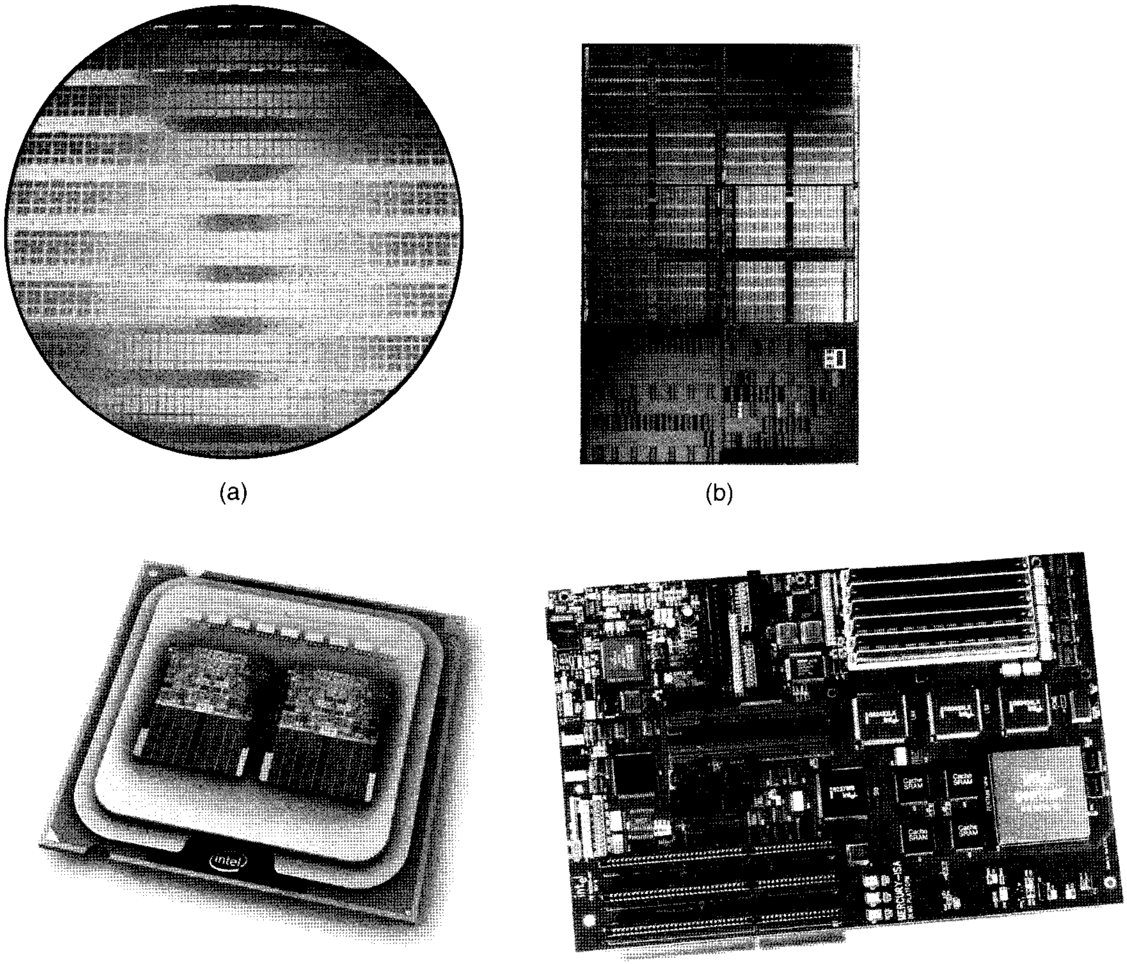
Typical parts produced: Computer processors, memory chips, printed circuit boards, and integrated circuits of all types.

28.1 Introduction

Although semiconducting materials have been used in electronics for a long time, it was the invention of the *transistor* in 1947 that set the stage for what would become one of the greatest technological advancements in all of history. **Microelectronics** has played an increasing role in our lives ever since **integrated circuit (IC)** technology became the foundation for calculators, wrist watches, controls for home appliances and automobiles, information systems, telecommunications, robotics, space travel, weaponry, and personal computers.

The major advantages of today's ICs are their very small size and low cost. As fabrication technology becomes more advanced, the size of devices made (such as transistors, diodes, resistors, and capacitors) continues to decrease. Consequently, more components can be put onto a **chip**—a small piece of semiconducting material on which the circuit is fabricated. In addition, mass production and automation have helped reduce the cost of each completed circuit.

Typical chips produced today have sizes that are as small as 0.5 × 0.5 mm and, in rare cases, can be more than 50 × 50 mm. In the past, no more than 100



(a) (b)

(c) (d)

FIGURE 28.1 (a) A 300-mm wafer with a large number of dies fabricated onto its surface; (b) detail view of an Intel 45-nm chip, including a 153-Mbit SRAM (static random access memory) and logic test circuits; (c) image of the Intel® Itanium® 2 processor; and (d) Pentium® processor motherboard. *Source*: Courtesy of Intel Corporation.

devices could be fabricated on a single chip; new technologies now allow densities in the range of 10 million devices per chip (Fig. 28.1). This magnitude of integration has been termed **very large scale integration** (VLSI). Some of the most advanced ICs may contain more than 100 million devices, termed **ultralarge-scale integration** (ULSI). The Intel Itanium® processors, for example, recently surpassed 2 billion transistors.

More recent advances include **wafer-scale integration** (WSI), in which an entire silicon wafer is used to build a single device. This approach has been of greatest interest in the design of massively parallel supercomputers, including **three-dimensional integrated circuits**, (3DICs) which use multiple layers of active circuits that maintain connections both horizontally and vertically.

This chapter describes the processes that are currently in use in the fabrication of microelectronic devices and integrated circuits and that follow the outline shown in Fig. 28.2. The major steps in fabricating a **metal-oxide-semiconductor field-effect transistor** (MOSFET), which is one of the dominant devices used in modern IC technology, are shown in Fig. 28.3.

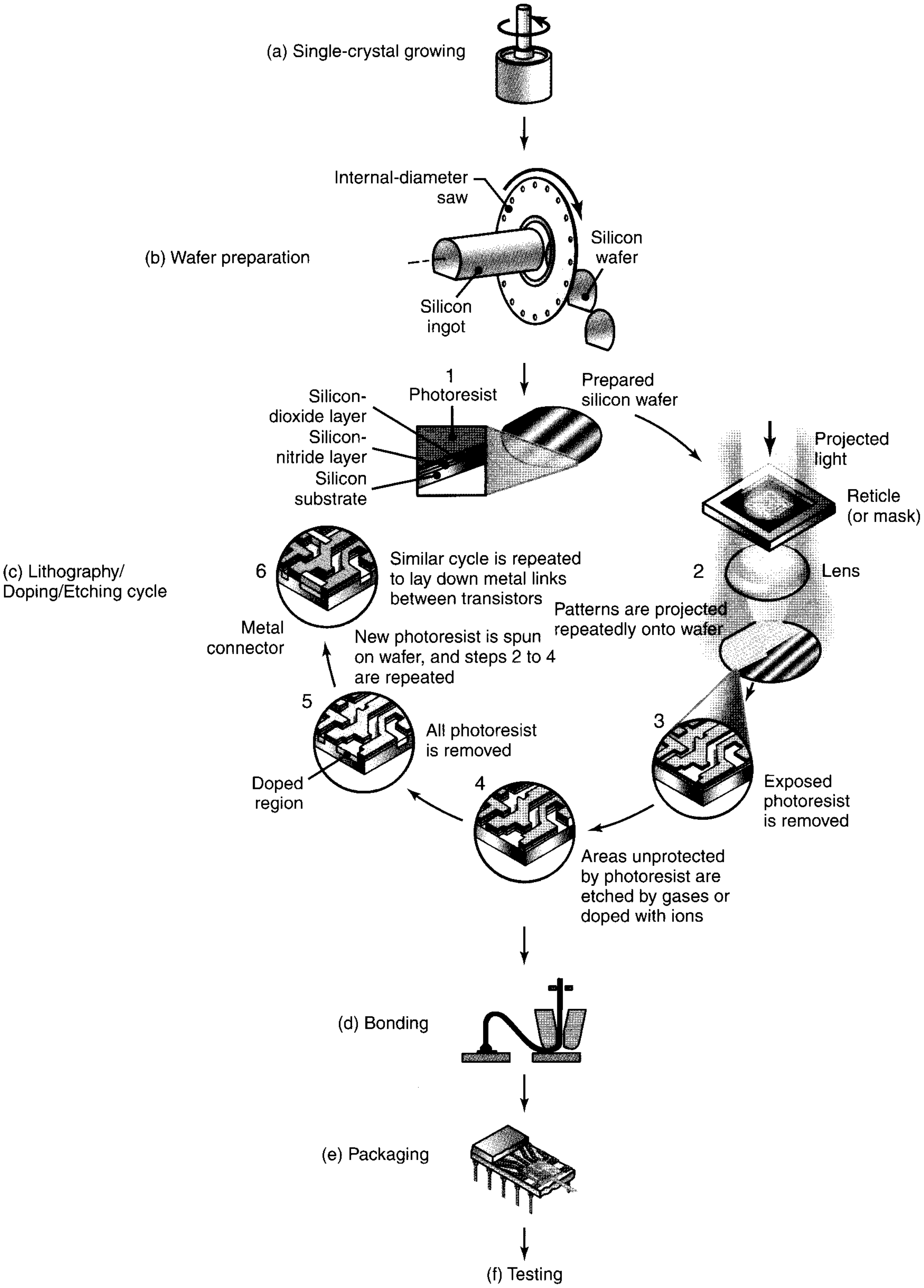


FIGURE 28.2 Outline of the general fabrication sequence for integrated circuits.

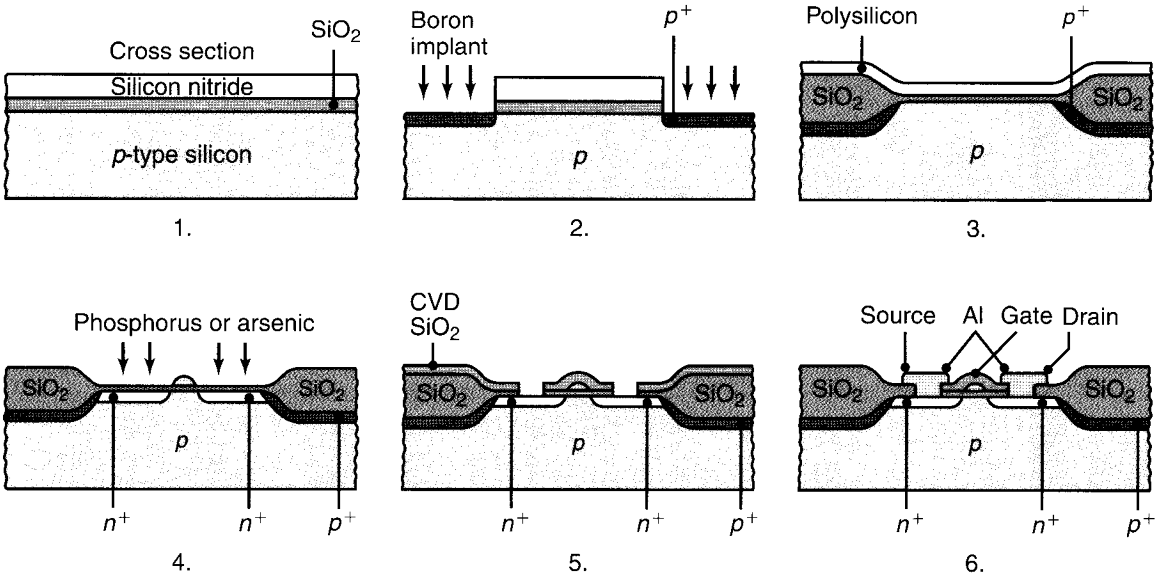


FIGURE 28.3 Cross-sectional views of the fabrication of a MOS transistor. *Source*: After R.C. Jaeger.

This chapter first introduces the fundamental properties of semiconductors and of silicon and then describes each of the major fabrication steps. Packaging of the integrated circuits and assembling them onto circuit boards are also discussed. Finally, the chapter describes current trends and forecasts in the microelectronics industry.

28.2 Clean Rooms

Clean rooms are essential for the production of most integrated circuits—a fact that can be appreciated by noting the scale of manufacturing to be performed. Integrated circuits are typically a few millimeters in length, and the smallest features in a transistor on the circuit may be as small as a few tens of nanometers. This size range is smaller than particles that we don't normally consider harmful, such as dust, smoke, perfume, and bacteria. However, if these contaminants are present on a silicon wafer during processing, they can seriously compromise the performance of the entire device. Thus, it is essential that all of the potentially harmful particles be eliminated from the IC manufacturing environment.

There are various levels of clean rooms, which are defined by the **class** of the room. The classification system refers to the number of 0.5-μm or larger particles within a cubic metre of air. Thus, a Class-0.35 clean room has 0.35×10^3 or fewer such particles per cubic metre. Clearly, the size and the number of particles are important in defining the class of a clean room, as shown in Fig. 28.4. Most clean rooms for microelectronics manufacuring range from Class 0.035 to Class 0.35. In comparison, the contamination level in modern hospitals is on the order of 350,000 particles per cubic metre.

To obtain 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.

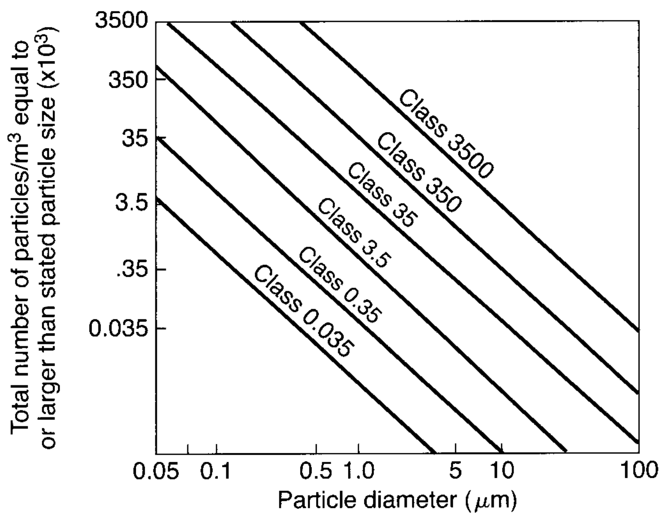


FIGURE 28.4 Allowable particle size counts for different clean-room classes.

The largest source of contaminants in a clean room is the workers themselves. Skin particles, hair, perfume, makeup, clothing, bacteria, and viruses are given off naturally by people and in sufficiently large numbers to quickly compromise a Class-3.5 clean room. For these reasons, most clean rooms require special coverings, such as white laboratory coats, gloves, and hairnets, as well as the avoidance of perfumes and makeup. The most stringent clean rooms require full-body coverings called *bunny suits*. There are other stringent precautions as well. For example, the use of a pencil, instead of a ballpoint pen, can produce objectionable graphite particles, and special clean-room paper is required to prevent the accumulation of paper particles in the air.

Clean rooms are designed such that the cleanliness at critical processing locations is greater than in the clean room in general. This is accomplished by directing the filtered ventilating air so that it displaces ambient air and directs dust particles away from the process—a goal that can be facilitated by laminar-flow hooded work areas.

28.3 Semiconductors and Silicon

As the name suggests, **semiconductor materials** have electrical properties that lie between those of conductors and insulators and that 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. These 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 be controlled through the creation of regions with different doping types and concentrations.

Although the earliest electronic devices were fabricated on *germanium*, **silicon** has become the industry standard. The abundance of alternative forms of silicon in the Earth is second only to that of oxygen, making it attractive economically. Silicon's main advantage over germanium is its large energy gap (1.1 eV) 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 purposes. By contrast, germanium oxide is water soluble and unsuitable for electronic devices. Furthermore, the oxidized form of silicon allows the production of **metal-oxide-semiconductor** (MOS) devices, which are the basis for MOS transistors. These materials make up memory devices, processors, and the like, and are by far the largest volume of semiconductor material produced worldwide.

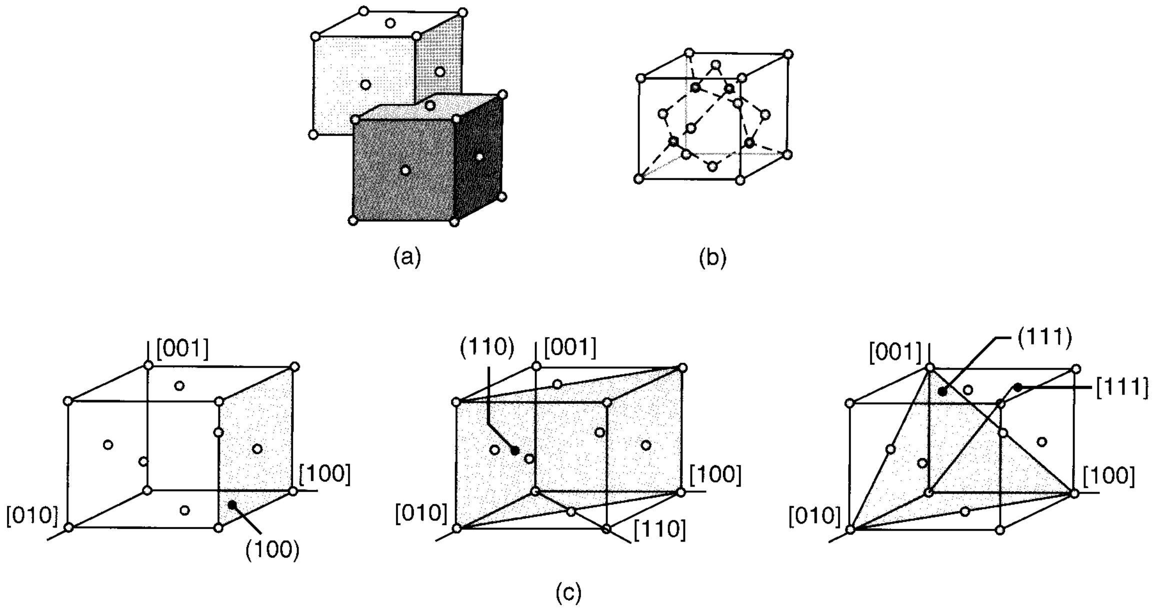


FIGURE 28.5 Crystallographic structure and Miller indices for silicon. (a) Construction of a diamond-type lattice from interpenetrating face-centered cubic cells; one of eight penetrating cells is shown. (b) Diamond-type lattice of silicon; the interiors have been shaded in color. (c) Miller indices for a cubic lattice.

Structure of Silicon. The crystallographic structure of silicon is a diamond-type fcc structure, as shown in Fig. 28.5, along with the *Miller indices* of an fcc material. Miller indices are a useful notation for identifying planes and directions within a unit cell. A crystallographic plane is defined by the reciprocal of its intercepts with the three axes. Since anisotropic etchants preferentially remove material in certain crystallographic planes, the orientation of the silicon crystal in a wafer is an important consideration.

In spite of its advantages, silicon has a larger energy gap (1.1 eV) than germanium oxide and, therefore, has a higher maximum operating temperature (about 200°C). This limitation 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 devices such as lasers and light-emitting diodes (LEDs).

Devices fabricated on gallium arsenide also have much higher operating speeds than those fabricated on silicon. Some of gallium arsenide's disadvantages 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 required for semiconductor device fabrication. The process begins by heating silica and carbon together in an electric furnace, which results in a 95-to-98% pure polycrystalline silicon. This material is converted to an alternative form, commonly trichlorosilane, which in turn is 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**, described in Section 11.5. The process 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 of the CZ process is a cylindrical single-crystal ingot, typically 100 to 300 mm in diameter and over 1 m in length. However, the technique does not allow for exact control of the ingot diameter. Therefore, ingots commonly are grown a few millimeters larger than the required size and are ground to a precise diameter. Silicon **wafers** then are produced from silicon ingots by a sequence of machining and finishing operations, as illustrated in Fig. 28.6.

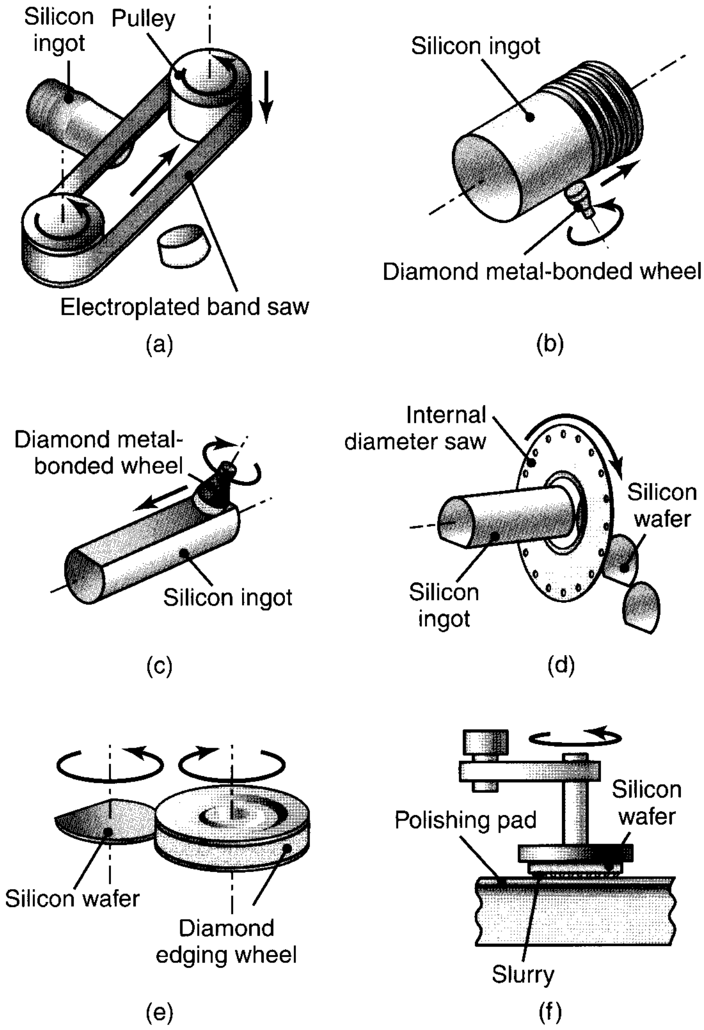


FIGURE 28.6 Finishing operations on a silicon ingot to produce wafers: (a) sawing the ends off the ingot; (b) grinding of the end and cylindrical surfaces of a silicon ingot; (c) machining of a notch or flat; (d) slicing of wafers; (e) end grinding of wafers; (f) chemical–mechanical polishing of wafers.

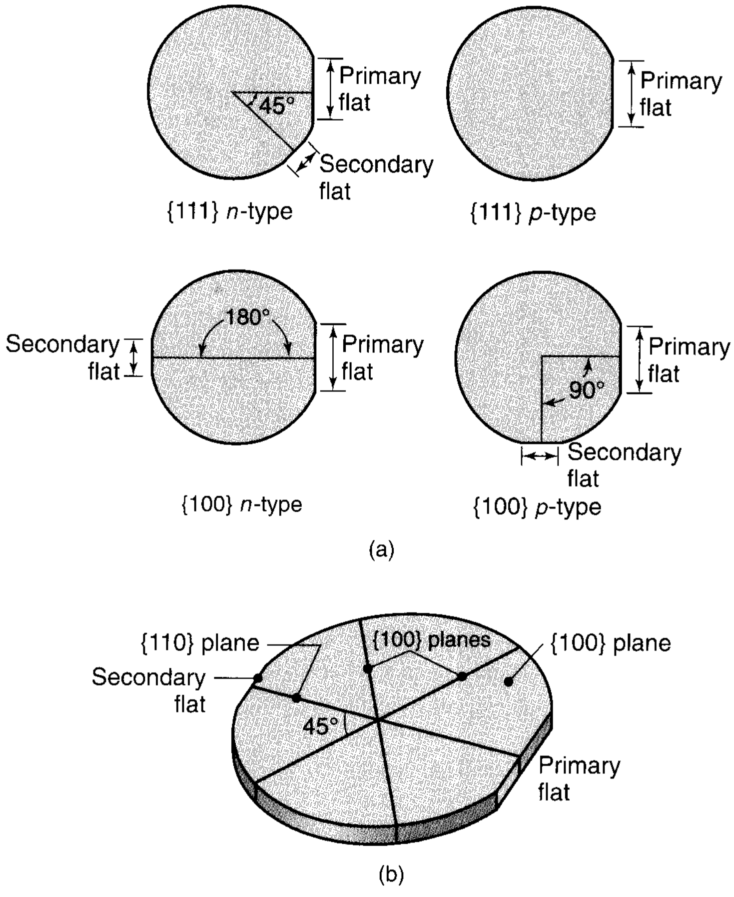


FIGURE 28.7 Identification of single-crystal wafers of silicon. This identification scheme is common for 150-mm diameter wafers, but notches are more common for larger wafers.

Next, the crystal is sliced into individual wafers by using an inner-diameter diamond-encrusted blade (Fig. 24.25f). In this method, a rotating, ring-shaped blade with its cutting edge on the inner diameter 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 necessary physical support for the absorption of temperature variations and the mechanical support needed during subsequent fabrication.

The wafer is then ground along its edges with a diamond wheel. This operation gives the wafer a rounded profile that is more resistant to chipping. Finally, the wafers must be polished and cleaned to remove surface damage caused by the sawing process. This is commonly performed by *chemical–mechanical polishing* (also referred to as *chemical–mechanical planarization*), as described in Section 26.7.

In order to properly control the manufacturing process, it is important to determine the orientation of the crystal in a wafer. Wafers, therefore, have notches or flats machined into them for identification, as shown in 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. Wafers are also identified by a laser scribe mark produced by the 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 unique wafer identification code.

Device fabrication takes place over the entire wafer surface. 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 be easily handled and transferred during processing steps. Because of the small device size and large wafer diameter, thousands of individual circuits can be placed on one wafer. Once processing is completed, the wafer is sliced into individual **chips**, each containing one complete integrated circuit.

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, and many chips are produced at the same time, as shown in Fig. 28.1a. The number of chips that can be produced is dependent on the cross-sectional area of the wafer. For this reason, a number of advanced-circuit manufacturers have moved towards using larger single-crystal cylinders, and 300-mm diameter wafers now have increasingly become common.

28.5 Film Deposition

Films of many different types are used extensively in microelectronic-device processing, particularly insulating and conducting films. Commonly deposited films include polysilicon, silicon nitride, silicon dioxide, tungsten, titanium, and aluminum. In some instances, the wafers merely serve as a mechanical support on which custom *epitaxial layers* are grown.

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 (notably carbon and oxygen), improved device performance, and the 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 applications, the film must inhibit the passage of dopants and concurrently display an ability 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 they often are used in combination.

Conductive films are used primarily for device interconnection. These films must have a low resistivity, be capable of carrying large currents, and be suitable for connection to terminal packaging leads with wire bonds. Generally, aluminum and copper are used for this purpose. Increasing circuit complexity has required up to six levels of conductive layers, all of which must be separated by insulating films.

Film Deposition. Films may be *deposited* by a number of techniques, which involve a variety of pressures, temperatures, and vacuum systems (see also Chapter 34):

- One of the simplest and oldest methods is **evaporation**, used primarily for depositing metal films. In this process, the metal is heated in a vacuum to its point of vaporization; upon evaporation, the metal forms a thin layer on the substrate surface. The heat for evaporation usually is generated by a heating filament or electron beam.
- **Sputtering** involves bombarding a target with high-energy ions (usually argon, Ar^+) in a vacuum. Sputtering systems generally include a DC power source to produce the energized ions. As the ions impinge on the target, atoms are knocked off and subsequently deposited on wafers mounted within the system. Although some argon may be trapped within the film, sputtering results in highly uniform coverage. Advances in this field 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 referred to as **low-pressure chemical-vapor deposition** (LPCVD) and is shown in Fig. 28.8b. Capable of coating hundreds of wafers at a time, this method results in a much higher production rate than that of atmospheric-pressure CVD and provides

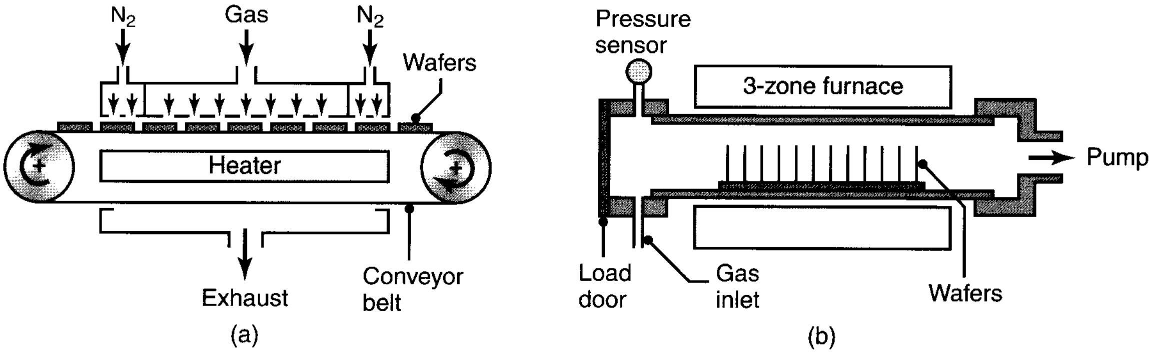


FIGURE 28.8 Schematic diagrams of (a) a continuous, atmospheric-pressure CVD reactor and (b) a low-pressure CVD. *Source*: After S.M. Sze.

superior film uniformity 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.

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, the heated substrate is brought into contact with a liquid solution containing the material to be deposited, a process called **liquid-phase epitaxy** (LPE).

Another high-vacuum process utilizes evaporation to produce a thermal beam of molecules that are deposited on the heated substrate. Called **molecular-beam epitaxy** (MBE), this process results 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. This level of control is 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 the previously described deposition techniques. Thermally grown oxides, described in this section, display a higher level of purity than 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, and a freshly sawed slice of silicon will grow a native oxide of 30 to 40 Å thickness ($Å = 10^{-7}$ mm) quickly.

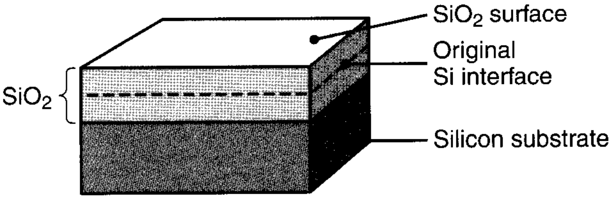


FIGURE 28.9 Growth of silicon dioxide, showing consumption of silicon. *Source*: After S.M. Sze.

- **Dry oxidation** is a relatively simple process and 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 able to pass 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 1000 Å thick, approximately 440 Å of silicon will be consumed. This requirement does not present a problem, as substrates always are grown sufficiently thick.

  One important effect of the consumption of silicon 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, processing parameters must be 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, but it suffers from a lower oxide density and, therefore, a lower dielectric strength. The common practice in industry is to combine both 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.

  The foregoing two oxidation methods 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 of oxidizing only certain areas is called **selective oxidation** and 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. There are many forms of lithography, but the most common form used today is **photolithography**. Electron-beam and X-ray lithography are of great interest because of their ability to transfer patterns of higher resolution, which is a necessary feature for the increased miniaturization of integrated circuits. Most IC applications can be manufactured successfully with photolithography.

Photolithography. Photolithography uses a **reticle**, also called a **photomask** or **mask**, 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 5× to 20× larger, although 10×

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.

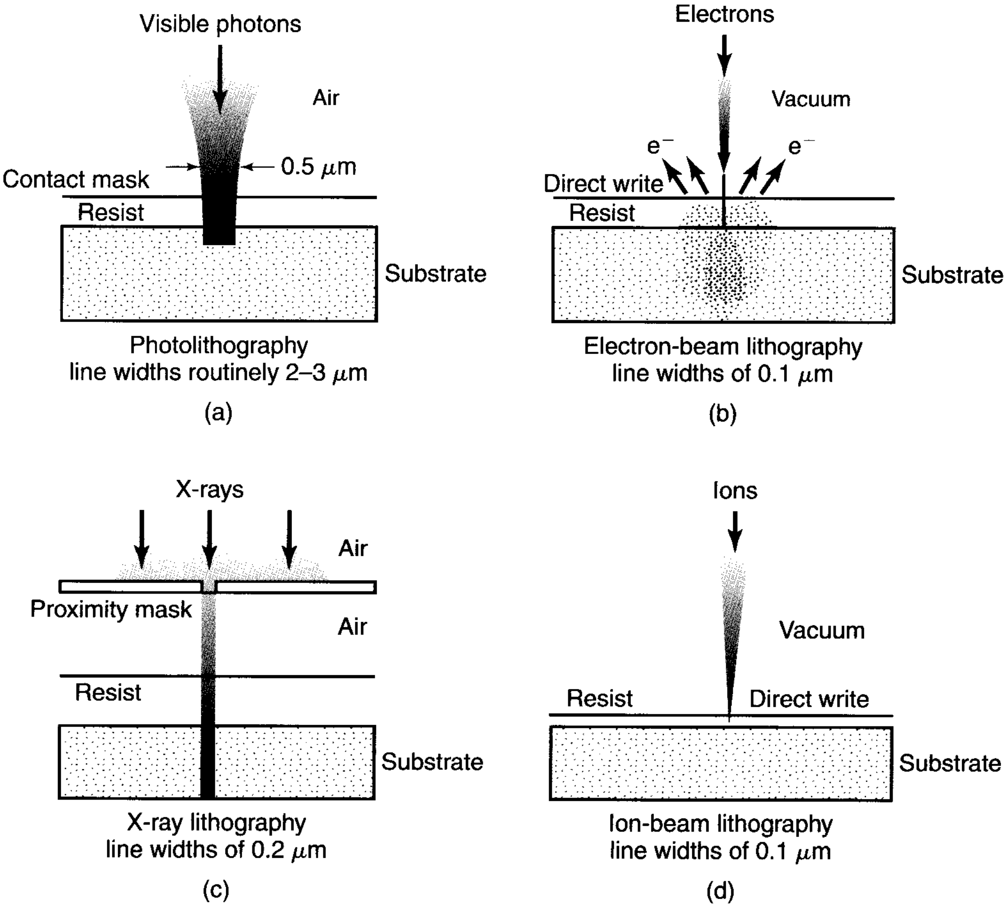


FIGURE 28.10 Comparison of lithography techniques. Note that (a) and (c) involve masking to achieve pattern transfer, while (b) and (d) scribe the pattern without a mask, known as *direct writing*.

magnification is the most common. Enlarged images then are focused onto a wafer through a lens system; this operation is referred to as *reduction lithography*.

In current practice, the *lithographic process* is applied to each microelectronic circuit as many as 25 times—each time using a different reticle to define the different areas of the working devices. 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; see Section 38.4) has had a major impact on reticle design and generation.

Cleanliness is especially important in lithography, and many manufacturers now use robotics and specialized wafer-handling apparatus 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 an organic polymer known as a **photoresist** (PR).

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, in order to deliver the polymer in liquid form.

Photoresist layers 0.5 to 2.5 μm thick are produced by applying the PR 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 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 ultraviolet (UV) radiation. 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, with negative masking, the photoresist can swell and distort, making it unsuitable for small geometries. Newer negative photoresist materials do not have this problem.

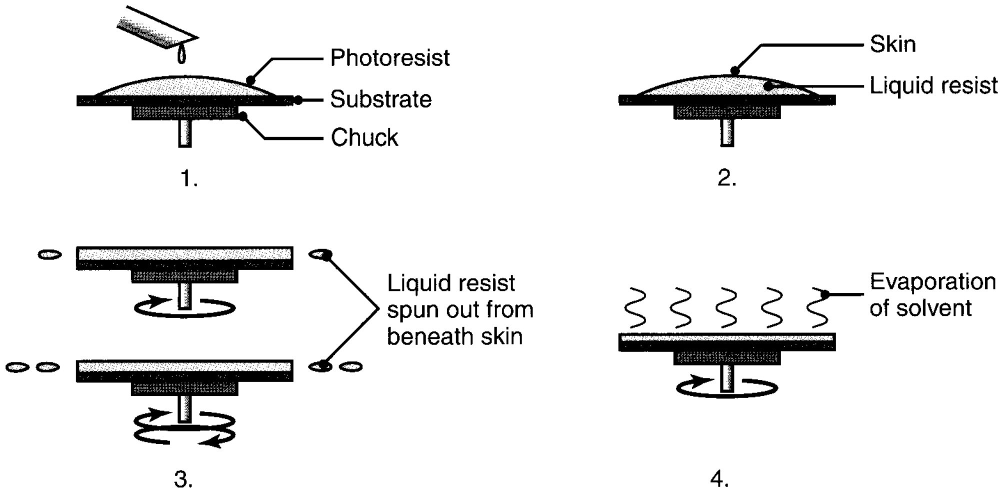


FIGURE 28.11 Spinning of an organic coating on a wafer.

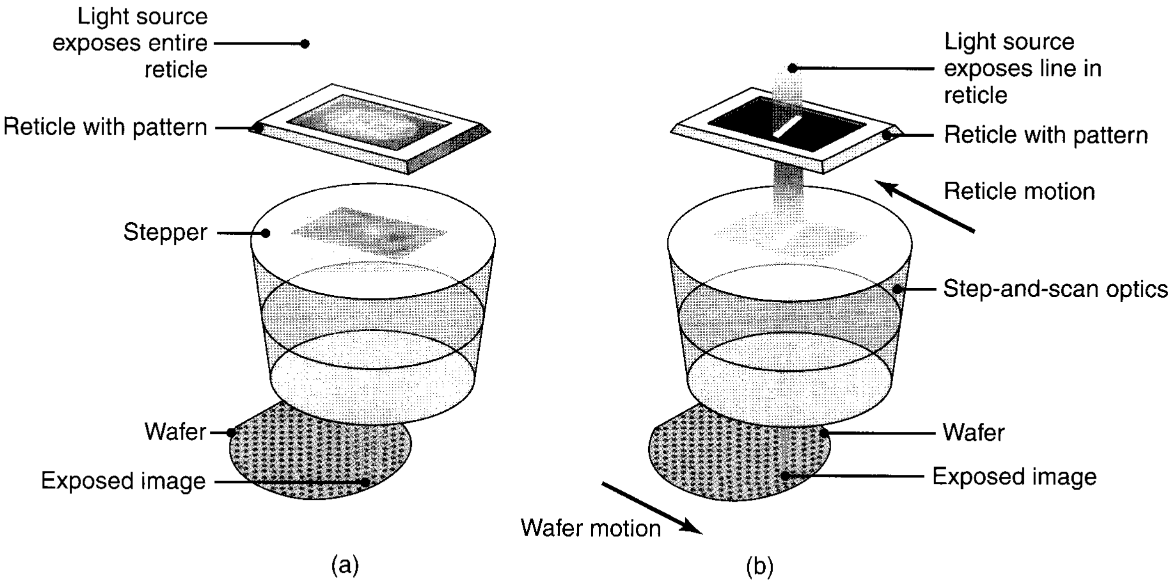


FIGURE 28.12 Schematic illustration of (a) wafer stepper technique for pattern transfer, and (b) step-and-scan technique.

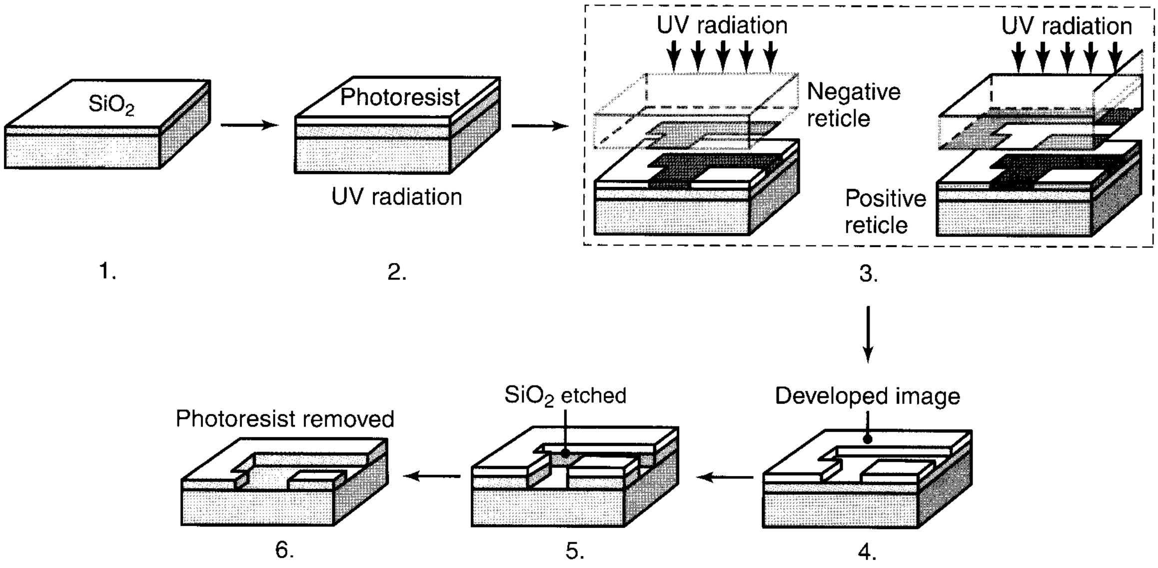


FIGURE 28.13 Pattern transfer by photolithography. Note that the mask in Step 3 can be a positive or negative image of the pattern.

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 (baking the wafer to about 150° 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).

After lithography, the developed photoresist must be removed in a process called **stripping.** In *wet stripping*, the photoresist is dissolved by solutions such as acetone or strong acids. *Dry stripping* involves exposing the photoresist to an oxygen plasma, which also is referred to as **ashing.** Dry stripping has become more popular, because it (a) does not involve the disposal of consumed hazardous chemicals and (b) is easier to control and can result in exceptional surfaces. Wet stripping solutions tend to lose potency in use.

One of the major issues in lithography is **line width**, the width of the smallest feature imprintable on the silicon surface. As circuit densities have escalated over the years, device sizes and features have become smaller and smaller. Today, commercially feasible minimum line widths are 45 nm, with considerable research directed at obtaining 32-nm or smaller line widths.

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. Alternatively, fluids with a refractive index higher than that of water are being investigated to increase the resolutions of immersion lithography.

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 obtain features in the range from 30 to 100 nm. The waves are focused through highly reflective molybdenum–silicon mirrors, instead of glass lenses, which absorb EUV light, 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 the 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. Furthermore, the aspect ratio (defined as the ratio of depth to lateral dimension) can be higher than 100 with X-ray lithography, but 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; however, the LIGA process (see Section 29.3) fully exploits the benefits of the technique.

Electron-beam and Ion-beam Lithography. Like X-ray lithography, *electron-beam* (e-beam) and *ion-beam* (i-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, using software (called *direct writing*). 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 techniques is that electron and ion beams have to be maintained in a vacuum, which significantly increases equipment complexity and production cost. Furthermore, the scan time for a wafer is much slower than that for other lithographic methods.

SCALPEL. Figure 28.14 shows the SCALPEL (*sc*attering with *a*ngular *l*imitation *p*rojection *e*lectron-beam *l*ithography) process, in which a mask is produced from about a 0.1-μm–thick membrane of silicon nitride and is 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 the SCALPEL process is the small-sized masks that are currently in use; however, 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 expensive as the reticle.

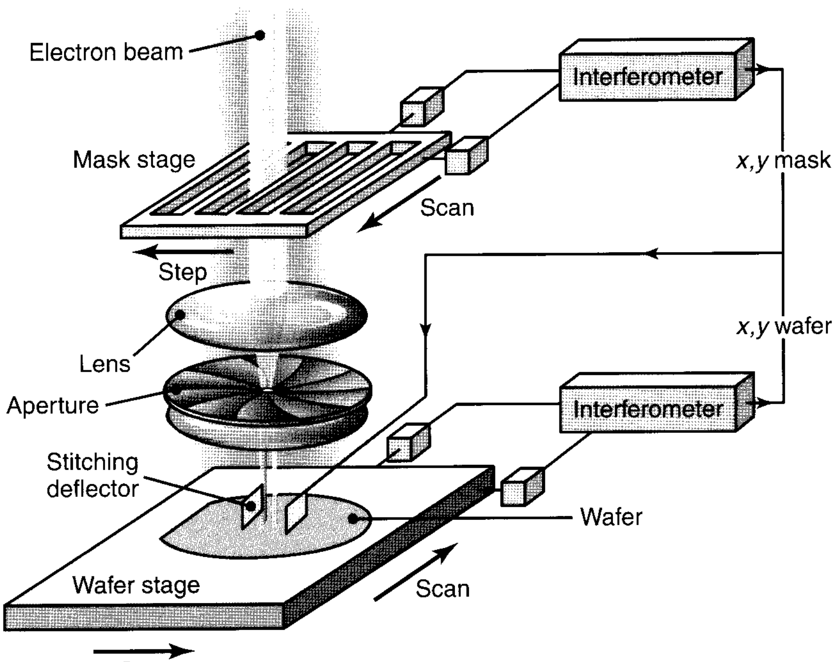


FIGURE 28.14 Schematic illustration of the SCALPEL process.

EXAMPLE 28.1 Moore's Law

G. Moore, an inventor of the IC and past chairman of Intel Corporation, 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, and the resulting estimate is now widely known as Moore's law. This law has been remarkably accurate. Figure 28.15 shows the historical progression of feature size in *dynamic random access memory* (DRAM) bits, as well as projected future developments. Looking ahead, however, one can see some major impediments to the continued reliability of Moore's law. Among the more important ones are the following:

- To produce ever smaller features in the 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. However, above a certain limit, the doping atoms cluster together. The result is 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 smaller voltages for proper operation. However, since the power consumption is still relatively high, very large currents are needed between the power-conversion devices and central-processing units of modern microprocessors. These large currents result in resistive heating, thus compounding the heat-extraction problems.
- Quantum effects play a large role at the small length scales, and system stability becomes an issue.

Much research is being 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 that, in turn, have been overcome.

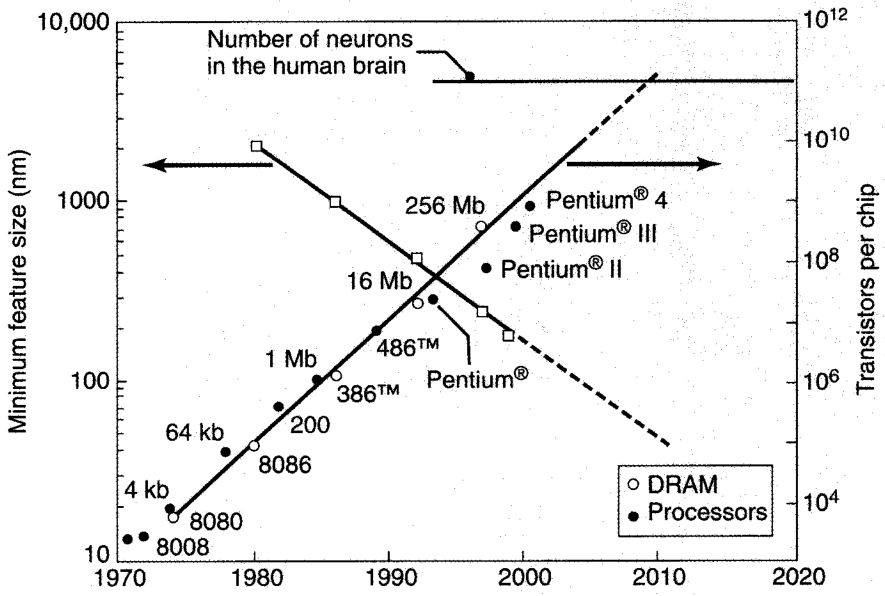


FIGURE 28.15 Illustration of Moore's law. *Source*: After M. Madou.

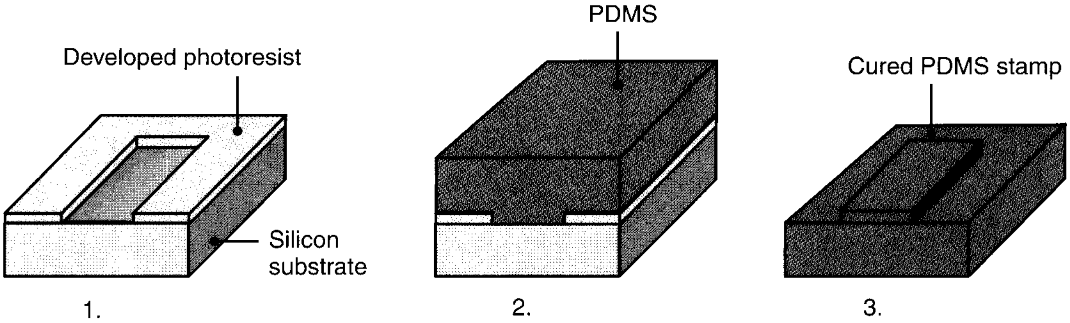


FIGURE 28.16 Production of a polydimethylsiloxane (PDMS) mold for soft lithography. 1. A developed photoresist is produced through standard lithography (see Fig. 28.13). 2. A PDMS stamp is cast over the photoresist. 3. The PDMS stamp is peeled off the substrate to produce a stamp. The stamp shown has 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 previously described. The master mold is then used to produce an elastomeric pattern, or stamp, as shown in Fig. 28.16. An elastomer that has been commonly used for the stamp is silicone rubber, or polydimethylsiloxane (PDMS), because it is chemically inert, is not hygroscopic (it does not swell with humidity), and has good thermal stability, strength, durability, and surface properties.

Several PDMS stamps can be produced from the same pattern, and each stamp can be used several times. Some of the common soft lithography processes are the following:

a. **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; see Section 33.6). This thin film can serve as a mask for selective wet etching, described shortly, or it can be used to impart a desired chemistry onto the surface.

b. **Microtransfer molding** (μTM). In this process, shown in Fig. 28.17a, the recesses in the PDMS mold are filled with a liquid polymer precursor and then pushed against a surface. After the polymer has cured, the mold is peeled off, leaving behind a pattern suitable for further processing.

c. **Micromolding in capillaries.** In this technique (MIMIC), shown in Fig. 28.17b, 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. The liquid can be a thermosetting polymer, a ceramic sol gel, or suspensions of solids within liquid solvents. Good pattern replication can occur, 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).

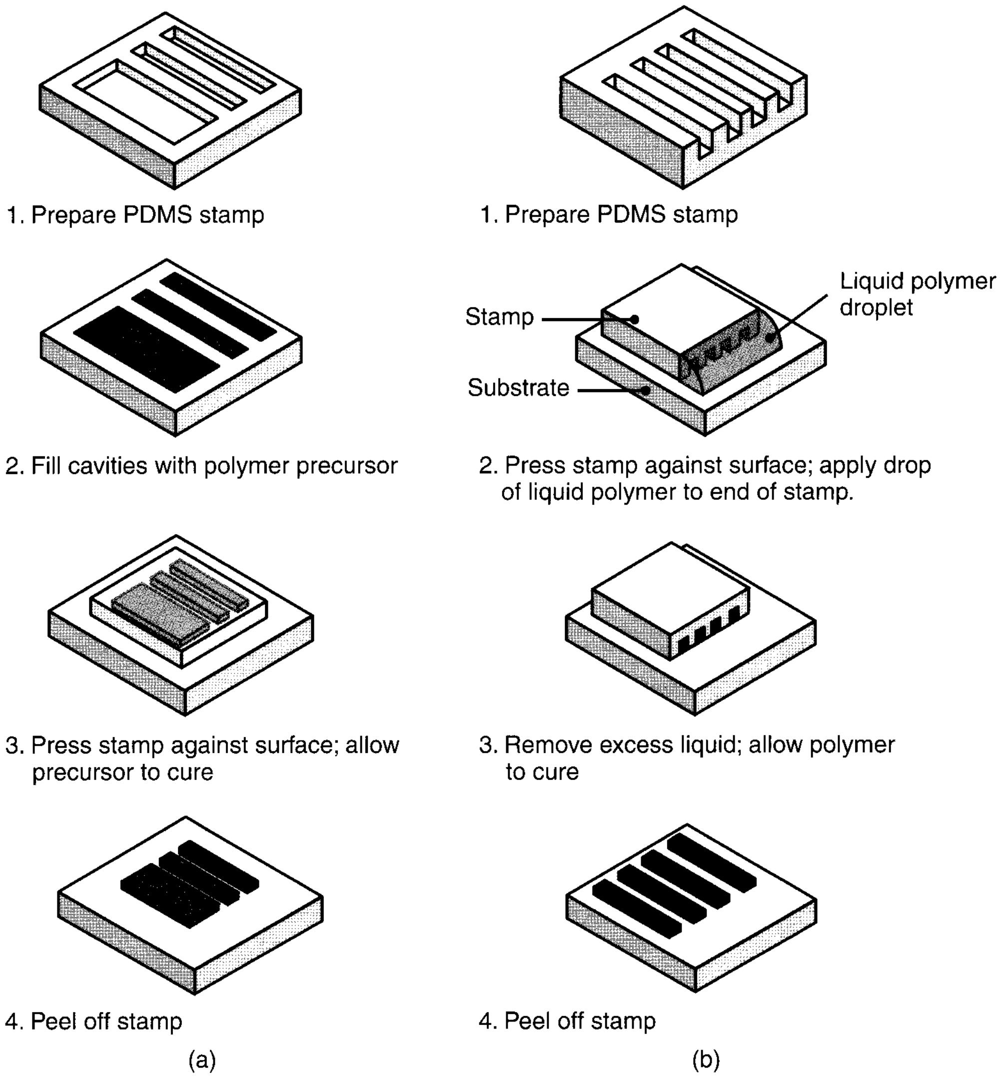


FIGURE 28.17 Soft lithography techniques. (a) Microtransfer molding (μTM), and (b) micromolding in capillaries (MIMIC). *Source*: After Y. Xia and G.M. Whitesides.

28.8 Etching

Etching is the process by which entire films or particular sections of films are removed, and it plays an important role in the fabrication sequence. One of the key criteria in this process is **selectivity**—that is, the ability to etch one material without etching another. A summary of etching processes and etchants is given in Tables 28.2 and 28.3.

In silicon technology, an etching process must etch the silicon-dioxide layer effectively and with minimal removal of 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 also 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** (defined as the ratio of the etch rates of the two films) can range from 1:1 to 100:1.

TABLE 28.2

General Characteristics of Silicon Etching Operations

	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.

28.8.1 Wet Etching

Wet etching involves immersing the wafers in a liquid solution, usually acidic. One 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. This isotropy results in *undercuts* beneath the mask material (see, for example, Figs. 27.3b and 28.18a) and limits the resolution of geometric features in the substrate.

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 limits the speed of the process, agitation or stirring of the solution can be employed to increase etching rates. If the second condition limits the speed of the process, the etching rate will depend strongly on temperature, etching material, and solution composition. Reliable etching thus requires both good temperature control and repeatable stirring capability.

Isotropic Etchants. These etchants are widely used for

- Removing damaged surfaces.
- Rounding of sharply etched corners to avoid stress concentrations.
- Reducing roughness after anisotropic etching.
- Creating structures in single-crystal slices.
- Evaluating defects.

Fabrication of microelectronic devices and microelectromechanical systems, described in Chapter 29, requires the precise machining of structures, done through masking. However, masking is a challenge with isotropic etchants. 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, and (b) produce rounded cavities. Furthermore, the etch rate is highly sensitive to agitation, and therefore, 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 desire to produce well-defined, extremely small structures.

TABLE 28.3

Comparison of Etch Rates for Selected Etchants and Target Materials

		Etch rate (nm/min)[a]							
Etchant	Target material	Polysilicon n^+	Polysilicon, undoped	Silicon dioxide	Silicon nitride	Phosphosilicate glass, annealed	Aluminum	Titanium	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_2O)	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% NH_4F and 8.3% HF by weight.

Source: After K. Williams and R. Muller.

Such small features cannot be attained through isotropic etching because of the poor definition resulting from undercutting of masks.

Anisotropic Etching. This situation takes place when etching is strongly dependent on compositional or structural variations in the material. There are two basic kinds of anisotropic etching: *orientation-dependent etching* (ODE) and *vertical etching*, although most vertical etching is done with dry plasmas and is described later. Orientation-dependent etching commonly occurs in a single crystal when etching takes place at different rates in different directions, as shown in Fig. 28.18b.

When used properly, anisotropic etchants produce geometric shapes with walls defined by the crystallographic planes that resist the etchants. For example, Fig. 28.19 shows the vertical etch rate for silicon as a function of temperature. As can be seen, etching 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

$$AR = \frac{E_1}{E_2} \qquad (28.1)$$

where E is the etch rate in a crystallographic direction of interest, and the subscripts refer to two directions or materials of interest. The anisotropy ratio is unity for isotropic etchants and can be as high as 400/200/1 for (110)/(100)/(111) silicon. The <111> planes always etch the slowest, but the <100> and <110> planes can be controlled through etchant chemistry.

Masking is also a concern in anisotropic etching, but for different reasons than with isotropic etching. Anisotropic etching is slow, typically 3 μm/min; thus, anisotropic etching through a wafer may take several hours. Silicon oxide may etch too rapidly to use as a mask; hence, a high-density silicon-nitride mask may be needed.

Often, it is important to rapidly halt the etching process. This is typically the case when thin membranes are to be manufactured or features with very precise thickness control are needed. Conceptually, rapid halting can be

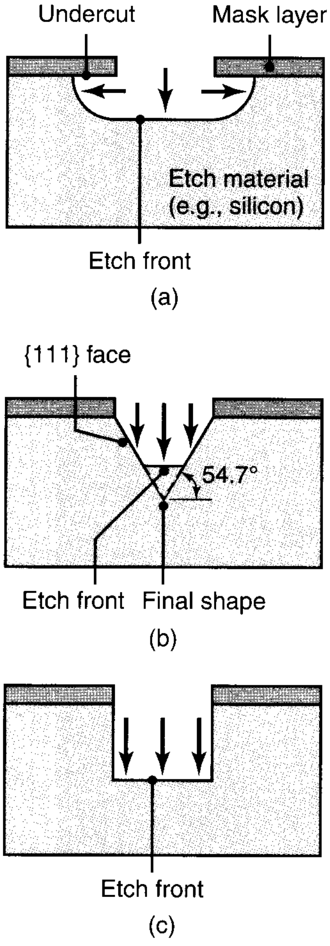


FIGURE 28.18 Etching directionality. (a) Isotropic etching: Etch proceeds vertically and horizontally at approximately the same rate, with significant mask undercut. (b) Orientation-dependent etching (ODE): Etch proceeds vertically, terminating on {111} crystal planes with little mask undercut. (c) Vertical etching: Etch proceeds vertically with little mask undercut. *Source*: Courtesy of K.R. Williams, Agilent Laboratories.

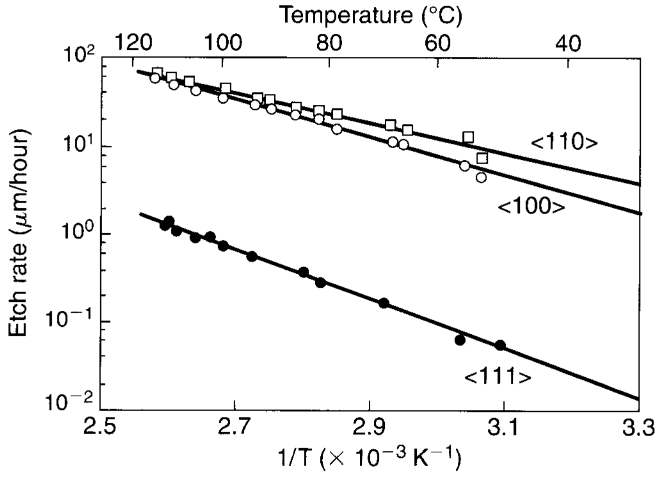


FIGURE 28.19 Etch rates of silicon in different crystallographic orientations, using ethylene-diamine/pyrocatechol-in-water as the solution. *Source*: After H. Seidel.

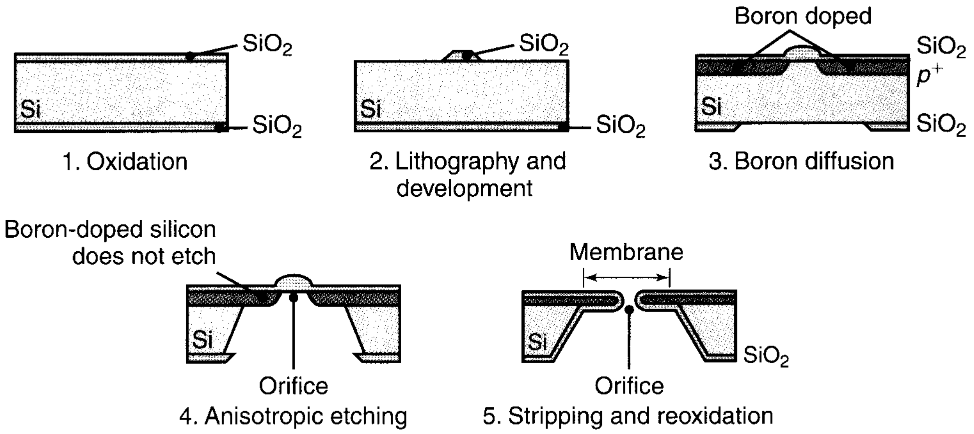


FIGURE 28.20 Application of a boron etch stop and back etching to form a membrane and orifice. *Source*: After I. Brodie and J.J. Murray.

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 etched depth.

The most common approach to obtain uniform feature sizes across a wafer is to use a boron etch stop, where a boron layer is diffused or implanted into silicon. Examples of common etch stops are the placement of a boron-doped layer beneath silicon or the placement of silicon dioxide (SiO_2) beneath silicon nitride (Si_3N_4). Since 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**. An example of the boron etch-stop approach is shown in Fig. 28.20.

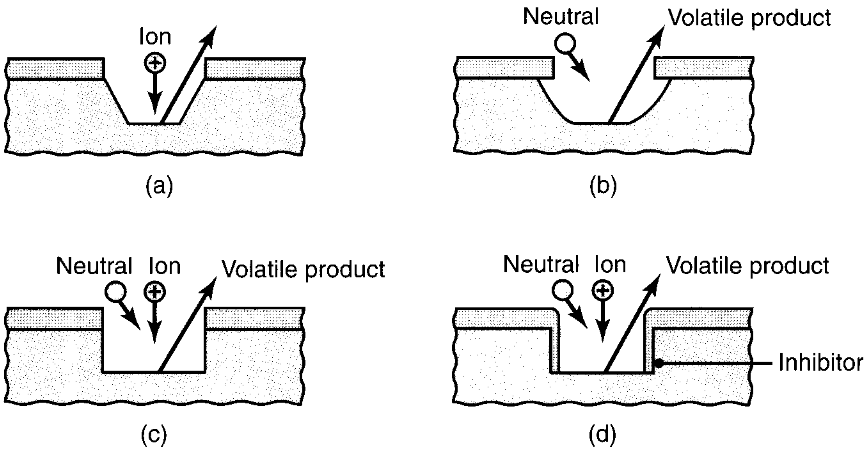
Several etchant formulations have been developed, including hydrofluoric acid, phosphoric acid, mixtures of nitric acid and hydrofluoric acid, potassium hydrochloride, and mixtures of phosphoric acid, nitric acid, acetic acid, and water. Wafer cleaning is done with a solution consisting of sulfuric acid and peroxide (*Piranha solution*, a trade name). Photoresist can be removed with these solutions, although acetone is used more commonly for this purpose.

28.8.2 Dry Etching

Modern integrated circuits are etched exclusively by *dry etching*, which involves the use of chemical reactants in a low-pressure system. In contrast to the wet-etching process, dry etching can have a high degree of directionality, resulting in highly anisotropic etching profiles (Fig. 28.18c). Also, the dry-etching process requires only small amounts of the reactant gases, whereas the solutions used in the wet-etching process 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.

Sputter Etching. *Sputter etching* 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. 28.21). 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



(a) (b) (c) (d)

FIGURE 28.21 Machining profiles associated with different dry-etching techniques: (a) sputtering; (b) chemical; (c) ion-enhanced energetic; (d) ion-enhanced inhibitor. *Source*: After M. Madou.

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 following:

- 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.
- Sputter etching is slow, with etch rates limited to tens of nanometers per minute.
- The photoresist is difficult to remove.

Reactive Plasma Etching. Also referred to as **dry chemical etching**, *reactive plasma etching* involves chlorine or fluorine ions (generated by RF excitation) and other molecular species that diffuse into and chemically react with the substrate, forming a volatile compound, which is then removed by a vacuum system. The mechanism of reactive plasma etching is shown in Fig. 28.22 and is as follows:

1. A reactive species, such as CF_4, is produced and dissociates upon impact with energetic electrons to produce fluorine atoms.
2. The reactive species then diffuses into the surface.
3. It becomes adsorbed.
4. The reactive species chemically reacts to form a volatile compound.
5. The reactant then desorbs from the surface.
6. It diffuses into the bulk gas, where it is removed by a vacuum system.

Some reactants polymerize on the surface and thus require additional removal, either with oxygen in the plasma reactor or by an external ashing operation. The electrical charge of the reactive species is not high enough to cause damage through impact on the surface, so no sputtering occurs. Thus, the etching is isotropic and undercutting of the mask takes place (Fig. 28.18a).

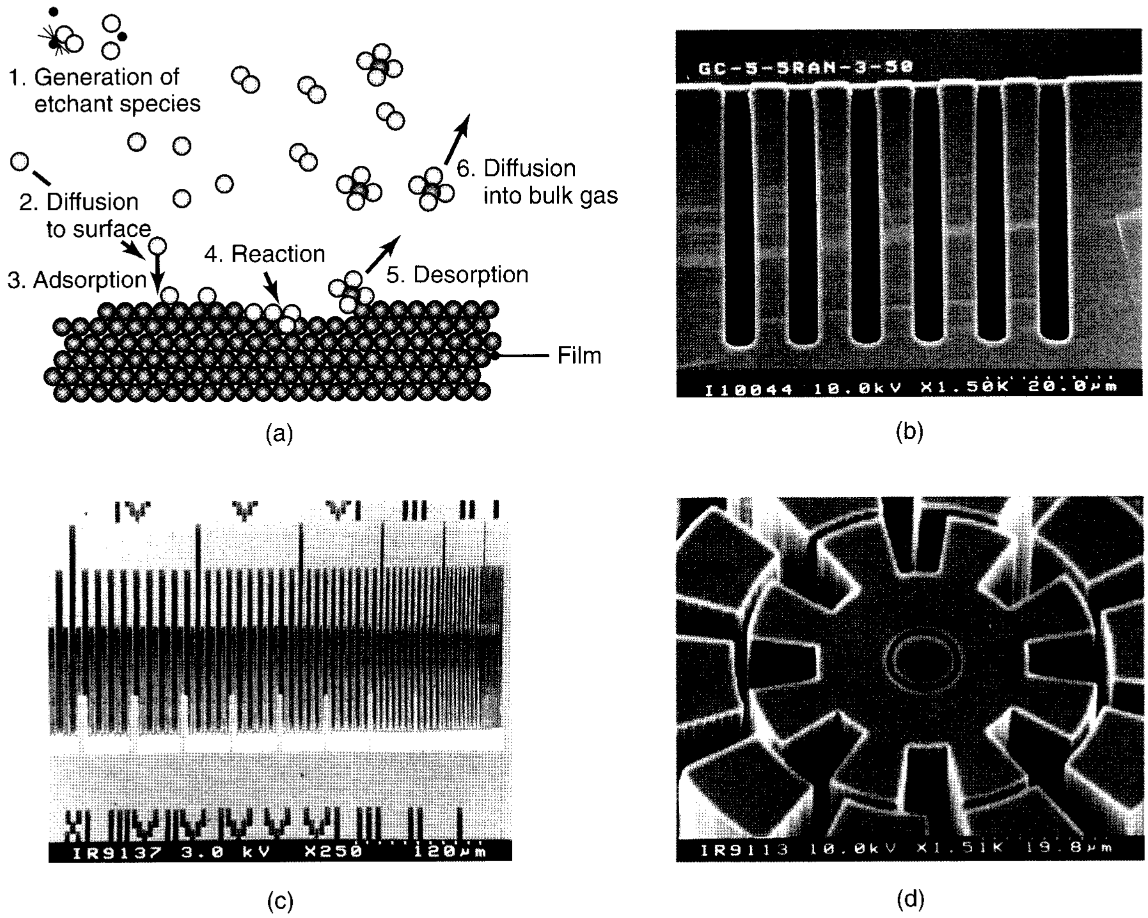


FIGURE 28.22 (a) Schematic illustration of reactive plasma etching. (b) Example of a deep reactive-ion etched trench. Note the periodic undercuts, or scallops. (c) Near-vertical sidewalls produced through deep reactive-ion etching (DRIE), an anisotropic-etching process. (d) An example of cryogenic dry 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.

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 this is 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, as shown in Fig. 28.22d.

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.
- 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 lowered to cryogenic temperatures, and then the CAIBE process takes place. The very low temperatures involved ensure that insufficient energy is available for a surface chemical reaction to take place, unless ion bombardment is normal to the surface. Oblique impacts, such as those occuring on sidewalls in deep crevices, cannot drive the chemical reactions.

Since dry etching is not selective, etch stops cannot be applied directly. Dry-etching reactions must be terminated when the target film is removed. Optical emission spectroscopy often is used to determine the end point of a reaction. Filters can be used to capture the wavelength of light emitted during a particular reaction. A noticeable change in light intensity at the point of etching will be detected.

EXAMPLE 28.2 Comparison of Wet and Dry Etching

Consider the case where a <100> wafer 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 <110> direction (see Fig. 28.5) of the wafer surface, as shown in Fig. 28.23.

Isotropic etching results in the cavity shown in Fig. 28.23a. 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. 28.23b. Because etching is much faster in the <100> and <110> directions than in the <111> direction, sidewalls defined by the <111> 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. 28.23c. The resultant pit is defined by <111> sidewalls and by a bottom in the <100> direction parallel to the surface. A rectangular mask and the resulting pit are shown in Fig. 28.23d. Deep reactive-ion etching is depicted in Fig. 28.23e. 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. 28.23f.

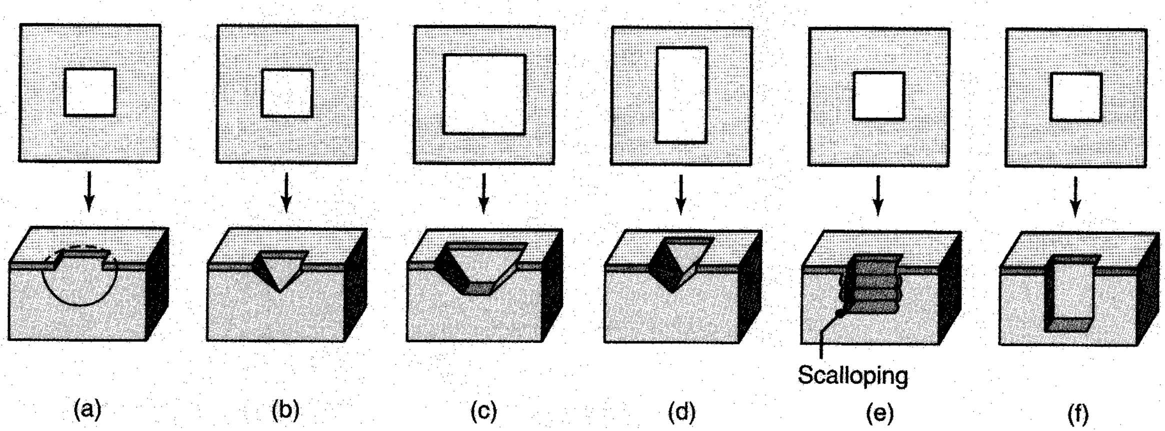


FIGURE 28.23 Various holes generated from a square mask in (a) isotropic (wet) etching; (b) orientation-dependent etching (ODE); (c) ODE with a larger hole; (d) ODE of a rectangular hole; (e) deep reactive-ion etching; and (f) vertical etching. *Source*: After M. Madou.

28.9 Diffusion and Ion Implantation

Recall that the electrical operation of microelectronic devices depends on regions that have different doping types and concentrations. The electrical character of these regions is altered through the introduction of dopants into the substrate, accomplished by the *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 elevated 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.

Because of the nature of diffusion, the dopant concentration is very high at the substrate surface and drops off sharply away from the surface. To obtain a more uniform concentration within the substrate, the wafer is heated further to drive in the dopants in a process called **drive-in diffusion**. Diffusion, whether desired or not, always occurs at high temperatures; this phenomenon is always taken into account during subsequent processing steps. Although the diffusion process is relatively inexpensive, it is highly isotropic.

Ion implantation is a much more extensive process and requires specialized equipment (Fig. 28.24; 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 uniform coverage of the substrate. The complete implantation operation must be performed in a vacuum.

The high-velocity impact of ions on the silicon surface damages the lattice structure and results in lower electron mobilities. This condition is undesirable, but 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.

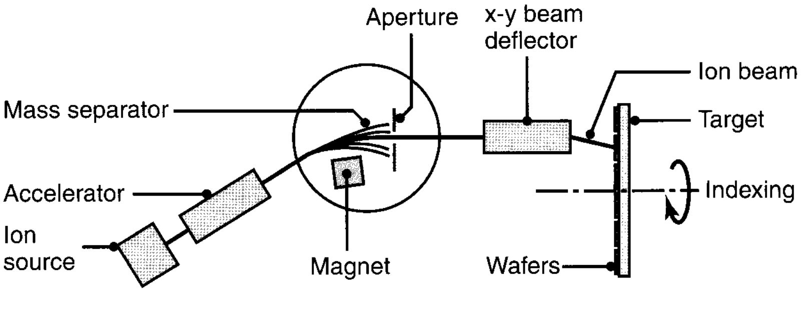


FIGURE 28.24 Schematic illustration of an apparatus for ion implantation.

EXAMPLE 28.3 Processing of a p-type Region in n-type Silicon

Suppose that we wish to create a p-type region within a sample of n-type silicon. Draw cross sections of the sample at each processing step in order to accomplish this.

Solution See Fig. 28.25. This simple device is known as a *pn-junction diode*, and the physics of its operation is the foundation for most semiconductor devices.

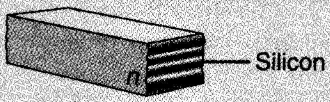
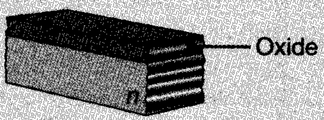
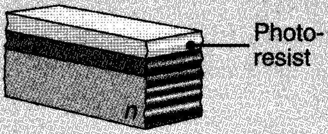
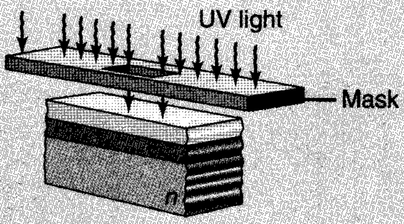
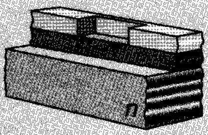
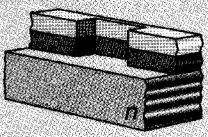
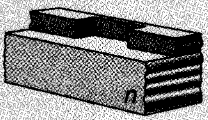
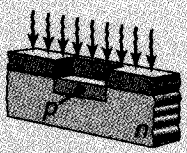
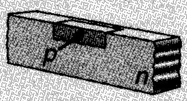
Processing step	Cross section	Description
1.		Sample of n-type silicon
2.		Grow silicon dioxide by oxidation
3.		Apply photoresist
4.		Expose photoresist, using appropriate lithographic mask
5.		Develop photoresist
6.		Etch silicon dioxide
7.		Remove photoresist
8.		Implant boron
9.		Remove silicon dioxide

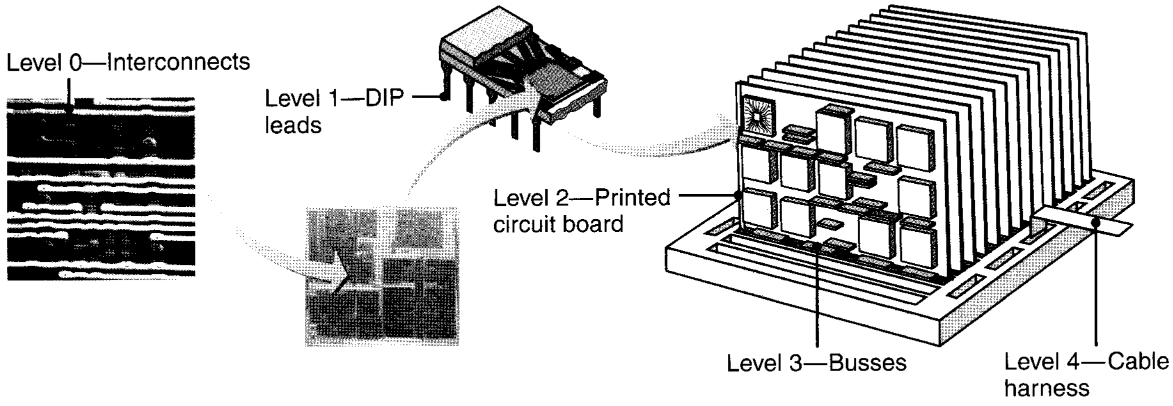
FIGURE 28.25 Fabrication sequence for a pn-junction diode.

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, and this must take place on a number of levels (Fig. 28.26). **Interconnections** are made with metals that exhibit low electrical resistance and good adhesion to dielectric insulator surfaces. Aluminum and aluminum–copper alloys remain the most commonly used materials for this purpose in VLSI technology today.

However, as device dimensions continue to shrink, electromigration has become more of a concern with aluminum interconnects. **Electromigration** is the 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 the problem include (a) the addition of sandwiched metal layers such as tungsten and titanium, and, more recently, (b) the use of pure copper, which displays lower resistivity and has significantly better electromigration performance than aluminum.

Metals are deposited by standard deposition techniques, and interconnection patterns are generated through lithographic and etching processes, as previously described. Modern ICs typically have one to six layers of metallization, each layer of metal being insulated by a dielectric.



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 28.26 Connections between elements in the hierarchy for integrated circuts.

Planarization (i.e., producing a planar surface) of these interlayer dielectrics is critical to the reduction of metal shorts and the line width variation of the interconnect. A common method used to achieve a planar surface is a uniform oxide-etch process that smoothens out the peaks and valleys of the dielectric layer.

However, today's standard for planarizing high-density interconnects has quickly become **chemical–mechanical polishing** (CMP), described in 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 and can polish a wafer to within 0.03 μm of being perfectly flat, with a R_q roughness (see Section 4.4) on the order of 0.1 nm for a new, bare silicon wafer.

Layers of metal are connected together by **vias**, and access to the devices on the substrate is achieved through **contacts** (Fig. 28.27). In recent years, as devices have 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. 28.28). Each chip, also known as a **die**, is tested by a computer-controlled probe platform containing needle-like 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 dies). These probes consist of transistors and interconnect structures that measure various processing parameters, such as resistivity, contact resistance, and electromigration.
2. **Direct probe.** This approach uses 100% testing on the bond pads of each die.

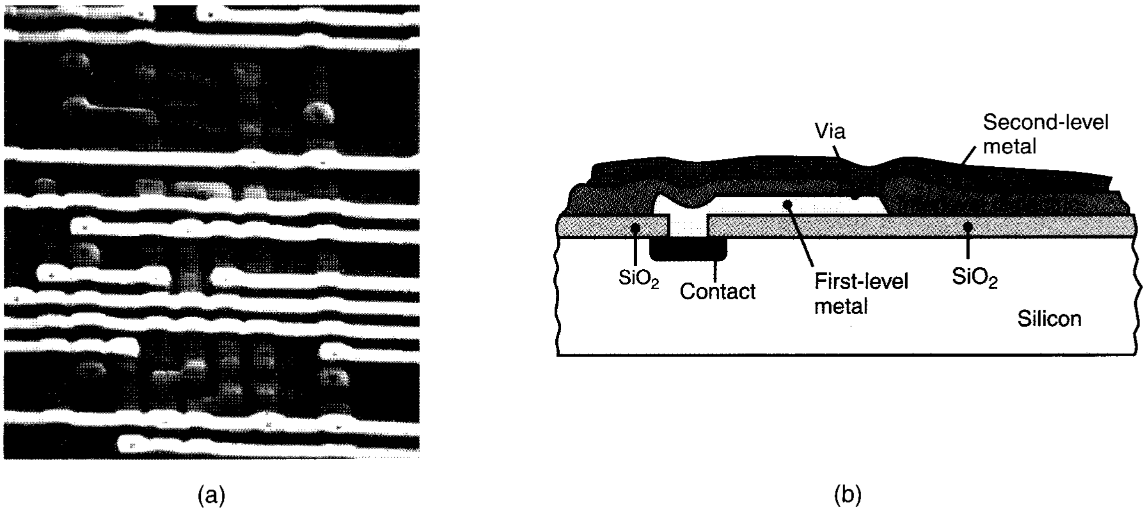


FIGURE 28.27 (a) Scanning-electron microscope (SEM) photograph of a two-level metal interconnect. Note the varying surface topography. (b) Schematic illustration of a two-level metal interconnect structure. *Source*: (a) Courtesy of National Semiconductor Corporation. (b) After R.C. Jaeger.

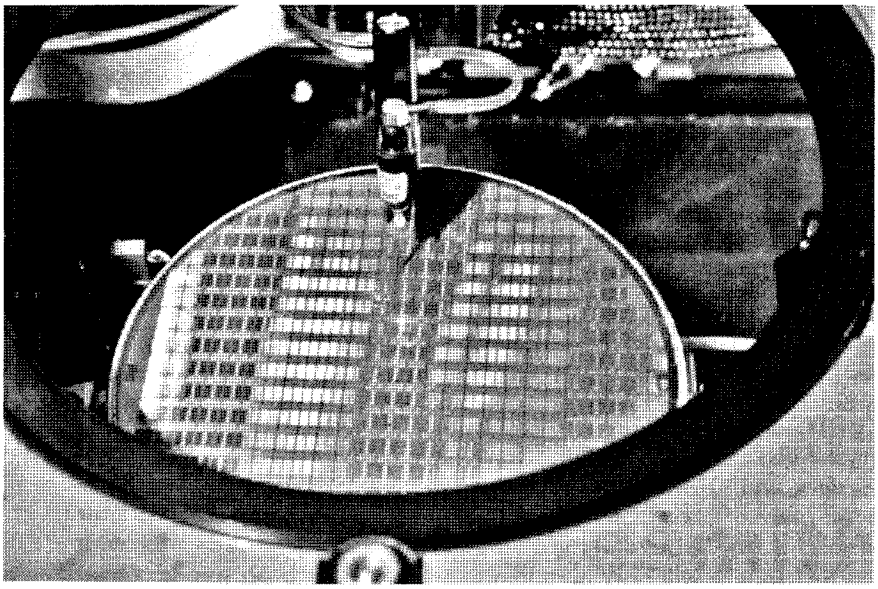


FIGURE 28.28 A probe (top center) checking for defects in a wafer; an ink mark is placed on each defective die. *Source*: Courtesy of Intel Corp.

The platform scans across the wafer and uses computer-generated timing waveforms to test whether each circuit is functioning properly. If a defective chip is encountered, 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.

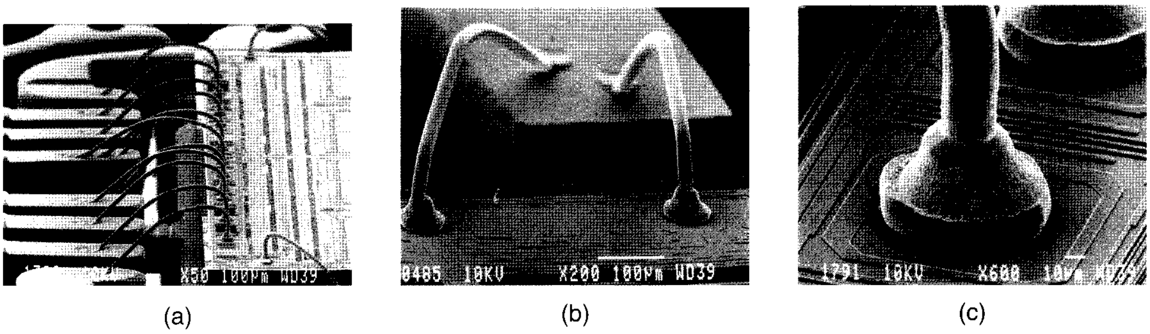
After the wafer-level testing is completed, back grinding may be done to remove a large amount of the original substrate. The final die thickness depends on the packaging requirement, but anywhere from 25 to 75% of the wafer thickness may be removed. After back grinding, each die is separated from the wafer. *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 an epoxy cement; another method makes use of a eutectic bond made by heating metal-alloy systems (see Section 4.3). One widely used mixture is 96.4% Au and 3.6% Si, and it 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. 28.29). The bonding pads on the die are typically drawn at 75 to 100 μm per side, and the bond wires are attached by means of thermocompression, ultrasonic, or thermosonic techniques (Fig. 28.30).

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



(a) (b) (c)

FIGURE 28.29 (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.

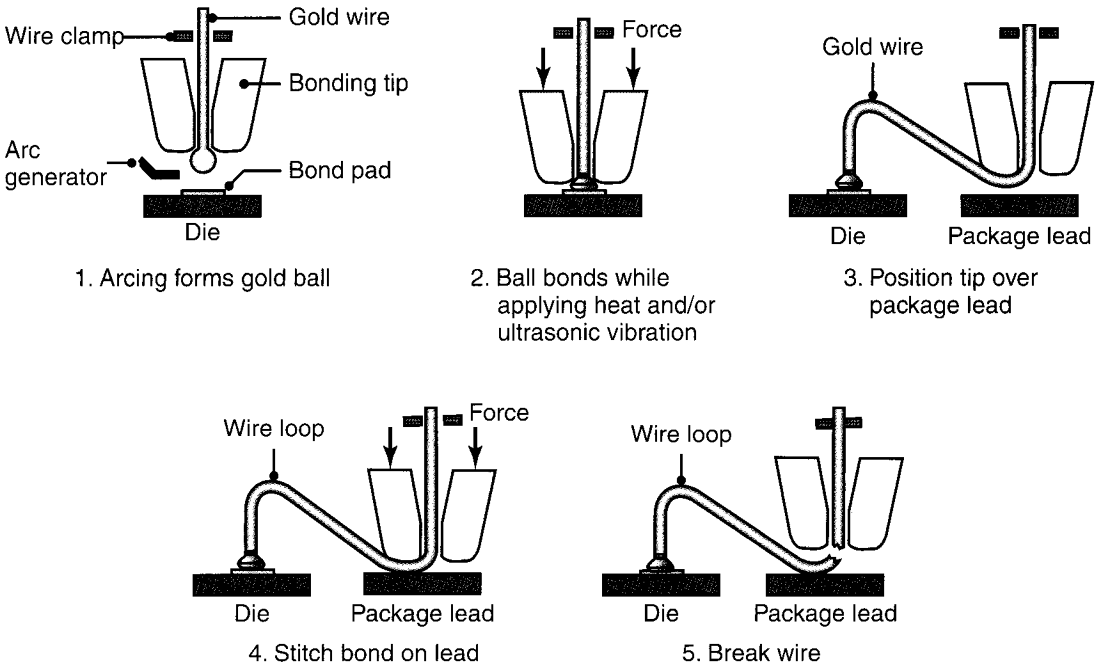


FIGURE 28.30 Schematic illustration of thermosonic welding of gold wires from package leads to bonding pads.

produced on the wafer, but 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 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 temperature resistance.

Packages are produced from polymers, metals, or ceramics. Metal containers are produced from alloys such as Kovar (an iron–cobalt–nickel alloy with a low coefficient of expansion; see Section 3.6), which provide a hermetic seal and good thermal conductivity, but are limited in the number of leads that can be used. Ceramic packages usually are produced from aluminum oxide (Al_2O_3), are hermetic,

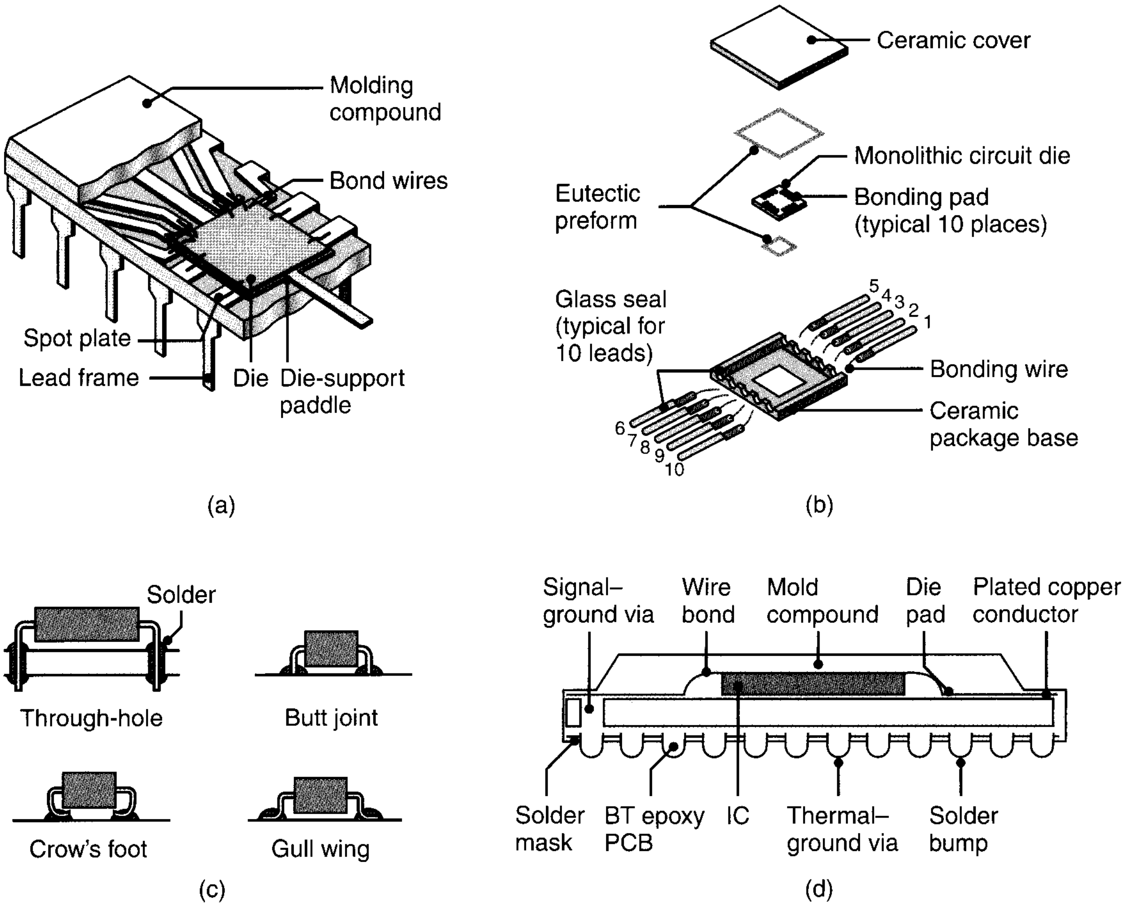


FIGURE 28.31 Schematic illustrations of various IC packages: (a) dual-in-line package (DIP); (b) flat ceramic package; (c) common surface-mount configurations; (d) ball-grid arrays.

and have good thermal conductivity, but have higher lead counts than metal packages; however, they are also more expensive. Plastic packages are inexpensive and have high lead counts, but they have high thermal resistance and are not hermetic.

An older style of packaging is the **dual-in-line package** (DIP), shown schematically in Fig. 28.31a. Characterized by low cost and ease of handling, DIP packages are made of thermoplastics, epoxies, or ceramics and can have from 2 to 500 external leads. Designed for use over a broader temperature range and in high-reliability and military applications, ceramic packages cost considerably more than plastic packages.

Figure 28.31b 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. For that reason, it usually is affixed permanently to a multiple-level circuit board in which the low profile of the flat package is necessary.

Surface-mount packages have become common for today's integrated circuits. Some examples are shown in Fig. 28.31c, where it can be seen 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), while 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, described in 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. However, PGAs and other in-line and surface-mount packages are extremely susceptible to plastic deformation of the wires 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), as shown in Fig. 28.31d. 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 **reflow soldering** (Section 32.3.3), the solder serves to center the BGA by surface tension, resulting in well-defined electrical connections for each ball. After the chip has been sealed in the package, it undergoes final testing. Because one of the main purposes of packaging is isolation from the environment, testing at this stage usually involves heat, humidity, mechanical shock, corrosion, and vibration. Destructive tests also are performed to investigate 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 on the circuit board. 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 its operation.

Flip-chip on Board. The *Flip-chip on Board* (FOB) technology, illustrated in Fig. 28.32, 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 28.33 illustrates the major categories of Silicon in package (SiP) designs. Although these packages can be

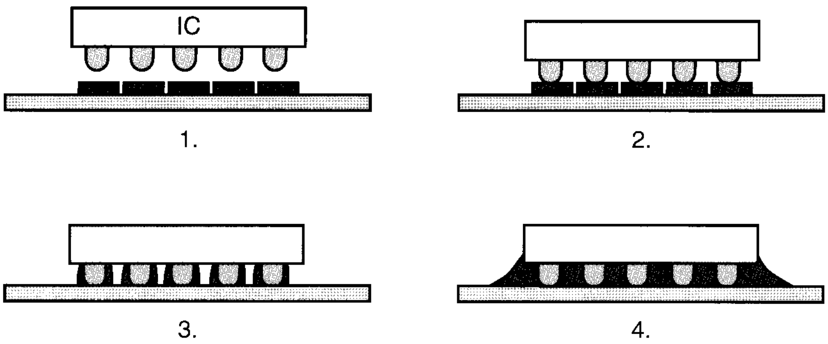


FIGURE 28.32 Illustration of flip-chip technology. Flip-chip package with 1. solder-plated metal balls and pads on the printed circuit board; 2. flux application and placement; 3. reflow soldering; 4. encapsulation.

integrated horizontally, vertical integration through stacked or embedded structures (Figs. 28.33b 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 as well, namely, that

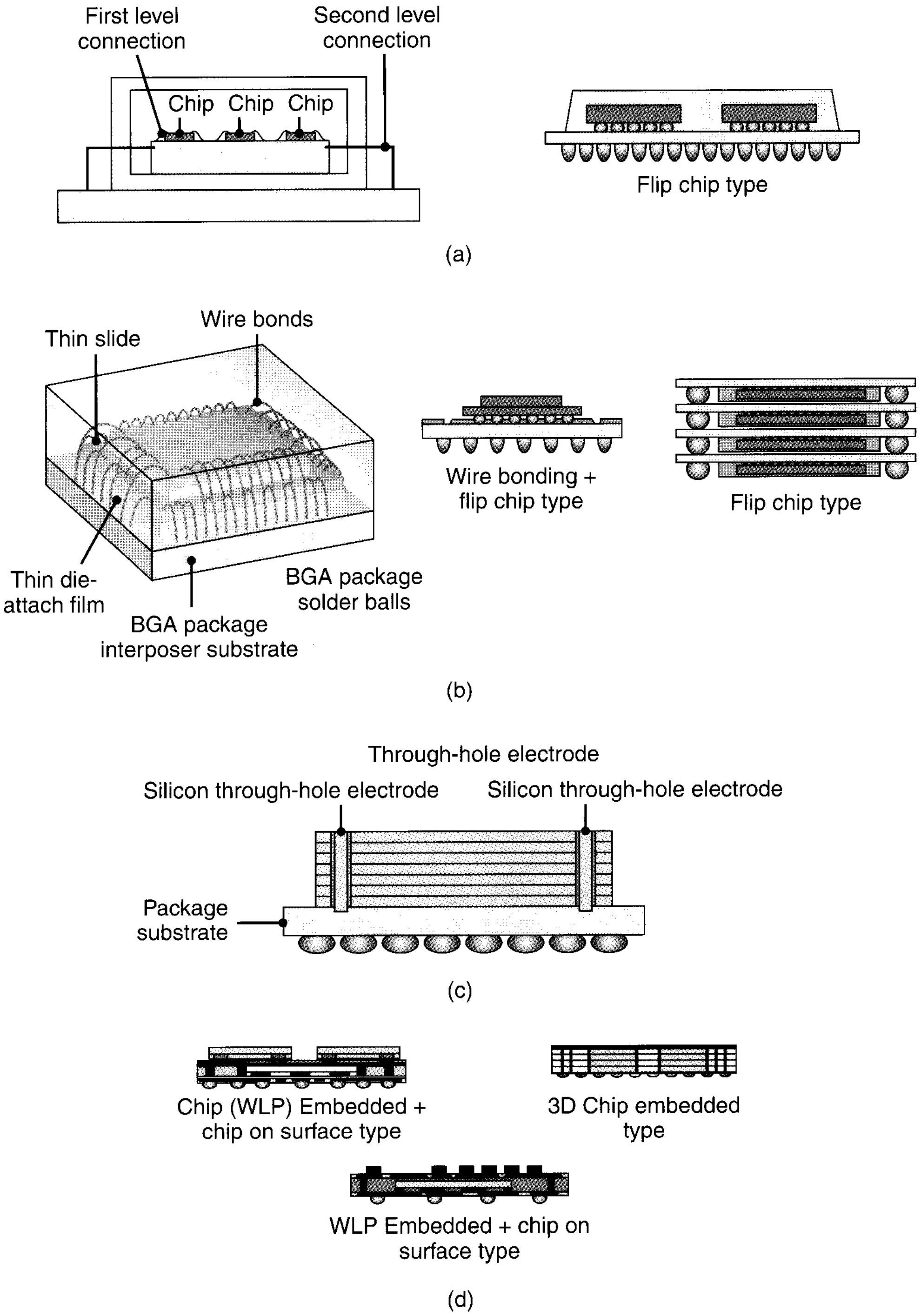


FIGURE 28.33 Major categories of system-in-package designs. (a) Horizontal placement, or multichip modules (MCMs); (b) interposer-type stacked structure; (c) interposerless stacked structure with through-silicon vias; (d) embedded structure. WLP = wafer level package.

they present a reduced size and less noise, and cross talk between chips can be better isolated; also, individual chips can be upgraded more easily. 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, as shown in Fig. 28.33a. To preserve area on a circuit board, chips and/or flip chips can be stacked and bonded to a circuit board, as illustrated in Fig. 28.33b. 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, but it has been noted that this is perhaps a case of 3D integration of a wafer, as shown in Fig. 28.33c.

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, bonding yield, packaging yield, and 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 the etch).

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, so the overall failure rate in entire systems is correspondingly higher.

Equally important in failure analysis is determining the *failure mechanism*—that is, the actual process that causes the device to fail. Common failures due to processing involve the following:

- 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).
- Other failures originating in improper chip mounting, poorly formed wire bonds, or loss of the package's hermetic seal.

Because device lifetimes are very long, it is impractical to study device failure under normal operating conditions. One method of studying failures efficiently is **accelerated life testing**, which 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.

28.13 Printed Circuit Boards

Packaged ICs seldom are used alone; rather, 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 and 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 (which 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. 28.34). **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 must have locations where electrical connectivity is established between the features on both sides of the board; this is accomplished with vias, as shown in Fig. 28.34. 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, 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 (see Section 9.2.1). 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.

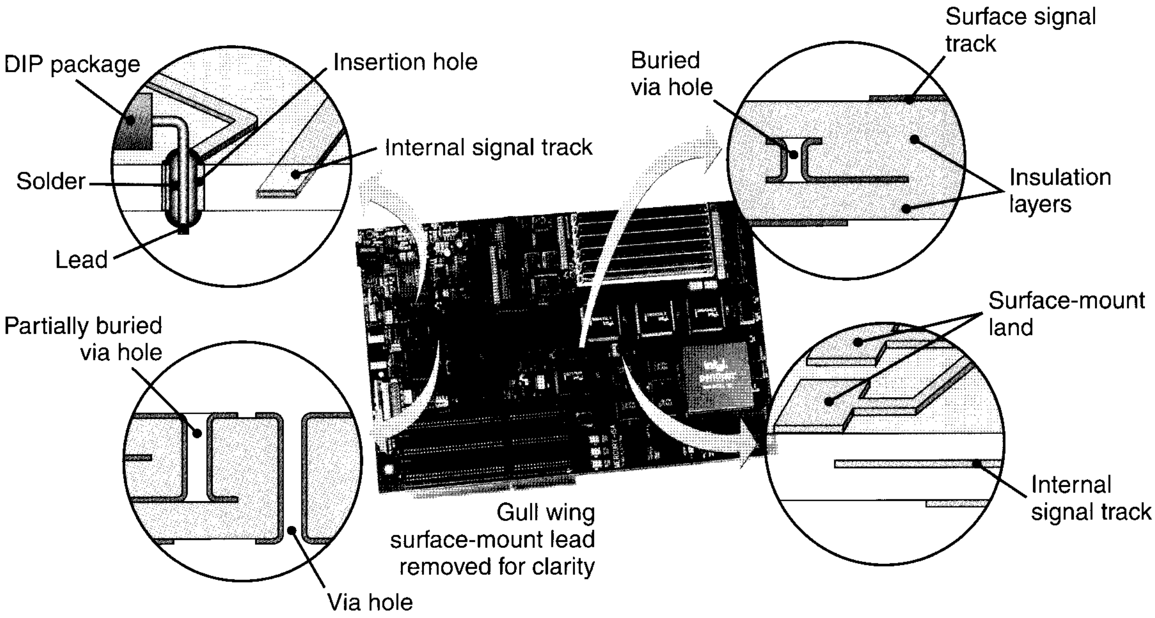


FIGURE 28.34 Printed circuit board structures and design features.

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 be drilled simultaneously to increase production rates.

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 serve to define the connections, tracks, and lands on the circuit board.

The ICs and other discrete components are then fastened to the board by soldering. 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* (see Section 32.3.3 and Example 32.1) are the preferred methods of soldering ICs onto circuit boards.

Some of the design considerations in laying out PCBs are the following:

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. 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.
2. 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.
3. The spacing of ICs is determined mainly by the need to remove heat during the operation. Sufficient clearance between packages and adjacent boards is required to allow forced airflow and heat convection.
4. There should also be sufficient space around each IC package to allow for reworking and repairing without disturbing adjacent devices.

SUMMARY

- The microelectronics industry continues to develop rapidly, and possibilities for new device concepts and circuit designs appear to be endless. The fabrication of microelectronic devices and integrated circuits involves several different types of processes, many of which have been adapted from those of other fields of manufacturing.
- A rough shape of single-crystal silicon is first obtained from 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. However, 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.
- Lastly, the packaged circuit and other discrete devices are soldered to a printed circuit board for final installation.

KEY TERMS

Accelerated-life testing
Bonding
Chemical–mechanical polishing
Chemical-vapor deposition
Chip
Chip on board
Contacts
Czochralski process
Die
Diffusion
Dopants
Dry etching
Dry oxidation
Dual-in-line package
Electromigration
Epitaxy
Etching
Evaporation
Failure rate
Film deposition
Flip-chip on board
Gallium arsenide
Integrated circuit
Ion implantation
Line width
Lithography
Masking
Metal-oxide-semiconductor field-effect transistor
Metallization
Microcontact printing
Micromolding in capillaries
Microtransfer molding
Oxidation
Packaging
Photoresist
Planarization
Postbaking
Prebaking
Printed circuit board
Registration
Reliability
Reticle
SCALPEL
Selective oxidation
Selectivity
Semiconductor
Silicon
Soft lithography
Sputtering
Surface-mount package
System in package
Very large scale integration
Vias
Wafer
Wafer-scale integration
Wet etching
Wet oxidation
Wire bonding
Yield

BIBLIOGRAPHY

Campbell, S.A., **The Science and Engineering of Microelectronic Fabrication**, 2nd ed., Oxford University Press, 2001.

Coombs, C.F., Jr. (ed.), **Printed Circuits Handbook**, 6th ed., McGraw-Hill, 2007.

Doering, R., and Nishi, Y., **Handbook of Semiconductor Manufacturing Technology**, 2nd. ed., CRC Press, 2008.

Harper, C.A. (ed.), **Electronic Packaging and Interconnection Handbook**, 4th ed., McGraw-Hill, 2004.

Humpston, G., and Jacobson, D.M., **Principles of Soldering**, ASM International, 2004.

International Technology Roadmap for Semiconductors, published annually, www.itrs.org.

Jackson, K.A., and Schroter, W. (eds.), **Handbook of Semiconductor Technology, Vol. 2: Processing of Semiconductors**, Wiley, 2000.

Judd, M., and Brindley, K., **Soldering in Electronics Assembly**, 2nd ed., Newnes, 1999.

Khandour, R.S., **Printed Circuit Boards**, McGraw-Hill, 2005.

Lambert, L.P., **Soldering for Electronic Assemblies**, Marcel Dekker, 1988.

Lau, J.H. (ed.), **Electronic Packaging: Design, Materials, Process, and Reliability**, McGraw-Hill, 1998.

Madou, M.J., **Manufacturing Techniques for Microfabrication and Nanotechnology**, CRC Press, 2009.

———, **Applications of Microfabrication and Nanotechnology**, CRC Press, 2009.

Mahajan, S., and Harsha, K.S.S., **Principles of Growth and Processing of Semiconductors**, McGraw-Hill, 1998.

Marks, L., **Printed Circuit Assembly Design**, McGraw-Hill, 2000.

Matisoff, B.S., **Handbook of Electronics Manufacturing**, 3rd ed., Chapman & Hall, 1996.

May, G.S., and Spanos, C.J., **Fundamentals of Semiconductor Manufacturing and Process Control**, Wiley, 2006.
Ohring, M., **Reliability and Failure of Electronic Materials and Devices**, Academic Press, 1999.
Rizvi, S., **Handbook of Photomask Manufacturing Technology**, CRC Press, 2005.
Runyan, W.R., and Shaffner, T.J., **Semiconductor Measurements and Instrumentation**, 2nd ed., McGraw-Hill, 1998.
Schroder, D.K., **Semiconductor Material and Device Characterization**, 3rd ed., Wiley-Interscience, 2006.
Taur, Y., and Ning, T.H., **Fundamentals of Modern VLSI Devices**, Cambridge, 1998.
Tummala, R., Rymaszewski, E.J., and Klopfenstein, A.G., **Microelectronics Packaging Fundamentals**, Chapman & Hall, 1999.
Ulrich, R.K., and Brown, W.D. (eds.), **Advanced Electronic Packaging**, Wiley, 2006.
Van Zant, P., **Microchip Fabrication**, 5th ed., McGraw-Hill, 2004.
Wolf, W.H., **Modern VLSI Design: Systems on Silicon**, 2nd ed., Prentice Hall, 1998
Yeap, G., **Practical Low Power Digital VLSI Design**, Kluwer Academic Pub., 1997.
Yu, P.Y., and Cardona, M., **Fundamentals of Semiconductors: Physics and Materials Properties**, Springer, 2006.

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, and DIP stand for?

28.4. Explain the differences between wet and dry oxidation.

28.5. What are the purposes of prebaking and postbaking in lithography?

28.6. Define selectivity and isotropy and their importance in relation to etching.

28.7. Compare the diffusion and ion-implantation processes.

28.8. Explain the difference between evaporation and sputtering.

28.9. What are the levels of interconnection?

28.10. Which is cleaner, a Class-0.35 or a Class-0.035 clean room?

QUALITATIVE PROBLEMS

28.11. Comment on your observations regarding the contents of Fig. V.1.

28.12. Describe how *n*-type and *p*-type dopants differ.

28.13. How is epitaxy different from other techniques used for deposition? Explain.

28.14. Note that, in a horizontal epitaxial reactor (see Fig. P28.14), the wafers are placed on a stage (susceptor) that is tilted by a small amount, usually 1° to 3°. Explain why this is done.

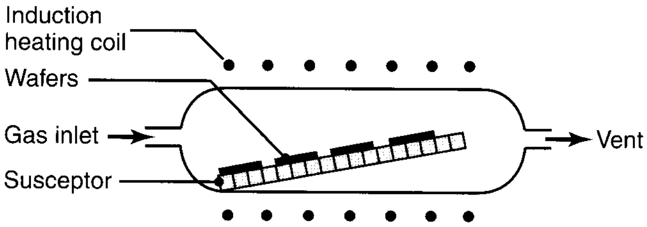


FIGURE P28.14

28.15. 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.

Effects of Manufacturing Changes

Change	Wafer yield	Number of functional chips
Increase wafer diameter		
Reduce chip size		
Increase process complexity		

28.16. 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 (say, ± 0.1 μm), 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?

28.17. 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.

28.18. Examine the hole profiles shown in Fig. P28.18 and explain how they might be produced.

28.19. Referring to Fig. 28.23, sketch the shape of the holes generated from a circular mask.

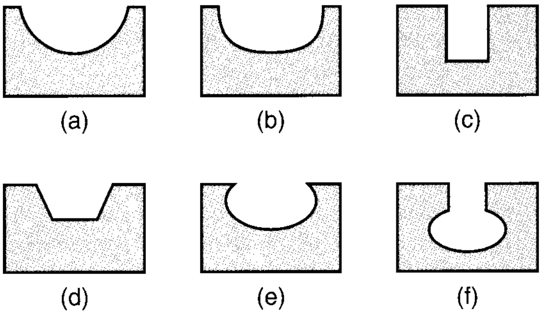


FIGURE P28.18

QUANTITATIVE PROBLEMS

28.20. A certain wafer manufacturer produces two equal-sized wafers, one containing 500 chips and the other containing 200. After testing, it is observed that 50 chips on each wafer are defective. What are the yields of these two wafers? Can any relationship be drawn between chip size and yield?

28.21. A chlorine-based polysilicon etching process displays a polysilicon-to-resist selectivity of 5:1 and a polysilicon-to-oxide selectivity of 60:1. How much resist and exposed oxide will be consumed in etching 3500 Å of polysilicon? What would the polysilicon-to-oxide selectivity have to be in order to reduce the loss to only 40 Å of exposed oxide?

28.22. During a processing sequence, three silicon-dioxide layers are grown by oxidation to 2500 Å, 4000 Å, and 1500 Å, respectively. How much of the silicon substrate is consumed?

28.23. 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 etching is perfectly isotropic)? What would be the minimum photoresist width if a perfectly anisotropic dry-etching process is used?

SYNTHESIS, DESIGN, AND PROJECTS

28.24. Describe products that would not exist today without the knowledge and techniques described in this chapter. Explain.

28.25. 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.

28.26. Describe your understanding of the important features of clean rooms and how they are maintained.

28.27. Make a survey of the necessity for clean rooms in various industries, including the medical, pharmacological, and aerospace industries, and what their requirements are.

28.28. 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.)

28.29. List and discuss the technologies that have enabled the manufacture of the products described in this chapter.

28.30. 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.

28.31. 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.

Fabrication of Microelectromechanical Devices and Systems and Nanoscale Manufacturing

CHAPTER 29

- 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.

Competing processes: Fine blanking, small scale machining, microforming.

EXAMPLES:

CASE STUDIES:

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 millimeters 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

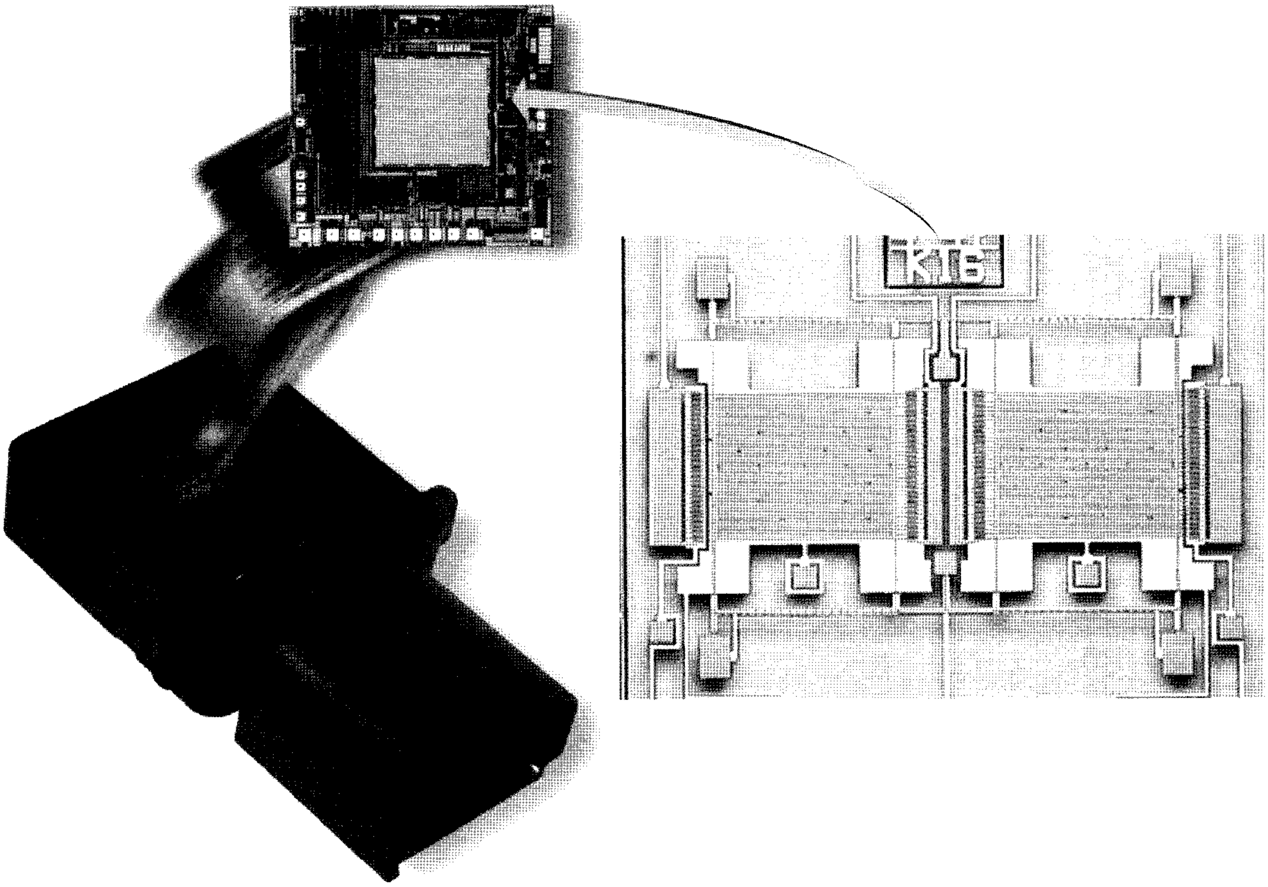


FIGURE 29.1 A gyroscope sensor used for automotive applications. The sensor combines mechanical and electronic systems. Perhaps the most widespread use of MEMS devices is in sensors of all kinds. *Source*: Courtesy of Motorola Corporation.

systems are rarer, but typical examples are air-bag sensors 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 29.5.

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 materials. Regardless, silicon often is used because several highly advanced and reliable manufacturing processes have been developed for microelectronic applications. This chapter emphasizes the manufacturing processes 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 still apply.

The area of MEMS and MEMS devices is rapidly advancing, and new processes or variations on existing processes are continually being developed. Although it has been suggested that MEMS technology can have widespread industrial applications, only a few industries, such as the computer, medical, and automotive industries, have exploited MEMS to date. Many of the processes described in this chapter have not yet become widespread, but are of interest to researchers and practitioners in MEMS.

29.2 Micromachining of MEMS Devices

The topics described in the preceding chapter dealt with the manufacture of integrated circuits and products that operate 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:

a. **Microelectronic devices** are semiconductor-based devices that often have the common characteristics associated with extreme miniaturization and use electrical principles in their design.

b. **Micromechanical devices** are products that are purely mechanical in nature and have dimensions between atomic length scales and a few millimeters. Very small gears and hinges are examples.

c. **Microelectromechanical devices** are products that combine mechanical and electrical or electronic elements at very small length scales. Most sensors are examples of microelectromechanical devices.

d. **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 micrometers to millimeters 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, and other, new processes have been developed that are compatible with the existing processing steps. The use of anisotropic etching techniques allows the fabrication of devices with well-defined walls and high aspect ratios; for this reason, some MEMS devices have been fabricated out of single-crystal silicon.

One of the difficulties associated with the use of silicon for MEMS devices is the high adhesion encountered at small length scales and the associated rapid wear. Most commercial devices are designed to avoid friction by, for example, using flexing springs instead of bushings. 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.

Silicon carbide, diamond, and metals (such as aluminum, tungsten, and nickel) have been investigated as potential MEMS materials. Lubricants also have been investigated. It is known that surrounding the MEMS device, for example, in a silicone oil practically eliminates adhesive wear (Section 33.5), but it also limits the 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.

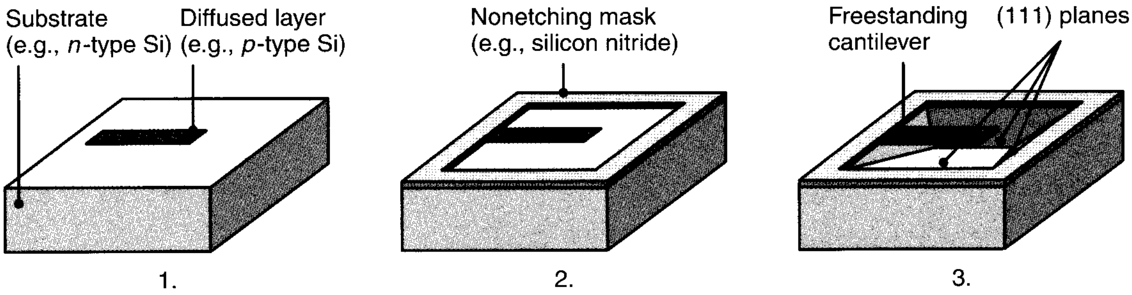


FIGURE 29.2 Schematic illustration of bulk micromachining. 1. Diffuse dopant in desired pattern. 2. Deposit and pattern-masking film. 3. Orientation-dependent etching (ODE) leaves behind a freestanding structure. *Source*: Courtesy of K.R. Williams, Agilent Laboratories.

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 etching down 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 be essentially 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.

29.2.2 Surface Micromachining

Bulk micromachining is useful for producing very simple shapes. It is restricted to single-crystal materials, since polycrystalline materials will not machine at different rates in different directions when wet etchants are used. Many MEMS applications require the use of other materials; hence, alternatives to bulk micromachining are needed. One such alternative is *surface micromachining*. The basic steps in surface micromachining are illustrated for silicon devices in Fig. 29.3. 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. 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 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.

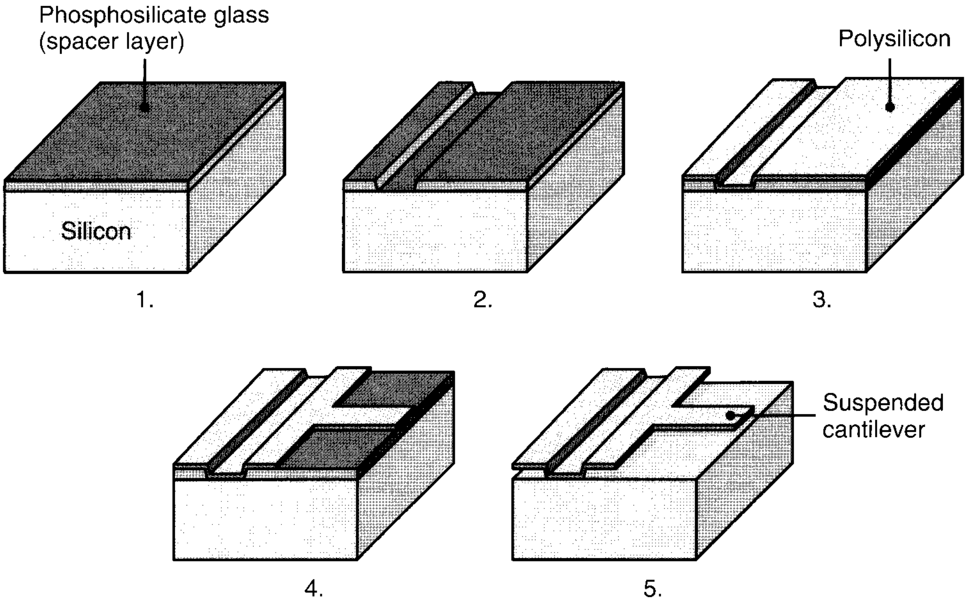


FIGURE 29.3 Schematic illustration of the steps in surface micromachining: 1. Deposition of a phosphosilicate glass (PSG) spacer layer; 2. Lithography and etching of spacer layer; 3. Deposition of polysilicon; 4. Lithography and etching of polysilicon; 5. Selective wet etching of PSG, leaving the silicon substrate and deposited polysilicon unaffected.

Figure 29.4 shows a microlamp that emits a white light when current is passed through it; it has been produced through a combination of surface and bulk micromachining. The top patterned layer is a 2.2-μm-thick 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-HF-etched phosphosilicate glass. The substrate is silicon, which is orientation-dependent etched.

The etchant used to remove the spacer layer must be selected carefully. 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 becomes a very difficult task, 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.

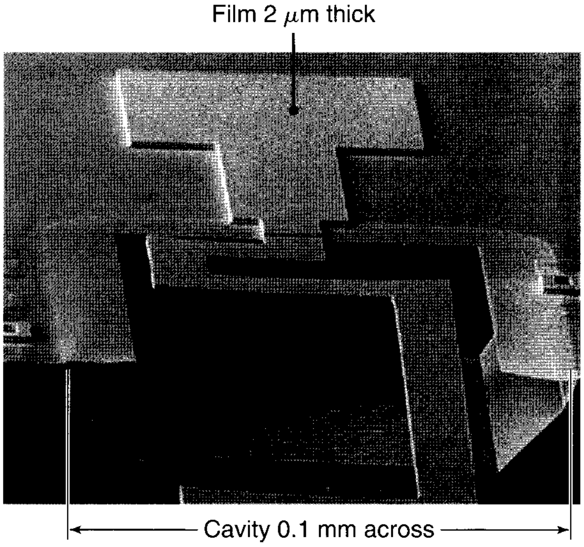


FIGURE 29.4 A microlamp produced from a combination of bulk and surface micromachining processes. *Source*: Courtesy of K.R. Williams, Agilent Technologies.

Another difficulty that must be overcome is **stiction** after wet etching. Consider 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 and results in capillary forces that can deform the film and cause it to contact 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.

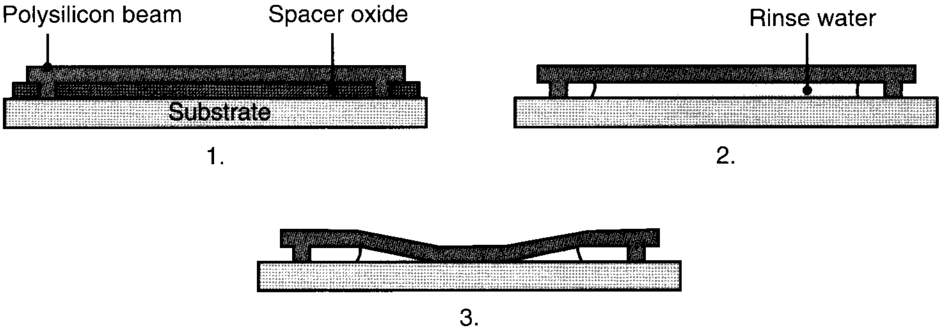


FIGURE 29.5 Stiction after wet etching: 1. Unreleased beam; 2. Released beam before drying; 3. Released beam pulled to the surface by capillary forces during drying. Once contact is made, adhesive forces prevent the beam from returning to its original shape. *Source*: After B. Bhushan.

EXAMPLE 29.1 Surface Micromachining of a Hinge

Surface micromachining is a widespread technology for the production of MEMS. Applications 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.

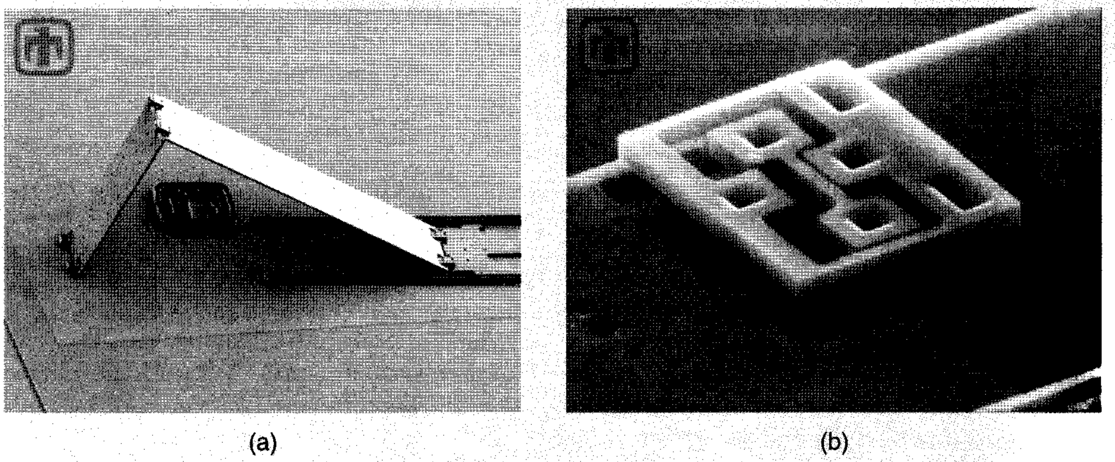

(a)
(b)

FIGURE 29.6 (a) SEM image of a deployed micromirror. (b) Detail of the micromirror hinge. *Source*: Courtesy of Sandia National Laboratories.

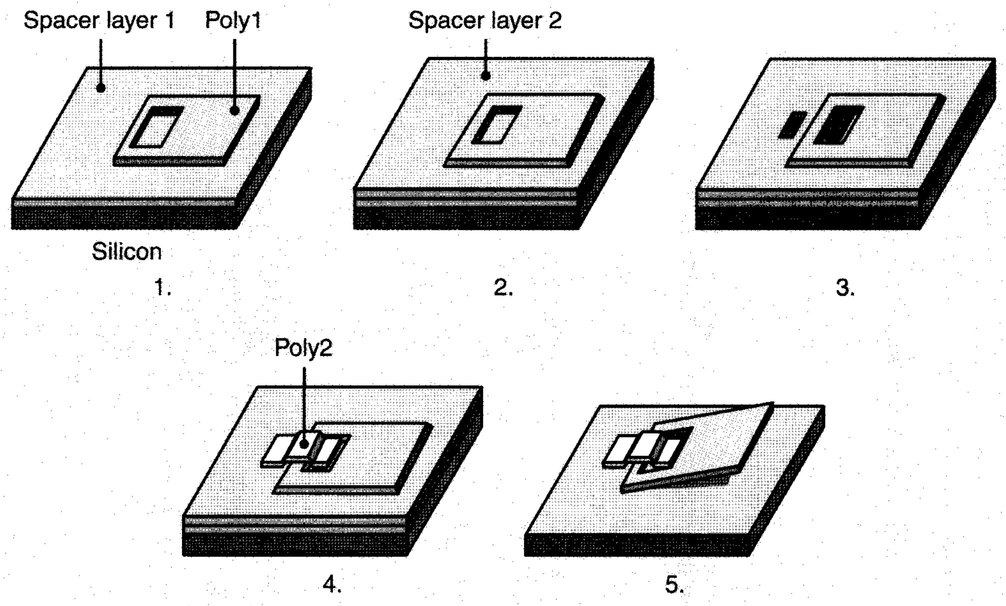


FIGURE 29.7 Schematic illustration of the steps required to manufacture a hinge. 1. Deposition of a phosphosilicate glass (PSG) spacer layer and polysilicon layer (see Fig. 29.3). 2. Deposition of a second spacer layer. 3. Selective etching of the PSG. 4. Deposition of polysilicon to form a staple for the hinge. 5. After selective wet etching of the PSG, the hinge can rotate.

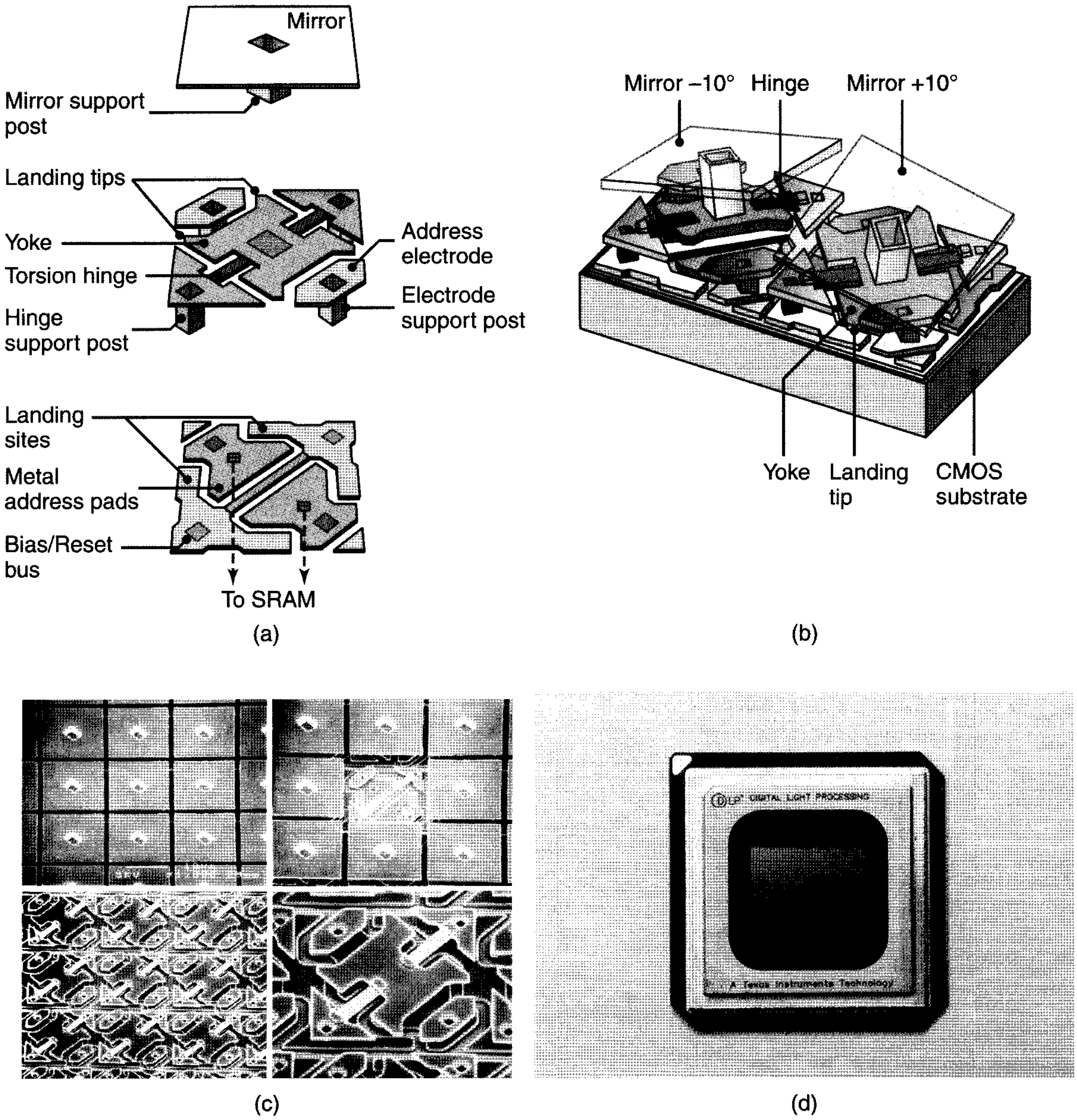
Figure 29.7 shows the cross section of a hinge during its manufacture. The following steps are involved in the production of the hinges:

1. A 2-μm-thick layer of phosphosilicate glass (PSG) 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 removed by wet etching.

Hinges such as these have very high friction. 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 control their deployment precisely.

CASE STUDY 29.1 Digital Micromirror Device

An example of a commercial MEMS-based product is the *digital pixel technology* (DPT™) device, illustrated in Fig. 29.8. This device uses an array of *digital micromirror devices* (DMD™) to project a digital image, as in computer-driven projection systems. The aluminum mirrors can be tilted so that

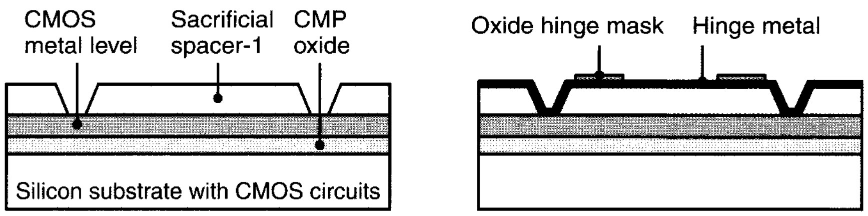


(a) (b)

(c) (d)

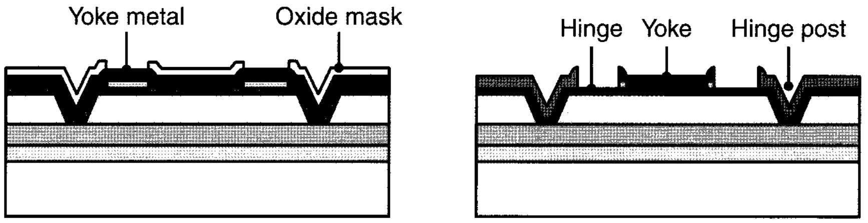
FIGURE 29.8 The Texas Instruments digital pixel technology (DPT™) device. (a) Exploded view of a single digital micromirror device (DMD™). (b) View of two adjacent DMD pixels. (c) Images of DMD arrays with some mirrors removed for clarity; each mirror measures approximately 17 μm on a side. (d) A typical 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.

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.



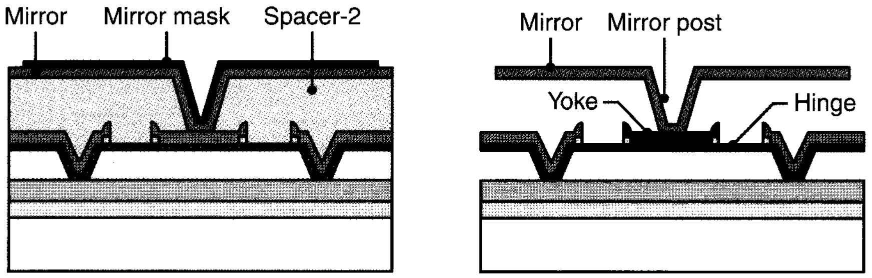
1. Pattern spacer-1 layer

2. Deposit hinge metal; deposit and pattern oxide hinge mask



3. Deposit yoke and pattern yoke oxide mask

4. Etch yoke and strip oxide



5. Deposit spacer-2 and mirror

6. Pattern mirror and etched sacrificial spacers

FIGURE 29.9 Manufacturing sequence for the Texas Instruments DMD device.

The fabrication steps for producing the DMD device are shown in Fig. 29.9. 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—low enough to ensure that no damage occurs to the electronic circuit.
- A thick silicon-dioxide layer is deposited and is chemical–mechanical polished 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. 29.10) 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

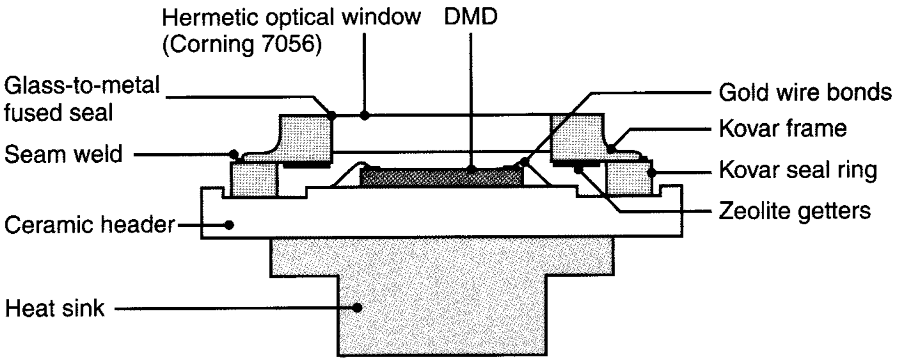


FIGURE 29.10 Ceramic flat-package construction used for the DMD device.

pixel technology is applied widely in digital projection systems, high-definition television, and other optical applications. 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.

SCREAM. Another method for making very deep MEMS structures is the SCREAM (*single-crystal silicon reactive etching and metallization*) process, depicted in Fig. 29.11. In this technique, standard lithography and etching processes produce trenches 10 to 50 μm deep, which are then protected by a layer of chemically vapor

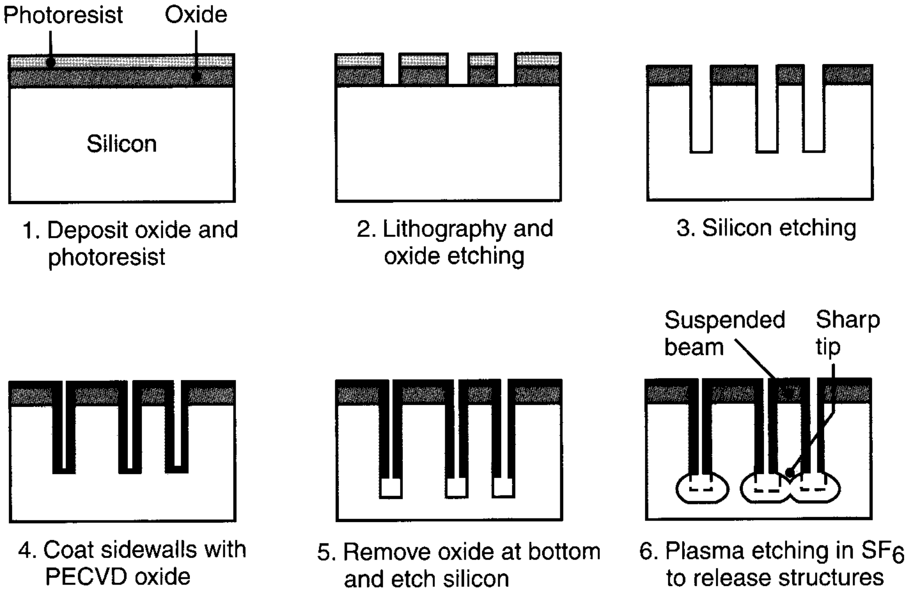


FIGURE 29.11 The SCREAM process. *Source*: After N. Maluf.

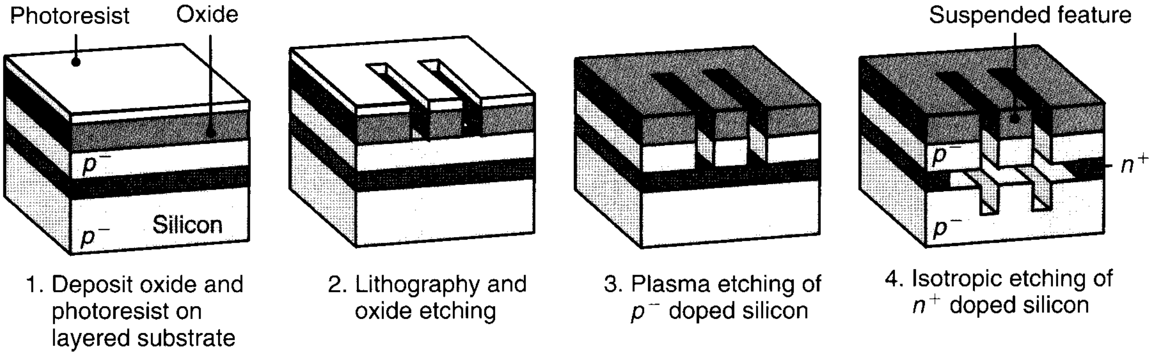


FIGURE 29.12 Schematic illustration of silicon micromachining by single-step plasma etching (SIMPLE) process.

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.

SIMPLE. An alternative to SCREAM is the SIMPLE (*silicon micromachining by single-step plasma etching*) technique, as depicted in Fig. 29.12. 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 as follows:

- 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.

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. 29.13. 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 fusion bonded to the oxide layer, and 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 etched by deep reactive-ion etching to the cavity in the first layer of silicon.

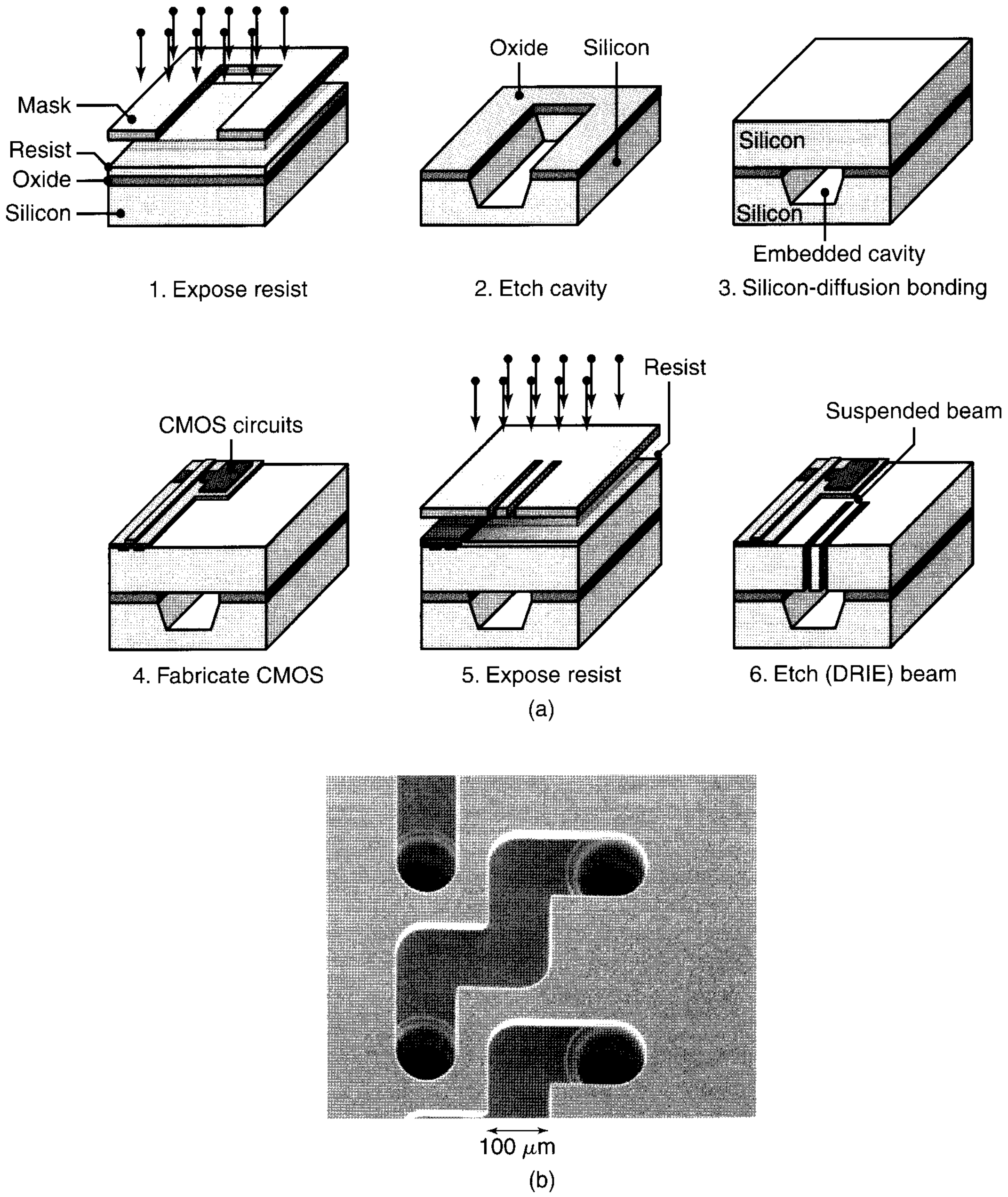


FIGURE 29.13 (a) Schematic illustration of silicon-diffusion bonding combined with deep reactive-ion etching to produce large, suspended cantilevers. (b) A microfluid-flow device manufactured by DRIE etching two separate wafers and then aligning and silicon-fusion bonding them together. Afterward, a 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.

EXAMPLE 29.2 Operation and Fabrication Sequence for a Thermal Ink-jet Printer

Thermal ink-jet printers are perhaps the most successful application of MEMS to date. These printers operate by ejecting nano- or picoliters (10^{-12} liter) of ink from a nozzle towards 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 (dots per inch) requires a nozzle spacing of approximately 20 μm.

The mode of operation of an ink-jet printer is shown in Fig. 29.14. When an ink droplet is to be generated and expelled, a tantalum resistor below a nozzle is heated. This heats a thin film of ink, so that a bubble forms within 5 microseconds, with internal pressures reaching 1.4 MPa. The bubble then expands rapidly, and as a result, 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 removed (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 were made with electroformed nickel nozzles, produced separately from the integrated circuitry, so a bonding operation was required 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. 29.15. A silicon wafer is 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 (95 dots per mm) or higher can be achieved.

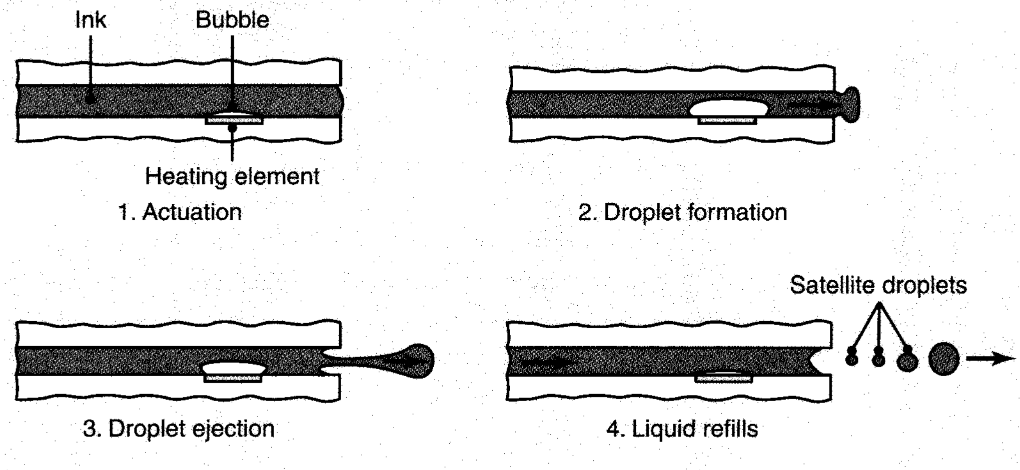


FIGURE 29.14 Sequence of operation of a thermal ink-jet printer. 1. Resistive heating element is turned on, rapidly vaporizing ink and forming a bubble. 2. Within 5 μm, the bubble has expanded and displaced liquid ink from the nozzle. 3. Surface tension breaks the ink stream into a bubble, which is discharged 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) are ejected, and surface tension of the ink draws more liquid from the reservoir. *Source*: After F.-G. Tseng.

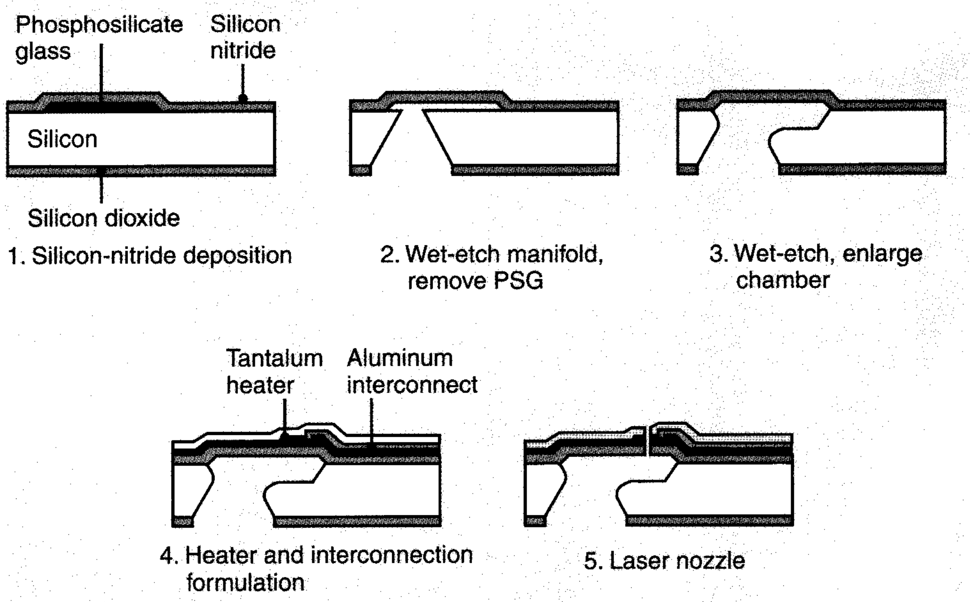


FIGURE 29.15 The manufacturing sequence for producing thermal ink-jet printer heads. *Source*: After F.-G.Tseng.

29.3 The LIGA Microfabrication Process

LIGA is a German acronym for the combined process of X-ray lithography, electrodeposition, and molding (in German, X-ray *l*ithographie, *g*alvanoformung, und *a*bformung). A schematic illustration of this process is given in Fig. 29.16.

The LIGA process involves the following steps:

1. A very thick (up to hundreds of microns) resist layer of polymethylmethacrylate (PMMA) is deposited onto a primary substrate.
2. The PMMA is exposed to columnated X-rays and is developed.
3. Metal is electrodeposited onto the primary substrate.
4. The PMMA is removed or stripped, resulting in a freestanding metal structure.
5. Plastic injection molding takes place.

Depending on the application, the final product from a LIGA process may consist of

- 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

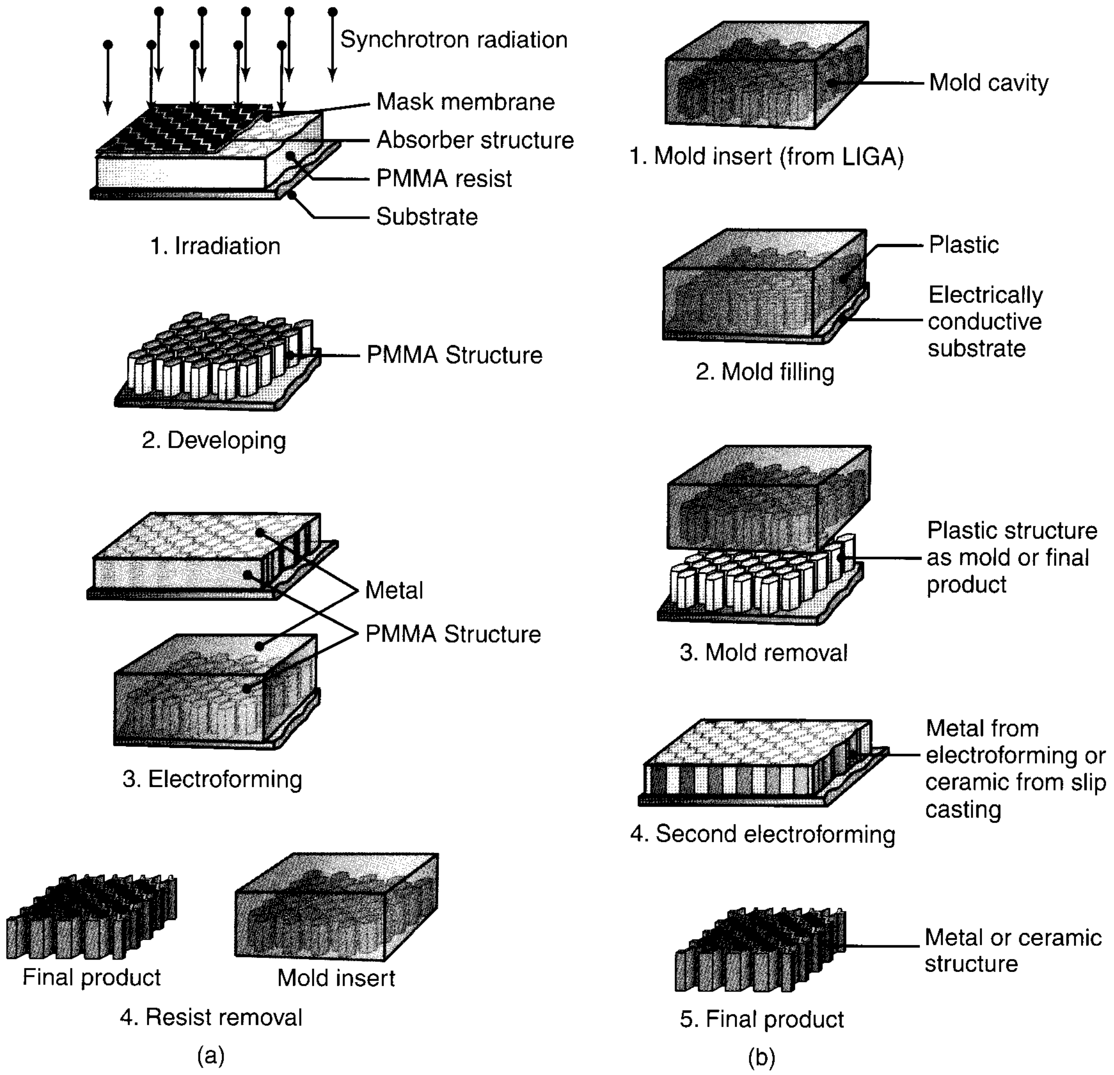


FIGURE 29.16 The LIGA (lithography, electrodeposition and molding) technique. (a) Primary production of a metal final product or mold insert. (b) Use of the primary part for secondary operations or replication. *Source*: Courtesy of IMM Institut für Mikrotechnik, Mainz, Germany.

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 grams per mole). 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.

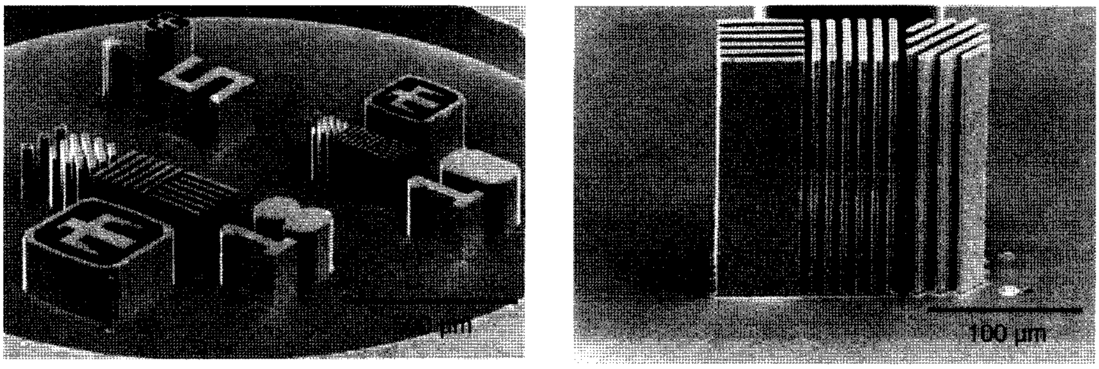
Two newer forms of LIGA are **UV-LIGA** and **Silicon-LIGA**. In *UV-LIGA*, special photoresists are used instead of PMMA, and they are exposed through ultraviolet lithography (Section 28.7). *Silicon-LIGA* uses deep reactive-ion-etched silicon (Section 28.8.2) as a preform for further operations. These 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. The nickel is deposited onto exposed areas of the substrate; it fills the PMMA structure and can even coat the resist (Fig. 29.16a). Nickel is the material of choice 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 has been 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: Conventional lapping leads to preferential removal of the soft PMMA and smearing of the metal. Planarization usually is accomplished with a diamond-lapping procedure 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 used for further processing. Examples of freestanding metal structures produced through the electrodeposition of nickel are shown in Fig. 29.17.

The processing steps used to make freestanding metal structures are extremely 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; 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 these micromolds.



(a) (b)

FIGURE 29.17 (a) Electroformed, 200-μm-tall nickel structures; (b) Detail of 5-μm-wide 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, which are available in powder form. These alloys are of interest because they can produce magnets that are an order of magnitude more powerful than conventional magnets (Table 29.2). Such materials can be used when effective miniature electromagnetic transducers are to be produced.

The processing steps involved in manufacturing these magnets are shown in Fig. 29.18. The PMMA

TABLE 29.2

Comparison of Properties of Permanent-magnet Materials

Material	Energy product (Gauss-Oersted × 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

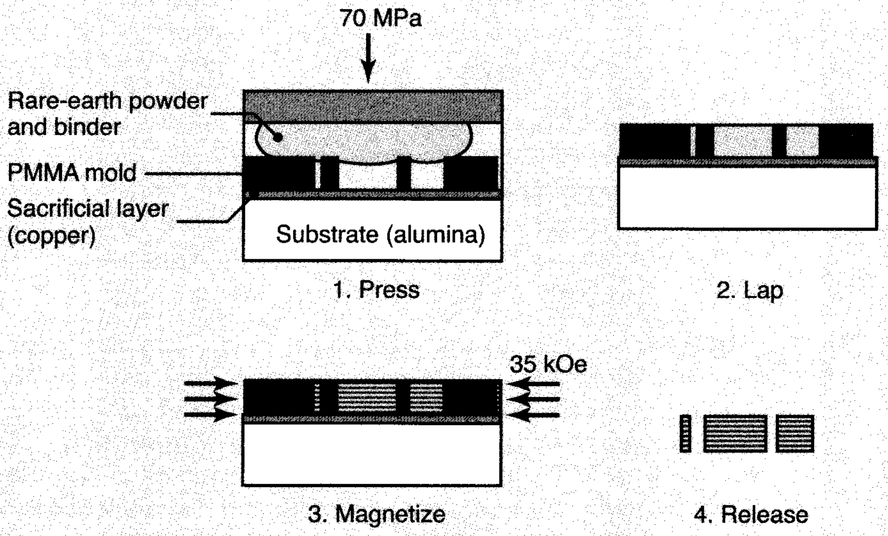


FIGURE 29.18 Fabrication process used to produce rare-earth magnets for microsensors. *Source*: Courtesy of T. Christenson, Sandia National Laboratories.

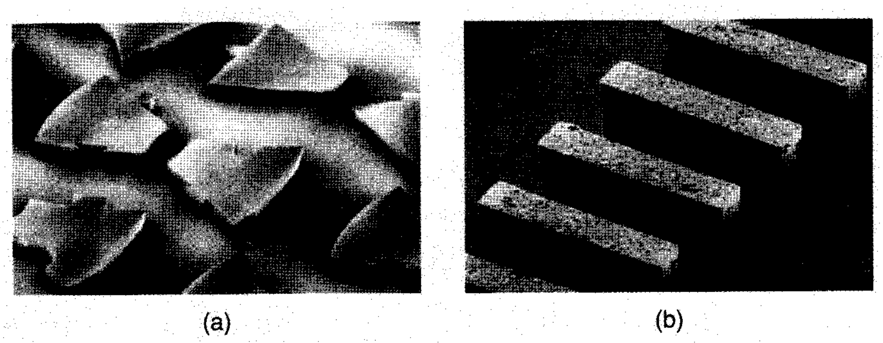


FIGURE 29.19 SEM images of $Nd_2Fe_{14}B$ permanent magnets. The powder particle size ranges from 1 to 5 μm, and the binder is a methylene-chloride-resistant epoxy. Mild distortion is present in the image due to magnetic perturbation of the imaging electrons. Maximum energy products of 9 MGOe have been obtained with this process. *Source*: Courtesy of T. Christenson, Sandia National Laboratories.

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. 19.22) 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. 29.19.

Source: Courtesy of T. Christenson, Sandia National Laboratories.

Multilayer X-ray Lithography. The LIGA technique is very powerful for producing MEMS devices with large aspect ratios and reproducible shapes. It is often useful to obtain a multilayer stepped structure that cannot be made directly through LIGA. For nonoverhanging geometries, direct plating can be applied. In this technique, 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 with 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. 29.20a. 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 one layer bonded to the other. Figure 29.20b shows an example of such a structure.

HEXSIL. The *HEXSIL* process, illustrated in Fig. 29.21, 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 first is 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 leads to good mold filling

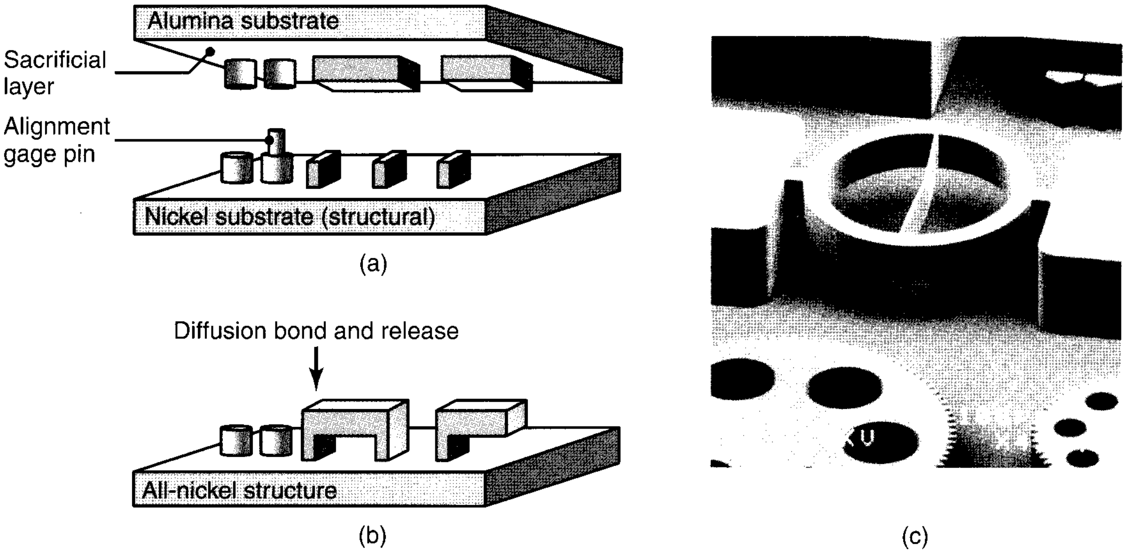


FIGURE 29.20 Multilevel MEMS fabrication through wafer-scale diffusion bonding. (a) Two wafers are aligned and assembled. (b) Resultant structure after diffusion bonding and removal of alumina substrate. (c) A suspended ring structure for measurement of tensile strain, formed by two-layer wafer-scale diffusion bonding. *Source*: (c) Courtesy of T. Christenson, Sandia National Laboratories.

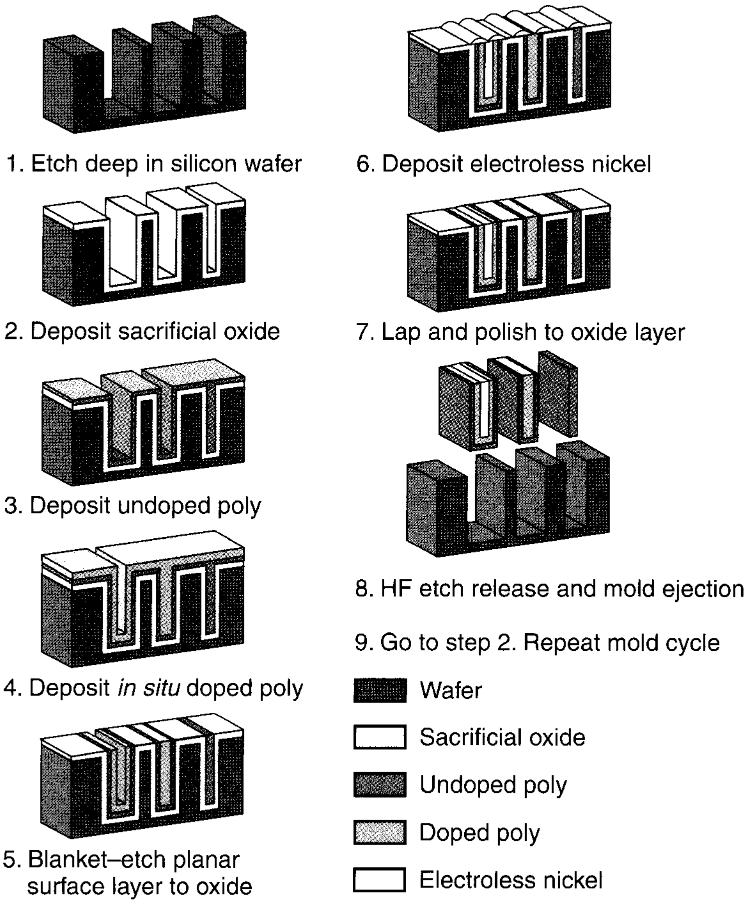


FIGURE 29.21 Illustration of the *hex*agonal honeycomb structure, *sil*icon micromachining, and thin-film deposition—the HEXSIL process.

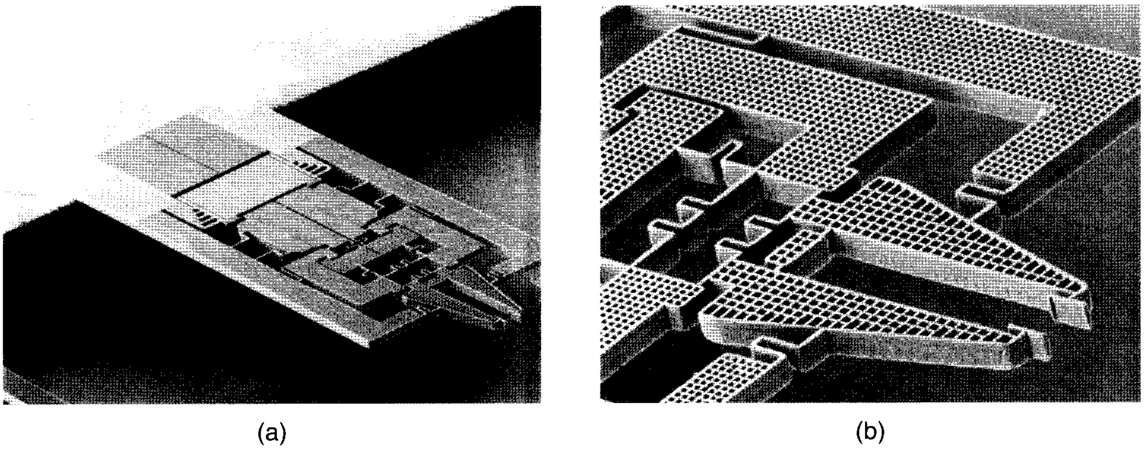


FIGURE 29.22 (a) SEM image of microscale tweezers used in microassembly and microsurgery applications. (b) Detailed view of gripper. *Source*: Courtesy of MEMS Precision Instruments (www.memspi.com).

and good shape definition. A doped-silicon layer follows, providing a resistive portion of the microdevice. Electroplated or electroless nickel plating is then deposited. Figure 29.21 shows various trench widths to demonstrate the different structures that can be produced in HEXSIL.

Microscale tweezers produced through the HEXSIL process are shown in Fig. 29.22. A thermally activated bar activates the tweezers, which have been used for microassembly and microsurgery applications.

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 opposed to material removal. Many of the advances in rapid prototyping also are applicable to MEMS manufacture for processes with sufficiently high resolution. Recall that *stereolithography* (Section 20.3.2) 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 0.05–0.25-mm diameter.

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, 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 needed in microstereolithography, since the smaller structures can be supported by the fluid.
- Parts with significant metal and ceramic content can be produced by suspending nanoparticles in the liquid photopolymer.

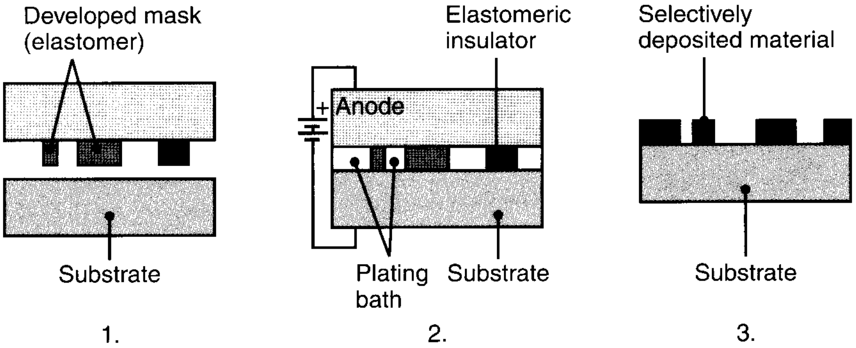


FIGURE 29.23 The instant-masking process: 1. Bare substrate; 2. During deposition, with the substrate and instant mask in contact; 3. The resulting pattern deposited. *Source*: Courtesy of A. Cohen, MEMGen Corporation.

The microstereolithography technique has a number of cost advantages, but the MEMS devices made by this method are difficult to integrate with the controlling circuitry.

Electrochemical Fabrication. Instant masking is a technique for producing MEMS devices (Fig. 29.23). The solid free-form fabrication of MEMS devices using instant masking is known as *electrochemical fabrication* (EFAB). 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.

CASE STUDY 29.2 Accelerometer for Automotive Air Bags

Accelerometers based on lateral resonators represent the largest commercial application of surface micromachining today and are used widely as sensors for automotive air-bag deployment systems. The sensor portion of such an accelerometer is shown in Fig. 29.24. A central mass is suspended over the substrate, but anchored through four slender beams, which act as springs to center the mass under static-equilibrium conditions. An acceleration causes the mass to deflect, reducing or increasing the clearance between the fins on the mass and the stationary fingers on the substrate. By measuring the electrical capacitance between the mass and fins, the deflection of the mass (and therefore the acceleration or deceleration of the system) can be directly measured. Figure 29.24 shows an arrangement for the measurement of acceleration in one direction, but commercial sensors employ several masses so that accelerations can be measured in multiple directions simultaneously.

Figure 29.25 shows the 50-g surface micromachined accelerometer (ADXL-50) with onboard signal conditioning and self-diagnostic electronics. The polysilicon sensing element (visible in the center of the die) occupies only 5% of the total die area, and the whole chip measures 500 × 625 μm. The mass is approximately 0.3 μg, and the sensor has a measurement accuracy of 5% over the ±50-g range.

Fabrication of the accelerometer proved to be a challenge, since it required a *complementary metal-oxide-semiconductor* (CMOS) fabrication sequence to be integrated closely with a surface micromachining approach. Analog Devices, Inc., was able to modify a CMOS production technique to directly incorporate surface micromachining. In the sensor

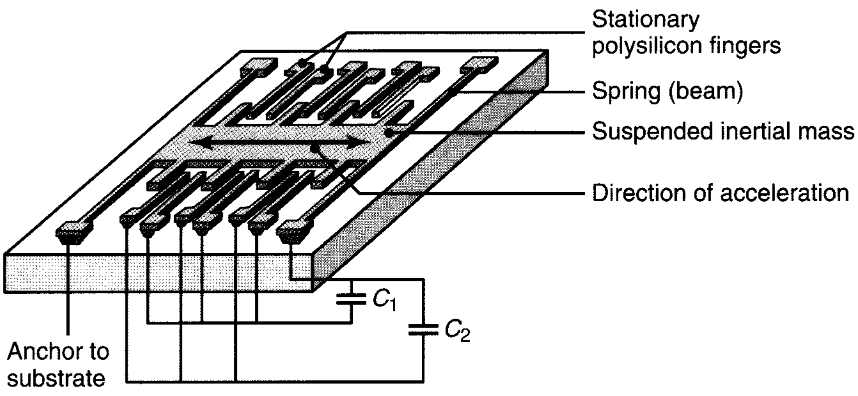


FIGURE 29.24 Schematic illustration of a microacceleration sensor. *Source*: After N. Maluf.

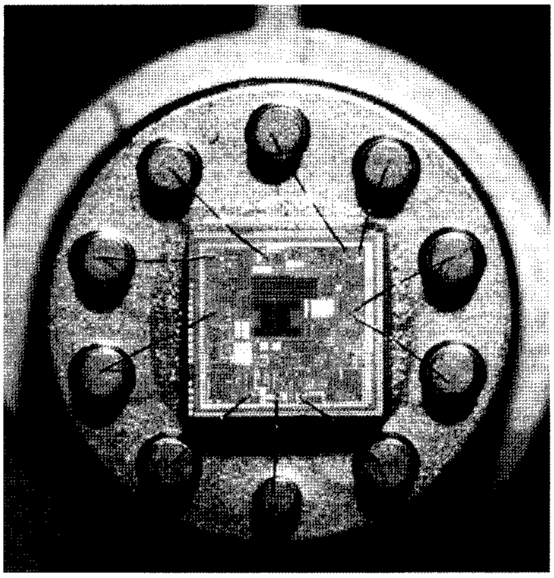


FIGURE 29.25 Photograph of Analog Devices' ADXL-50 accelerometer with a surface micromachined capacitive sensor (center), on-chip excitation, and self-test and signal-conditioning circuitry. The entire chip measures 0.500 × 0.625 mm. *Source*: From R.A. Core, *et al.*, *Solid State Technology*, pp. 39–47, October 1993.

design, $n+$ underpasses connect the sensor area to the electronic circuitry, replacing the usual heat-sensitive aluminum connect lines. Most of the sensor processing is inserted into the fabrication process right after a borosilicate-glass planarization.

After the planarization, a designated sensor region, or *moat*, is cleared in the center of the die (step 1 in Fig. 29.26). A thin oxide is then deposited to passi-

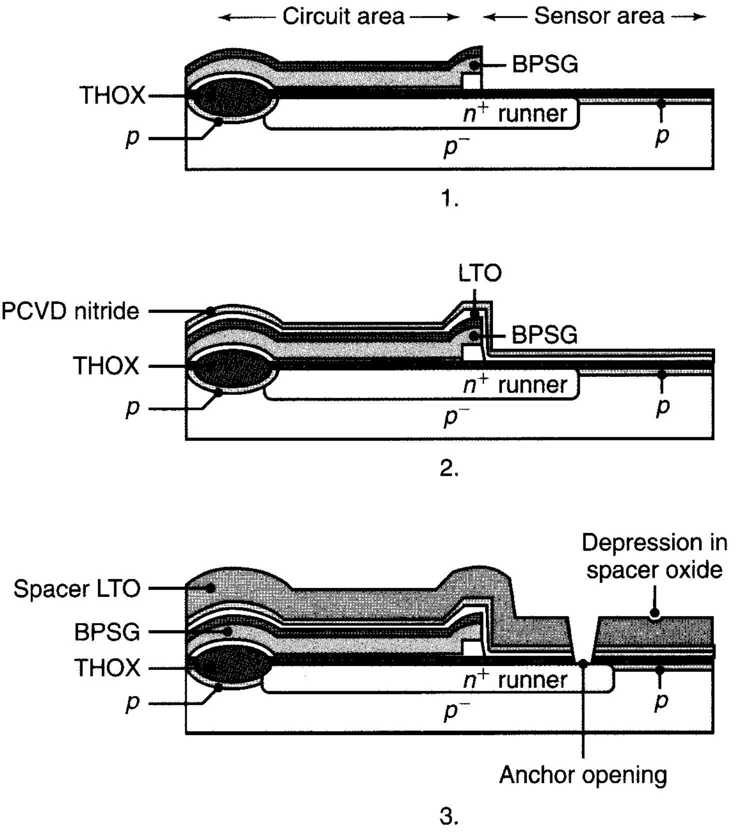


FIGURE 29.26 Preparation of IC chip for polysilicon. 1. Sensor area post-borophosphosilicate glass (BPSG) planarization and moat mask. 2. Blanket deposition of thin oxide and thin nitride layer. 3. Bumps and anchors made in low-temperature oxide (LTO) spacer layer. *Source*: After R.A. Core.

vate the $n+$ underpass connects and is followed by a thin, low-pressure chemical-vapor deposited (LPCVD) nitride to act as an etch stop for the final polysilicon released etching (step 2 in Fig. 29.26). The spacer or sacrificial oxide used is a 1.6-μm densified low-temperature oxide (LTO) deposited over the whole die (step 3 in Fig. 29.26).

In a first etching, small depressions that will form bumps or dimples on the underside of the polysilicon sensor are created in the LTO layer. These bumps will limit adhesive forces and sticking in case the sensor comes in contact with the substrate. A subsequent etching cuts anchors into the spacer layer to provide regions of electrical and mechanical contact (step 3 in Fig. 29.26). The 2-μm thick sensor of polysilicon layer is deposited, implanted, annealed, and patterned (step 1 in Fig. 29.27).

Metallization follows, starting with the removal of the sacrificial spacer oxide from the circuit area along with the LPCVD nitride and LTO layer. A low-temperature oxide is deposited on the polysilicon-sensor part, and contact openings appear in the IC part of the die, where platinum is deposited to form platinum silicide (step 2 in Fig. 29.27). The trimmable thin-film material, TiW barrier metal, and Al–Cu interconnect metal are sputtered on and patterned in the IC area.

The circuit area is then passivated in two sepa-

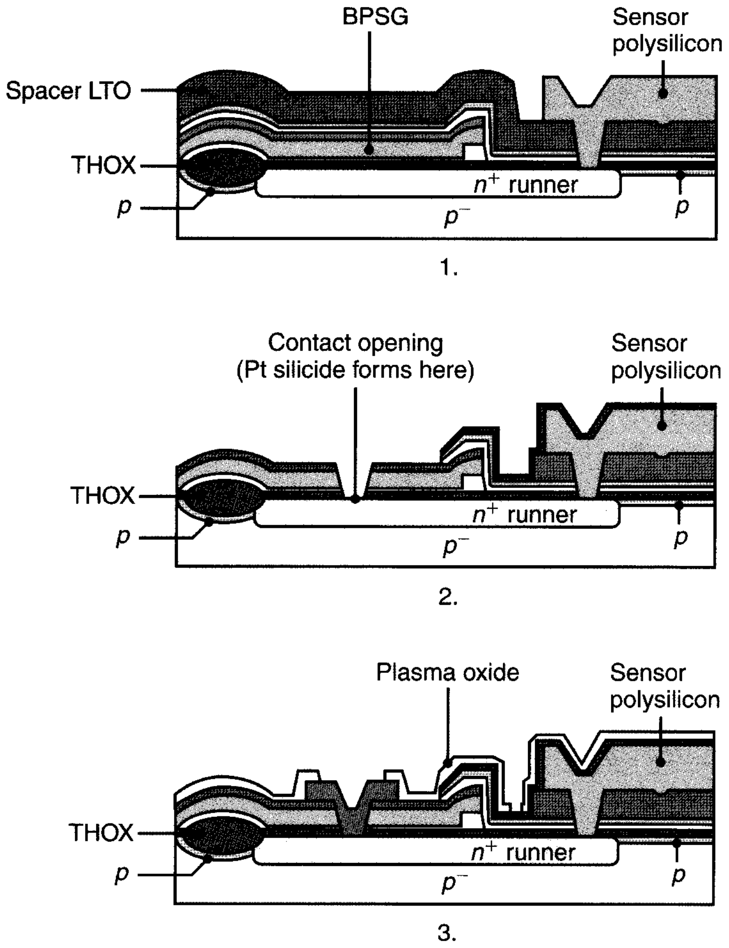


FIGURE 29.27 Polysilicon deposition and IC metallization. 1. Cross-sectional view after polysilicon deposition, implanting, annealing, and patterning. 2. Sensor area after removal of dielectrics from circuit area, contact mask, and platinum silicide. 3. Metallization scheme and plasma-oxide passivation and patterning. *Source*: After R.A. Core.

rate deposition steps. First, plasma oxide is deposited and patterned (step 3 in Fig. 29.27), followed by a plasma nitride (step 1 in Fig. 29.28), to form a seal with the previously deposited LCVD nitride. The nitride acts as a hydrofluoric-acid barrier in the subsequent etch release in surface micromachining. The plasma oxide left on the sensor acts as an etch stop for the removal of the plasma nitride (step 1 in Fig. 29.28). The sensor area is then prepared for the final release etch. The dielectrics are removed from the sensor, and the final protective resist mask is applied. The photoresist protects the circuit area from the long-term buffered oxide etch (step 2 in Fig. 29.28). The final device cross section is shown in step 3 in Fig. 29.28.

Source: Adapted from M. Madou, *Fundamentals of Microfabrication*, 2nd ed., CRC Press, 2002.

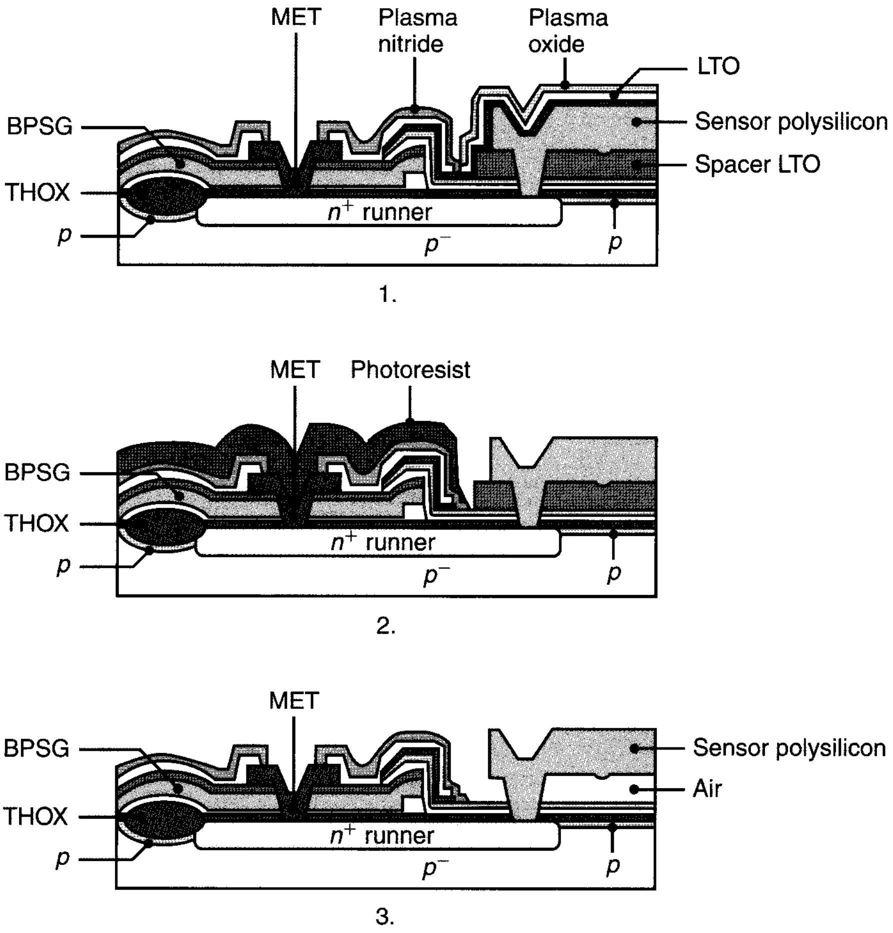


FIGURE 29.28 Prerelease preparation, and release. 1. Post-plasma nitride passivation and patterning. 2. Photoresist protection of the IC. 3. Freestanding, released polysilicon beam. *Source*: After R.A. Core.

29.5 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 with significant manufacturing relevance is. 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 innovation.

Nanoscale manufacturing techniques are outlined in Table 29.3. 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). At the other extreme, **bottom-up** approaches use small building blocks

TABLE 29.3

Comparison of Nanoscale Manufacturing Techniques

Characteristic	Top down			Both top-down and bottom-up	Bottom up	
Nanopatterning technique	Photolithography	Electron beam lithography	Nanoimprint lithography	Dip pen nanolithography	Microcontact printing	Scanning tunneling microscopy
Material flexibility	No	No	No	Yes	Yes	Limited
Resolution	~35 nm	~15 nm	~10 nm	14 nm	~100 nm	Atomic
Registration accuracy	High	High	High	Extremely high	Low	Extremely high
Speed	Very fast	Moderate	Fast	Slower, but scalable	Fast	Very slow
Cycle Time	Weeks	Days	Days–week	Hours	Days–weeks	Days
Cost						
Purchase	>\$10 M	>\$1 M	>\$500 K	<\$250 K	~\$200 K	>\$250 K
Operation	High	High	Moderate	Low	Moderate	Low

Source: Courtesy of NanoInk, Inc.

(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 20.3. However, when placed in the context of nanomanufacturing, bottom-up approaches suggest the manipulation and construction of products on an atomic or molecular scale.

Bottom-up approaches are widely used in nature (e.g., 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 use mainly atomic force microscopy (AFM) for the manipulation of materials on the nanoscale. Figure 29.29 is an illustration of an atomic-force microscope. A probe (Fig. 29.29b) 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 ($<1 \times 10^{-10}$ m).

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 as follows:

- Photolithography, electron-beam lithography, and **nanoimprint lithography**, all using soft lithography techniques, 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. 29.29b has a tip radius around 10 nm. By contacting and plowing across a surface, it can produce grooves up to a few microns thick. The spacing between lines depends on the groove depth needed.

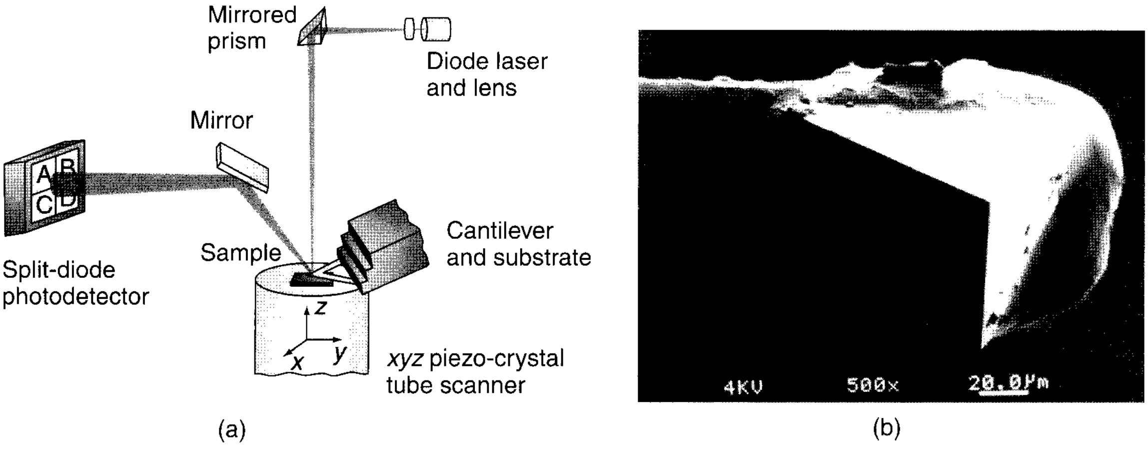


FIGURE 29.29 (a) Schematic illustration of an atomic-force microscope. A probe is mounted on a cylinder containing piezoelectric material; this arrangement allows translation of the probe in three dimensions. A laser reflected from a mirror on the back of the probe onto a set of photosensors allows measurement of the probe's location and monitoring of interactions with a sample surface. (b) Scanning-electron microscope image of a diamond-tipped stainless-steel cantilever suitable for nanolithography.

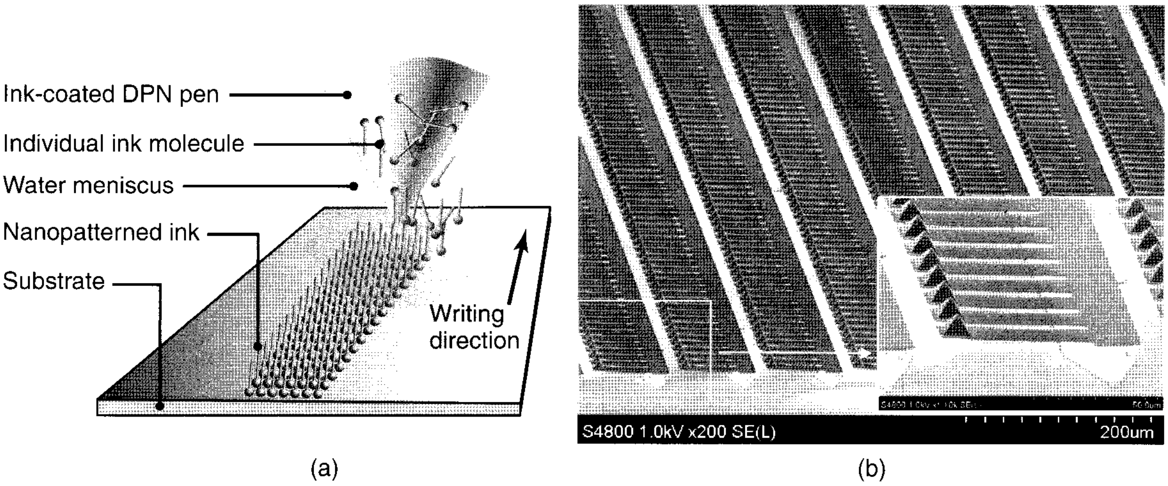


FIGURE 29.30 (a) Schematic illustration of dip pen nanolithography; (b) An array of pens used to produce identical patterns on surfaces. Commercial pen arrays can contain up to 55,000 pens; only a fraction of the available pens are shown. The inset highlights individual cantilevers, showing the 7.5-micrometer-high tips. *Source*: Courtesy of NanoInk, Inc.

- **Dip Pen Nanolithography.** This approach (Fig. 29.30) 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 (Fig. 29.30b), 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 can also 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

- The MEMS field is relatively new and developing rapidly. 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 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

Bulk micromachining
Diffusion bonding
Dip pen nanolithography
EFAB
Electrochemical fabrication
HEXSIL
LIGA
MEMS
Micromachining
Microstereolithography
Multilayer X-ray lithography
Planarization
Sacrificial layer
SCREAM
Silicon-LIGA
SIMPLE
Stiction
Surface micromachining
UV-LIGA

BIBLIOGRAPHY

Allen, J.J., **Microelectromechanical System Design**, CRC Press, 2006.

Berger, L.I., **Semiconductor Materials**, CRC Press, 1997.

Campbell, S.A., **The Science and Engineering of Microelectronic Fabrication**, 2nd ed., Oxford, 2001.

Electronic Materials Handbook, Vol. 1: **Packaging**, ASM International, 1989.

Elwenspoek, M., and Jansen, H., **Silicon Micromachining**, Cambridge University Press, 1998.

Elwenspoek, M., and Wiegerink, R., **Mechanical Microsensors**, Springer, 2001.

Gad-el-Hak, M. (ed.), **The MEMS Handbook**, 2nd ed. (3 vols.), CRC Press, 2006.

Harper, C.A. (ed.), **Electronic Packaging and Interconnection Handbook**, 3rd ed., McGraw-Hill, 2000.

———, **High-Performance Printed Circuit Boards**, McGraw-Hill, 2000.

Hwang, J.S., **Modern Solder Technology for Competitive Electronics Manufacturing**, McGraw-Hill, 1996.

Kovacs, G.T.A., **Micromachined Transducers Sourcebook**, McGraw-Hill, 1998.

Leondes, C.T., **MEMS/NEMS Handbook** (5 vols.), Springer, 2007.

Liu, C., **Foundations of MEMS**, Prentice-Hall, 2006.

Madou, M.J., **Manufacturing Techniques for Microfabrication and Nanotechnology**, CRC Press, 2009.

———, **Applications of Microfabrication and Nanotechnology**, CRC Press, 2009.

Mahajan, S., and Harsha, K.S.S., **Principles of Growth and Processing of Semiconductors**, McGraw-Hill, 1998.

Maluf, N., **An Introduction to Microelectromechanical Systems Engineering**, 2nd ed., Artech House, 2004.

Nishi, Y., and Doering, R. (eds.), **Handbook of Semiconductor Manufacturing Technologies**, Marcel Dekker, 2000.

Pierret, R.F., and Neudeck, G.W., **Semiconductor Device Fundamentals**, Addison-Wesley, 1996.

Quirk, M., and Serda, J., **Semiconductor Manufacturing Technology**, Prentice Hall, 2000.

Rockett, A., **The Materials Science of Semiconductors**, Springer, 2008.

Sze, S.M. (ed.), **Modern Semiconductor Devices Physics**, Wiley, 1997.

———, **Semiconductor Devices: Physics and Technology**, Wiley, 2001.

Taur, Y., and Ning, T.H., **Fundamentals of Modern VLSI Devices**, Cambridge, 1998.

van Zant, P., **Microchip Fabrication: A Practical Guide to Semiconductor Processing**, 4th ed., McGraw-Hill, 2000.

Varadan, V.K., Jiang, X., and Varadan, V., **Microstereolithography and other Fabrication Techniques for 3D MEMS**, Wiley, 2001.

Wise, K.D. (ed.), **Special Issue on Integrated Sensors, Microactuators and Microsystems (MEMS)**, Proceedings of the IEEE, 1998.

REVIEW QUESTIONS

29.1. Define MEMS, SIMPLE, SCREAM, and HEXSIL.

29.2. Why is silicon often used with MEMS devices?

29.3. Describe bulk and surface micromachining.

29.4. What is the purpose of a spacer layer in surface micromachining?

29.5. What is the main limitation to successful application of MEMS?

29.6. What are common applications for MEMS and MEMS devices?

29.7. What is LIGA? What are its advantages?

29.8. What is a sacrificial layer?

29.9. Explain the differences between stereolithography and microstereolithography.

QUALITATIVE PROBLEMS

29.10. Describe the difference between isotropic etching and anisotropic etching.

29.11. 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.

29.12. Which process or processes in this chapter allow the fabrication of products from polymers?

29.13. What is the difference between chemically reactive ion etching and dry-plasma etching?

29.14. 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.

29.15. Explain how you would produce a spur gear if its thickness was one-tenth of its diameter and its diameter was (a) 10 μm, (b) 100 μm, (c) 1 mm, (d) 10 mm, and (e) 100 mm.

29.16. List the advantages and disadvantages of surface micromachining compared with bulk micromachining.

29.17. What are the main limitations to the LIGA process? Explain.

29.18. Other than HEXSIL, what process can be used to make the microtweezers shown in Fig. 29.22? Explain.

QUANTITATIVE PROBLEMS

▶**29.19.** The atomic-force microscope probe shown in Fig. 29.29 has a stainless steel cantilever that is $450 \times 40 \times 2\ \mu\text{m}$. Using equations from solid mechanics, estimate the stiffness of the cantilever, and the force required to deflect the end of the cantilever by $1\ \mu\text{m}$.

▶**29.20.** Estimate the natural frequency of the cantilever in Problem 29.19. *Hint:* See Problem 3.21.

▶**29.21.** Tapping-mode probes for the atomic-force microscope are produced from etched silicon and have typical dimensions of $125\ \mu\text{m}$ in length, $30\ \mu\text{m}$ in width, and $3\ \mu\text{m}$ in thickness. Estimate the stiffness and natural frequency of such probes.

▶**29.22.** 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.

SYNTHESIS, DESIGN, AND PROJECTS

29.23. List similarities and differences between IC technologies described in Chapter 28 and miniaturization technologies presented in this chapter.

29.24. Figure I.8 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.

29.25. Referring to Fig. 29.5, design an experiment to find the critical dimensions of an overhanging cantilever that will not stick to the substrate.

29.26. Design an accelerometer by using (a) the SCREAM process and (b) the HEXSIL process.

29.27. Design a micromachine or device that allows the direct measurement of the mechanical properties of a thin film.

Joining Processes and Equipment

PART VI

When inspecting various common products, note that some products, such as paper clips, nails, steel balls for bearings, staples, screws and bolts, are made of only one component. Almost all products, however, are assembled from components that have been manufactured as individual parts. Even relatively simple products consist of at least two different components joined by various means. For example, (a) some kitchen knives have wooden or plastic handles that are attached to the metal blade with fasteners; (b) cooking pots and pans have metal, plastic, or wooden handles and knobs that are attached to the pot by various methods; and (c) the eraser of an ordinary pencil is attached with a brass sleeve.

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 and a grand piano has 12,000 parts. A C-5A transport plane has more than 4 million parts, and a Boeing 747-400 aircraft has more than 6 million parts. A typical automobile consists of 15,000 components, some of which are shown in Fig. VI.1.

Joining is an all-inclusive term covering processes such as welding, brazing, soldering, adhesive bonding, and mechanical fastening. These processes are an essential

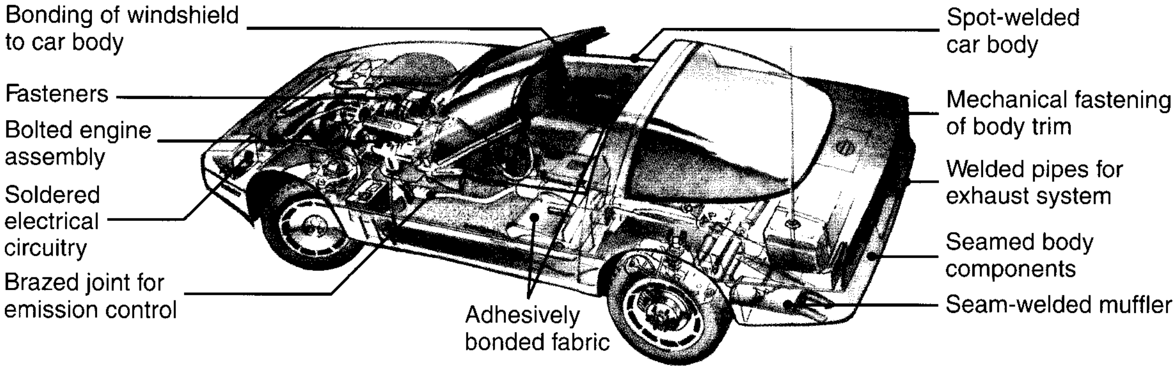


FIGURE VI.1 Various parts in a typical automobile that are assembled by the processes described in Part VI.

and important aspect of manufacturing and *assembly* operations, for one or more of the following reasons:

- 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.
- The product, such as a cooking pot with a handle, is *easier and more economical* to manufacture as individual components, which are then assembled into a product.
- Products such as hair dryers, appliances, and automobile engines need to be designed to be able to be *taken apart* for maintenance or replacement of their parts.
- *Different properties* may be desirable for functional purposes of the product. For example, surfaces subjected to friction, wear, corrosion, or environmental attack generally require characteristics that differ significantly 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, and (c) grinding wheels bonded to a metal backing (Section 26.2).
- *Transporting* the product in individual components and assembling them later may be easier and less costly than transporting the completed item. Metal or wood shelving, large toys, and machinery are assembled after the components or subassemblies have been transported to the appropriate site.

Although there are different ways of categorizing the wide variety of available joining processes, we will follow the classification by the American Welding Society

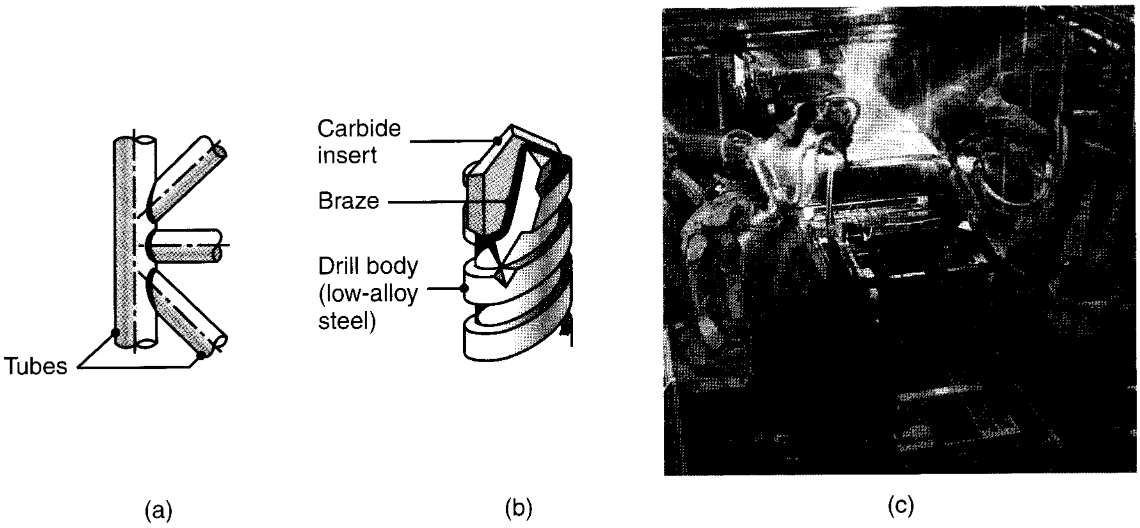


FIGURE VI.2 Examples of parts utilizing joining processes. (a) A tubular part fabricated by joining individual components. This product cannot be manufactured in one piece by any of the methods described in the previous chapters if it consists of thin-walled, large-diameter, tubular-shaped long arms. (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) Courtesy of Ford Motor Co.

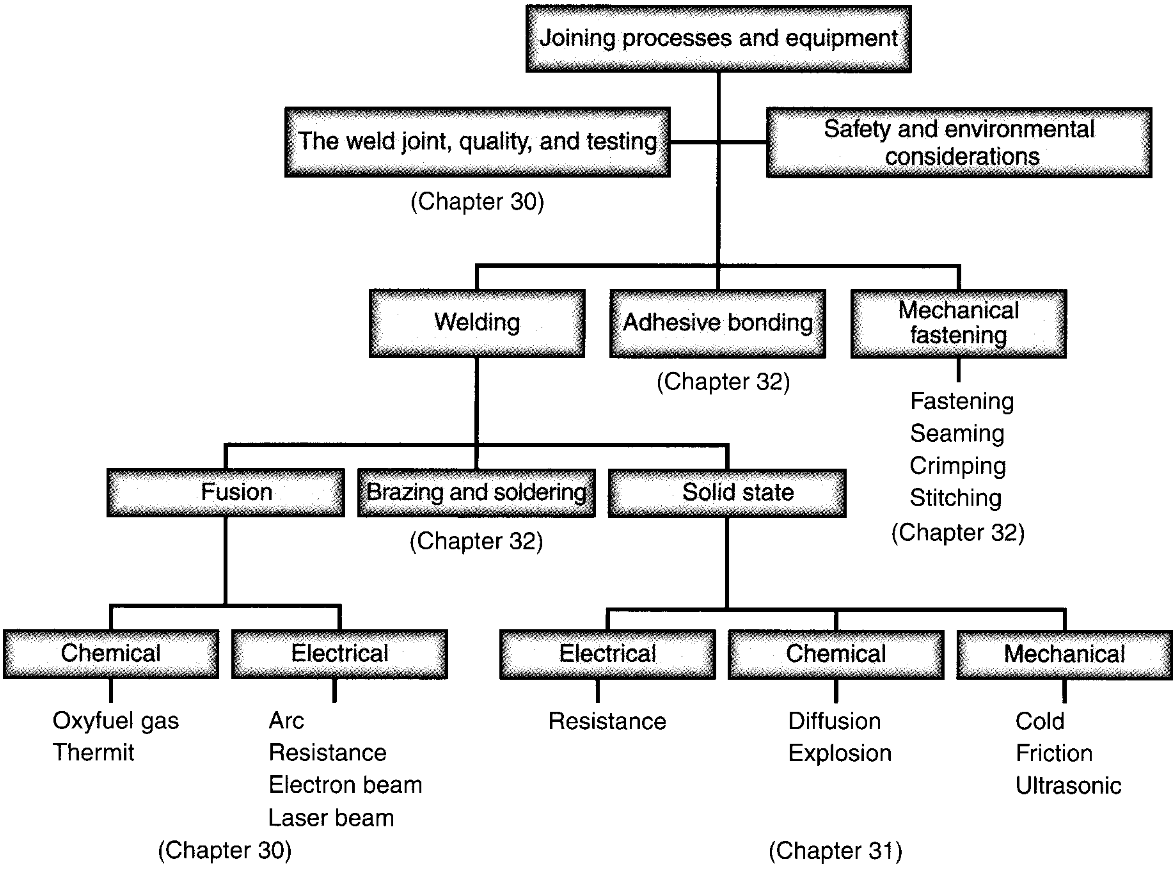


FIGURE VI.3 Outline of topics described in Part VI.

(AWS). Accordingly, joining processes fall into three major categories (see Figs. VI.3 and I.7f):

- **Welding**
- **Adhesive bonding**
- **Mechanical fastening.**

Table VI.1 lists the general relative characteristics of various joining processes. Welding processes, in turn, are generally classified into three basic categories:

- **Fusion welding**
- **Solid-state welding**
- **Brazing and soldering.**

As will be seen, 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 joint undergoes important metallurgical and physical changes, which, in turn, have a major

TABLE VI.1

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
Riveting	1	2	3	1	1	1	3	1	2
Fasteners	2	3	3	1	2	2	2	1	3
Seaming and crimping	2	2	1	3	3	1	3	1	1
Adhesive bonding	3	1	1	2	3	2	3	3	2

Note: 1 = very good; 2 = good; 3 = poor. For cost, 1 is the lowest.

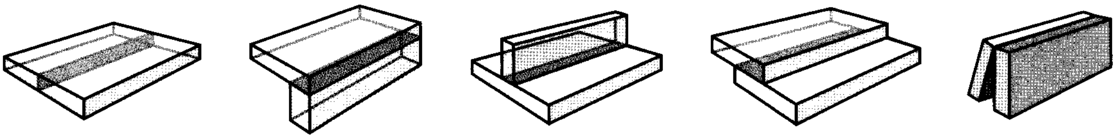

(a) Butt joint (b) Corner joint (c) T joint (d) Lap joint (e) Edge joint

FIGURE VI.4 Examples of joints that can be made through the various joining processes described in Chapters 30 through 32.

effect on the properties and performance of the welded component or structure. Some simple welded joints are shown 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 welding. **Soldering** uses similar filler metals (solders) and involves even lower temperatures.

Adhesive bonding has unique applications that require strength, sealing, thermal and electrical insulating, vibration damping, and resistance to corrosion between dissimilar metals. **Mechanical fastening** involves traditional methods of using various fasteners, especially 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.

Fusion-Welding Processes

CHAPTER 30

- This chapter describes fusion-welding processes, in which two pieces are joined together by the application of heat, which melts and fuses the interface; the operation is sometimes assisted with a filler metal.
- All fusion-welding processes are discussed in this chapter, beginning with oxyfuel–gas welding, in which acetylene and oxygen provide the energy 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 examined 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 discussed, in which electron beams or lasers provide highly focused heat sources.
- The chapter ends with a discussion of the weld joint, including quality, inspection, and testing procedures, along with a discussion of good weld design practices and process selection.

EXAMPLES:

30.1 Introduction

The welding processes described in this chapter involve the partial melting and fusion between two members to be joined. Here, **fusion welding** is defined as *melting together and coalescing* materials by means of heat. *Filler metals*, which are metals added to the weld area during welding, also may be used. Fusion welds made without the use of filler metals are known as *autogenous welds*.

The chapter describes the major classes of fusion-welding processes. It covers the basic principles of each process; the equipment used; the relative advantages, limitations, and capabilities of the process; and the economic considerations affecting process selection (Table 30.1). These processes include the oxyfuel–gas, arc, and high-energy-beam (laser-beam and electron-beam) welding processes, which have important and unique applications in modern manufacturing.

The chapter continues with a description of weld-zone features and the variety of discontinuities and defects that can exist in welded joints. The weldability of

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 2	Medium (5000+)
Gas metal arc	Semiautomatic or automatic	Most metals	Low to high	All	DC	2 to 3	Medium (3000+)
Gas tungsten arc	Manual or automatic	Most metals	Low to high	All	AC, DC	2 to 3	Medium (5000+)
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)

*1 = highest; 5 = lowest.

various ferrous and nonferrous metals and alloys are then reviewed. The chapter ends with a discussion of design guidelines for welding, giving several examples of good weld-design practices. As in all manufacturing processes, the economics of welding is a significant aspect of the overall operation. Welding processes, equipment, and labor costs are discussed in Section 31.8.

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 flame is the source of the heat that is used to melt the metals at the joint. The most common gas-welding process uses *acetylene*; the process is known as *oxyacetylene-gas welding* (OAW) and is typically used for structural metal fabrication and repair work.

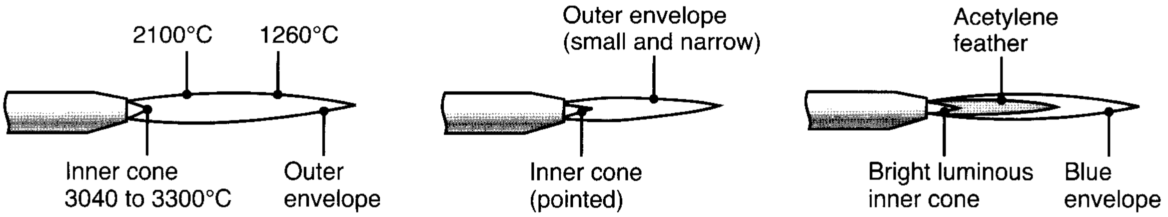
Developed in the early 1900s, OAW utilizes the heat generated by the combustion of acetylene gas (C_2H_2) in a mixture with oxygen. The heat is generated in accordance with a pair of chemical reactions. The primary combustion process, which occurs in the inner core of the flame (Fig. 30.1), involves the following reaction:

$$C_2H_2 + O_2 \longrightarrow 2CO + H_2 + \text{Heat}. \tag{30.1}$$

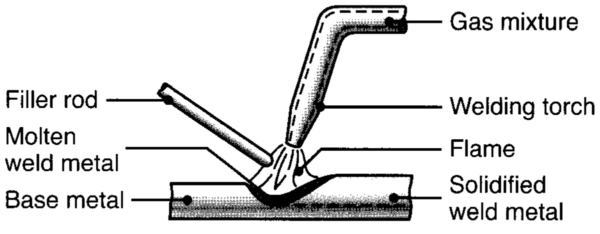
This reaction dissociates the acetylene into carbon monoxide and hydrogen and produces about one-third of the total heat generated in the flame. The secondary combustion process is

$$2CO + H_2 + 1.5O_2 \longrightarrow 2CO_2 + H_2O + \text{Heat}. \tag{30.2}$$

This reaction consists of the further burning of both the hydrogen and the carbon monoxide and produces about two-thirds of the total heat. Note that the reaction also produces water vapor. The temperatures developed in the flame can reach 3300°C.



(a) Neutral flame (b) Oxidizing flame (c) Carburizing (reducing) flame



(d)

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.

Flame Types. 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 a greater oxygen supply, the flame can be harmful (especially for steels), because it oxidizes the metal. For this reason, a flame with excess oxygen is 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** (a flame having excess acetylene; Fig. 30.1c). The temperature of a reducing flame is lower; hence, such a flame is suitable for applications requiring low heat, such as brazing, soldering, and flame-hardening operations.

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. Hence, they are used for welding (a) metals with low melting points (such as lead) and (b) parts that are thin and small. The flame with pure hydrogen gas is colorless; therefore, it is difficult to adjust the flame by eyesight.

Filler Metals. Filler metals are used to supply additional metal to the weld zone during welding. They 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 contributing to the formation of a stronger joint. The slag developed (compounds of oxides, fluxes, and electrode-coating materials) protects the molten puddle of metal against oxidation as it cools.

Welding Practice and Equipment. Oxyfuel–gas welding can be used with most ferrous and nonferrous metals for almost any workpiece thickness, but the relatively low heat input limits the process 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 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 consists basically of a **welding torch** connected by hoses to high-pressure gas cylinders and 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 is an important factor in safety. Oxygen and acetylene cylinders have different threads, so the hoses cannot be connected to the wrong cylinders. The low equipment cost is an attractive feature of oxyfuel–gas welding. Although it can be mechanized, this operation is essentially manual and, hence,

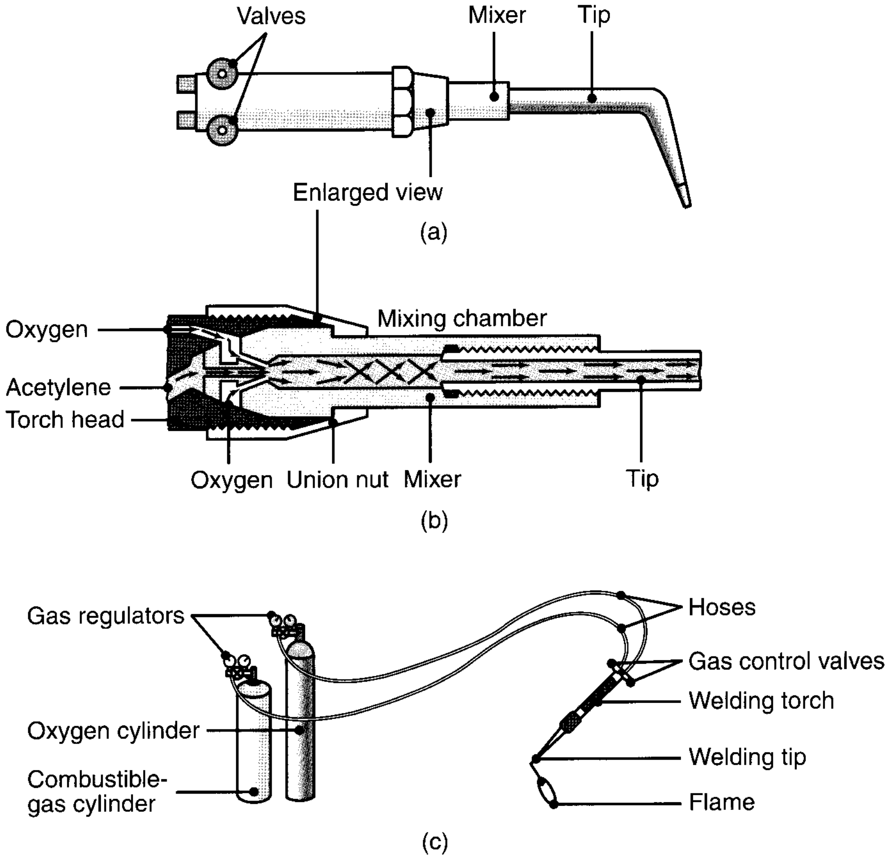


FIGURE 30.2 (a) General view of, and (b) cross section of, a torch used in oxyacetylene welding. The acetylene valve is opened first. The gas is lit with a spark lighter or a pilot light. Then the oxygen valve is opened and the flame adjusted. (c) Basic equipment used in oxyfuel–gas welding. To ensure 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.

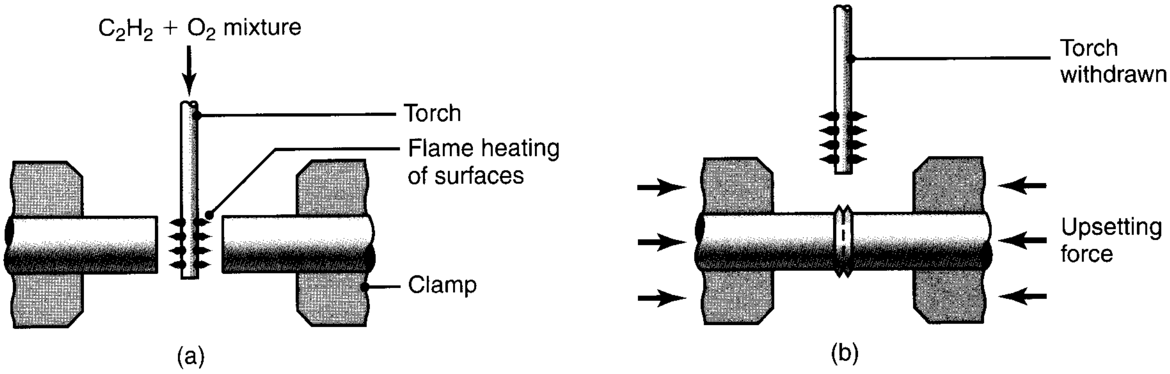


FIGURE 30.3 Schematic illustration of the pressure-gas welding process: (a) before and (b) after. Note the formation of a flash at the joint; later the flash can be trimmed off.

slow. However, it has the advantages of being portable, versatile, and economical for simple and low-quantity work.

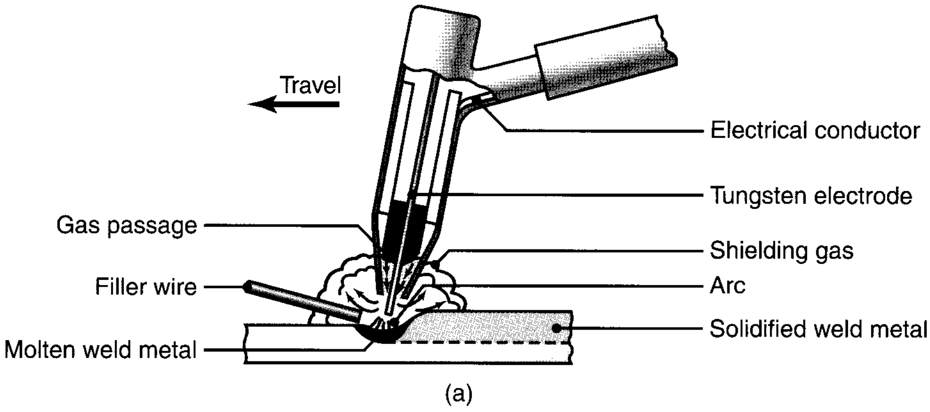
Pressure-gas Welding. In this method, the welding of two components starts with the heating of the interface by means of a torch using (typically) an oxyacetylene–gas mixture (Fig. 30.3a). After the interface begins to melt, the torch is withdrawn. A force is applied to press the two components together (Fig. 30.3b) and is maintained until the interface solidifies. Note the formation of a flash due to the upsetting of the joined ends of the two components.

30.3 Arc-welding Processes: Nonconsumable Electrode

In *arc welding*, developed in the mid-1800s, the heat required is obtained from electrical energy. The process involves either a *consumable* or a *nonconsumable 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, which are much higher than those developed in oxyfuel–gas welding.

In *nonconsumable-electrode* welding processes, the electrode is typically a **tungsten electrode** (Fig. 30.4). Because of the high temperatures involved, an externally supplied shielding gas is necessary to prevent oxidation of the weld zone. Typically, direct current is used, and its **polarity** (the direction of current flow) is important. The selection of current levels depends on such factors as the type of electrode, metals to be welded, and 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. 30.5a). 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 (Fig. 30.5b). Hence, DCEP is preferred for sheet metals and for joints with very 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. 30.5c).



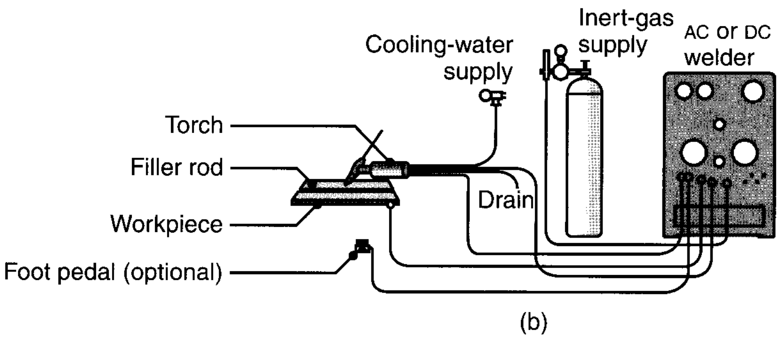


FIGURE 30.4 (a) The gas tungsten-arc welding process, formerly known as TIG (for tungsten–inert-gas) welding. (b) Equipment for gas tungsten-arc welding operations.

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 and varies from around 75% for shielded metal-arc welding to 90% for gas metal-arc welding and submerged-arc welding. The efficiency is an indication that not all of the available energy is beneficially used to melt material, because the heat is conducted through the workpiece, some is lost by radiation, and still more is lost by convection to the surrounding environment.

The heat input given by Eq. (30.3) melts a certain 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 material melted, 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:

$$v = e\frac{VI}{uA}. \tag{30.5}$$

Although these equations have been developed for arc welding, similar ones can be obtained for other fusion-welding operations as well, taking into account differences in weld geometry and process efficiency.

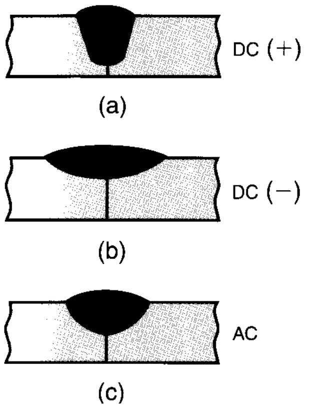


FIGURE 30.5 The effect of polarity and current type on weld beads: (a) DC current with straight polarity; (b) DC current with reverse polarity; (c) AC current.

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

EXAMPLE 30.1 Welding Speed for Different Materials

Consider a situation in which a welding operation is being performed with $V = 20$ volts, $I = 200$ A, and the cross-sectional area of the weld bead is 30 mm^2. Estimate the welding speed if the workpiece and electrode are made of (a) aluminum, (b) carbon steel, and (c) titanium. Use an efficiency of 75%.

Solution For aluminum, we note from Table 30.2 that the specific energy required 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), leading to $v = 10.3$ mm/s. For titanium, $u = 14.3$ J/mm^3, so that $v = 7.0$ mm/s.

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. 30.4a). Because the tungsten electrode is not consumed in this operation, a constant and 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 usually argon or helium (or a mixture of the two). Welding with GTAW may be done without filler metals—for example, in the welding of close-fit joints.

Depending on the metals to be welded, the power supply is either DC at 200 A or AC at 500 A (Fig. 30.4b). 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 the 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. Therefore, contact of the electrode with the molten-metal pool should be avoided.

The GTAW process is used for a wide variety of applications and metals, particularly aluminum, magnesium, titanium, and the refractory metals. It is especially suitable for thin metals. The cost of the inert gas makes this process more expensive than SMAW, but provides welds of very high quality and surface finish. GTAW is used in a variety of critical applications with a wide range of workpiece thicknesses and shapes. The equipment is portable.

Plasma-arc Welding. In *plasma-arc welding* (PAW), developed in the 1960s, a concentrated plasma arc is produced and directed towards the weld area. The arc is stable and reaches temperatures as high as 33,000°C. A **plasma** is an ionized hot gas composed of nearly equal numbers of electrons and ions. The plasma is initiated between the tungsten electrode and the orifice by a low-current pilot arc. What makes plasma-arc welding unlike other processes is that the plasma arc is concentrated because it is forced through a relatively small orifice. Operating currents usually are below 100 A, but they can be higher for special applications. When a filler metal is used, it is fed into the arc, as is done in GTAW. Arc and weld-zone shielding is supplied 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. 30.6a), the workpiece being welded is part of the electrical circuit. The arc transfers from the electrode to the workpiece—hence the term *transferred*.
- In the **nontransferred** method (Fig. 30.6b), the arc occurs 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).

Compared with other arc-welding processes, plasma-arc welding has better arc stability, less thermal distortion, and higher energy concentration, thus permitting deeper and narrower welds. In addition, 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 penetrate completely through the joint (known as the **keyhole technique**), with thicknesses as much as 20 mm for some titanium and aluminum alloys. In 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) often is used for butt and lap joints because of its higher energy concentration, better arc stability, and higher welding speeds. Proper training and skill are essential for operators who use this

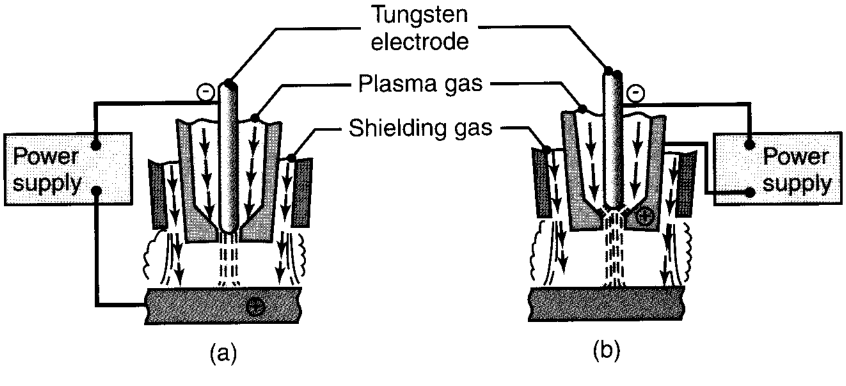


FIGURE 30.6 Two types of plasma-arc welding processes: (a) transferred and (b) nontransferred. Deep and narrow welds can be made by these processes at high welding speeds.

equipment. 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 in a shielding atmosphere of hydrogen gas. The arc is maintained independently of the workpiece or parts being welded. The hydrogen gas normally is diatomic (H_2), but where the temperatures are over 6,000°C near the arc, the hydrogen breaks down into its atomic form, simultaneously absorbing a large amount of heat from the arc. When the hydrogen strikes the cold surface of the workpieces to be joined, it recombines into its diatomic form and rapidly releases the stored heat. The energy in AHW can be varied easily by changing the distance between the arc stream and the workpiece surface. This process is being replaced by shielded metal-arc welding, mainly because of the availability of inexpensive inert gases.

30.4 Arc-welding Processes: Consumable Electrode

There are several consumable-electrode arc-welding processes.

30.4.1 Shielded Metal-arc Welding

Shielded metal-arc welding (SMAW) is one of the oldest, simplest, and most versatile joining processes. About 50% of all industrial and maintenance welding currently is performed by this process. 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. 30.7a). The electrodes are in the shapes of thin, long rods (hence, this process also is known as **stick welding**) that are held manually.

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 the base metal (the workpiece), the electrode metal, and substances from the coating on the electrode; this mixture forms 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 clamped to one terminal of the power source, while the other terminal is connected to the workpiece being welded (Fig. 30.7b). The current, which may be DC or AC, usually ranges from 50 to 300 A.

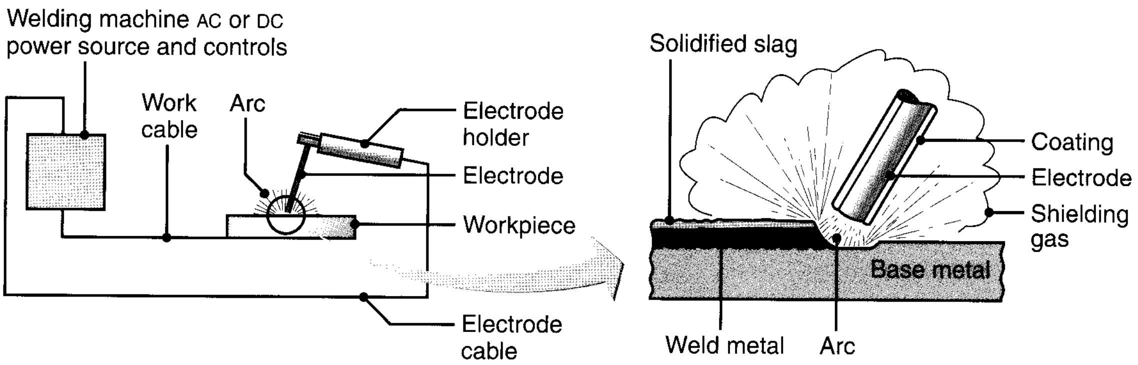


FIGURE 30.7 Schematic illustration of the shielded metal-arc welding process. About 50% of all large-scale industrial-welding operations use this process.

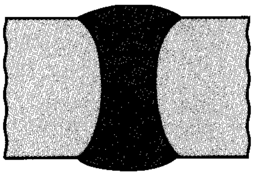


FIGURE 30.8 A deep weld showing the buildup sequence of eight individual weld beads.

For sheet-metal welding, DC is preferred because of the steady arc it produces. Power requirements generally are less than 10 kW.

The SMAW process has the advantages of being relatively simple, versatile, and requiring a smaller variety of electrodes. The equipment consists of a power supply, cables, and an electrode holder. The SMAW process commonly is used in general construction, shipbuilding, pipelines, and maintenance work. It is especially useful for work in remote areas where a portable fuel-powered generator can be used as the power supply. SMAW is best suited for workpiece thicknesses of 3 to 19 mm, although this range can be extended easily by skilled operators using multiple-pass techniques (Fig. 30.8).

The multiple-pass approach requires that the slag be removed after each weld bead. Unless removed completely, the solidified slag can cause severe corrosion of the weld area and lead to failure of the weld, but it also prevents the fusion of weld layers and, therefore, compromises the weld strength. Before another weld is applied, the slag should be removed completely—for example, by wire brushing or weld chipping. Consequently, both labor costs and material costs are high.

30.4.2 Submerged-arc Welding

In *submerged-arc welding* (SAW), the weld arc is shielded by a *granular flux* consisting 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. 30.9). The thick layer of flux completely covers the molten metal. It 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. The unused flux can be recovered (using a recovery tube), treated, and reused.

The consumable electrode is a coil of bare round wire 1.5 to 10 mm in diameter; it is fed automatically through a tube (**welding gun**). Electric currents typically range from 300 to 2000 A. The power supplies usually are 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 largely to welds in a flat or horizontal position having a backup piece. Circular welds can be made on pipes and cylinders—provided that they are rotated during welding. As Fig. 30.9 shows, the unfused flux can be recovered, treated, and reused. SAW is automated

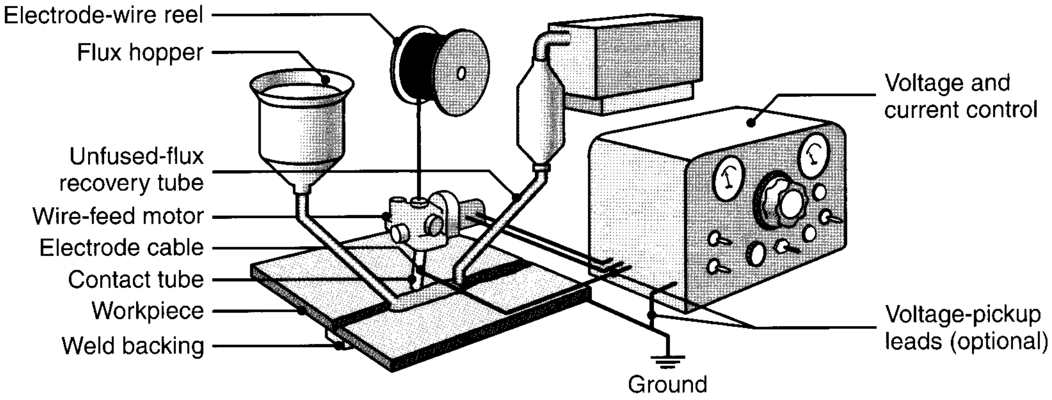


FIGURE 30.9 Schematic illustration of the submerged-arc welding process and equipment. The unfused flux is recovered and reused.

and is used to weld a variety of carbon and alloy steel and stainless-steel sheets or plates at speeds as high as 5 m/min. 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 to 10 times the amount of weld metal per hour as the SMAW process. Typical applications include thick-plate welding for shipbuilding and for 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 effectively inert atmosphere of argon, helium, carbon dioxide, or various other gas mixtures (Fig. 30.10a). The consumable bare wire is fed automatically through a nozzle into the weld arc by a wire-feed drive motor (Fig. 30.10b). In addition to using inert shielding gases, deoxidizers usually are present in the electrode metal itself in order to prevent oxidation of the molten-weld puddle. Multiple-weld layers can be deposited at the joint.

Metal can be transferred by three methods in the GMAW process:

1. In **spray transfer**, small, molten metal droplets from the electrode are transferred to the weld area at a rate of several hundred droplets per second. The

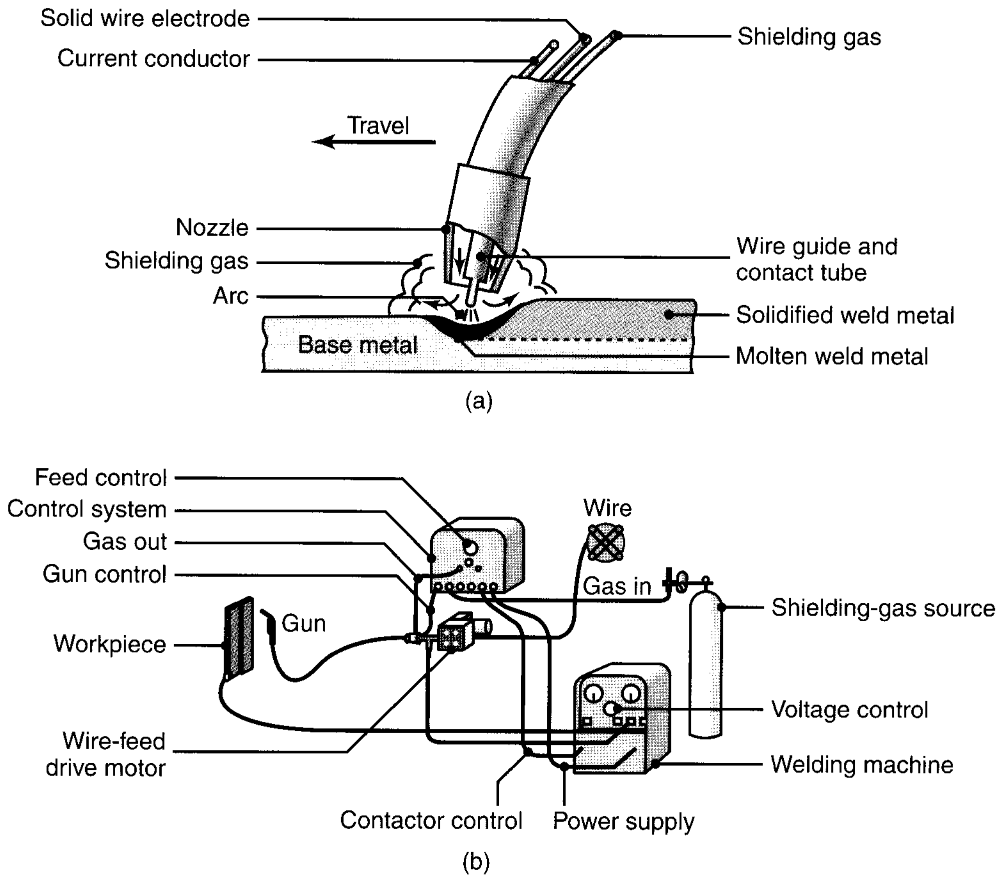


FIGURE 30.10 (a) Schematic illustration of the gas metal-arc welding process, formerly known as MIG (for metal inert-gas) welding. (b) Basic equipment used in gas metal-arc welding operations.

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 in this process can be reduced with the use of a **pulsed arc**, which superimposes high-amplitude pulses onto a low, steady current. The process can be used in all welding positions.

2. In **globular transfer**, carbon-dioxide-rich gases are utilized, and globules are propelled by the forces of the electric-arc transfer of the metal, resulting in considerable spatter. High welding currents are used, making it possible for greater weld penetration and higher welding speed than are achieved in spray transfer. Heavier sections commonly are joined by this method.
3. In **short circuiting**, the metal is transferred in individual droplets (more than 50 per second) 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. The power required 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; otherwise incomplete fusion may occur. The operation, which is easy to perform, is commonly used for welding ferrous metals in 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 relatively simple nature of the process, the training of operators is easy. The process is versatile, rapid, and economical, and welding productivity is double that of the SMAW process. The GMAW process can be automated easily and lends itself readily to robotics and to flexible manufacturing systems (see Chapters 37 and 39).

30.4.4 Flux-cored Arc Welding

The *flux-cored arc welding* (FCAW) process, illustrated in Fig. 30.11, is similar to gas metal-arc welding, except that the electrode is tubular in shape and is filled with

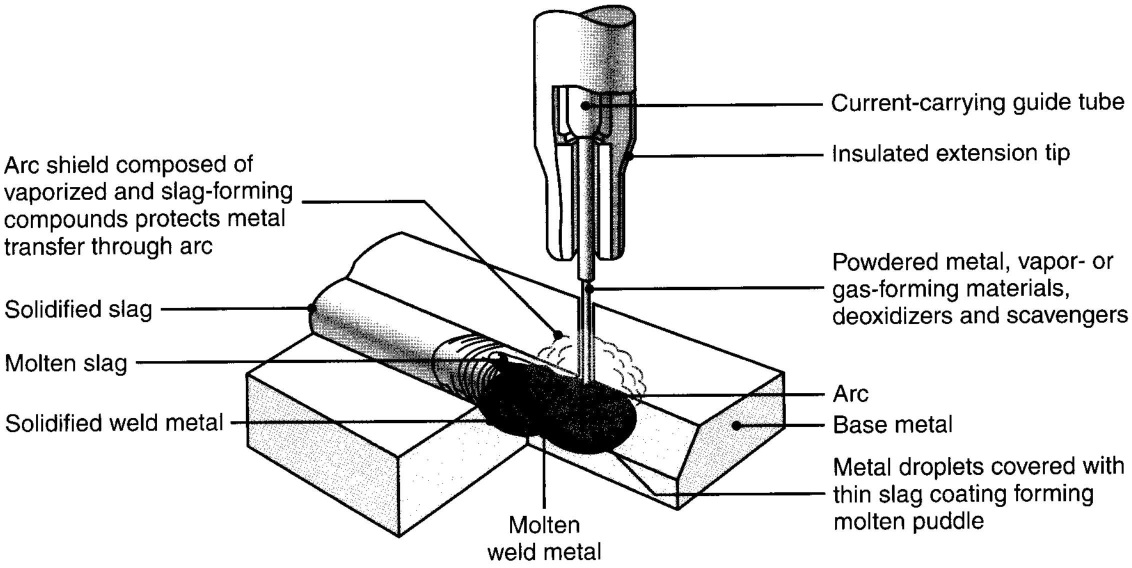


FIGURE 30.11 Schematic illustration of the flux-cored arc-welding process. This operation is similar to gas metal-arc welding, shown in Fig. 30.10.

flux (hence the term *flux-cored*). Cored electrodes produce a more stable arc, improve weld contour, and produce better mechanical properties of the weld metal. The flux in these electrodes is much more flexible than the brittle coating used on SMAW electrodes, so the tubular electrode can be provided in long coiled lengths.

The electrodes are usually 0.5 to 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 the welding of thinner materials not only possible, but often preferable. Also, small-diameter electrodes make it relatively easy to weld parts in different 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, so it is used for welding a variety of joints, mainly on steels, stainless steels, and nickel alloys. The higher weld-metal deposition rate of the FCAW process (compared with that of GMAW) has led to its use in the joining of sections of all thicknesses. The use of *tubular electrodes* with very small diameters has extended the use of this process to workpieces of smaller section size.

A major advantage of FCAW is the ease with which specific weld-metal chemistries can be developed. By adding alloying elements to the flux core, virtually any alloy composition can be produced. 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 used primarily for welding the edges of sections vertically and in one pass with the pieces placed edge to edge (butt joint). It is classified as a *machine-welding process*, because it requires special equipment (Fig. 30.12). The weld metal is deposited into a weld cavity between the two pieces to be joined. The space 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, with the workpiece rotating.

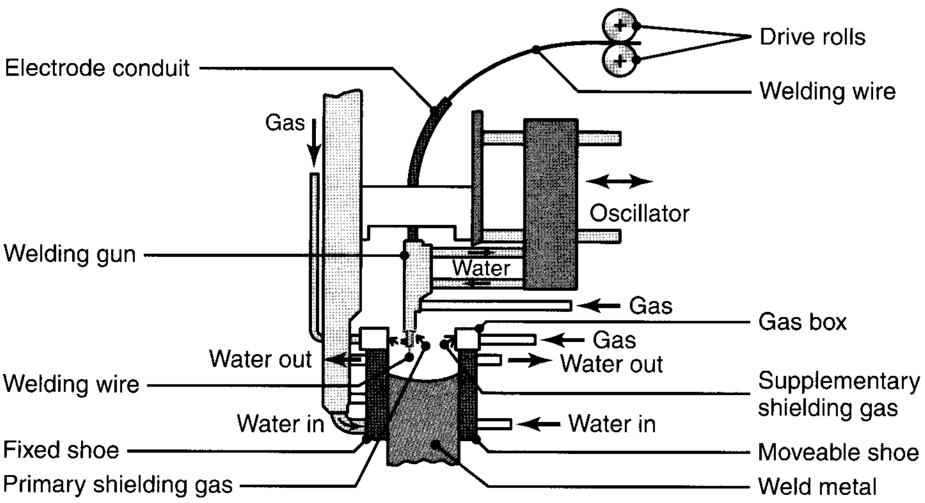


FIGURE 30.12 Schematic illustration of the electrogas-welding process.

Single or multiple electrodes are fed through a conduit, and a continuous arc is maintained by flux-cored electrodes at up to 750 A or solid electrodes at 400 A. Power requirements are about 20 kW. Shielding is done 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 provided either from an external source, from a flux-cored electrode, or from both.

The equipment for electrogas welding is reliable and training for operators is relatively simple. Weld thickness ranges from 12 to 75 mm on steels, titanium, and aluminum alloys. Typical applications are in the construction of bridges, pressure vessels, thick-walled and large-diameter pipes, storage tanks, and ships.

30.4.6 Electroslag Welding

Electroslag welding (ESW) and its applications are similar to electrogas welding (Fig. 30.13). The main difference is that the arc is started 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 as well as flux-cored electrodes may be used. The guide may be nonconsumable (conventional method) or consumable.

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 about 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 from 12 to 36 mm/min. Weld quality is good. This process is used for large structural-steel sections, such as heavy machinery, bridges, oil rigs, ships, and nuclear-reactor vessels.

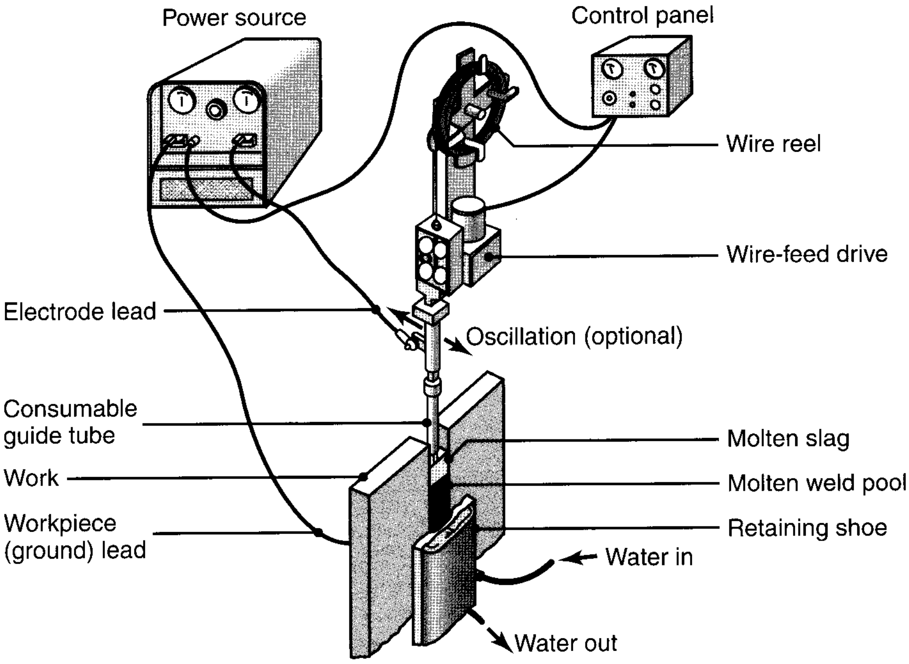


FIGURE 30.13 Equipment used for electroslag-welding operations.

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.

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 from 150 to 460 mm in length and 1.5 to 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 appear 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 supplier literature and in the various handbooks and references listed at the end of this chapter.

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	60,000 psi
E70XX	70,000
E110XX	110,000

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)

Note: Multiply pounds per square in. (psi) by 6.9×10^{-3} to obtain megapascals (MPa).

Electrode Coatings. Electrodes are *coated* with claylike materials that include silicate binders and powdered materials, such as oxides, carbonates, fluorides, metal alloys, cotton cellulose, and wood flour. The coating, which is brittle and takes part in complex interactions during welding, 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.

The deposited electrode coating or slag must be removed after each pass in order to ensure a good weld; a wire brush (manual or power) can be used for this purpose. Bare electrodes and wires made of stainless steels and aluminum alloys also are available. They 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. The process requires special equipment to focus the beam on the workpiece, typically in a vacuum. The higher the vacuum, the more the beam penetrates, and the greater is the depth-to-width ratio; thus, the methods are called EBW-HV (for "high vacuum") and EBW-MV (for "medium vacuum"); some materials also may also be welded by EBW-NV (for "no vacuum").

Almost any metal can be welded by EBW, and workpiece thicknesses can 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 (see **keyhole technique**, Section 30.3). Generally, no shielding gas, flux, or filler metal is required.

The EBW process has the capability of making high-quality welds that are almost parallel sided, are deep and narrow, and have small heat-affected zones (see Section 30.9). Depth-to-width ratios range between 10 and 30. The sizes of welds made by EBW are much smaller than those of welds made by conventional processes. With the use of automation and servo controls, parameters can be controlled accurately at welding speeds as high as 12 m/min.

Almost any metal can be butt or lap welded with this process at thicknesses up to 150 mm. Distortion and shrinkage in the weld area are minimal. The weld quality is good and of very high purity. Typical applications include the welding of aircraft, missile, nuclear, and electronic components, as well as gears and shafts for the automotive industry. Electron-beam welding equipment generates X-rays; hence, proper monitoring and periodic maintenance 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 the beam can be focused onto a very small area, it has high energy density and deep-penetrating capability. The beam can be directed, shaped, and focused precisely on the workpiece. Consequently, this process is suitable particularly for welding deep and narrow joints (Fig. 30.14) with depth-to-width ratios typically ranging from 4 to 10.

Laser-beam welding has become extremely popular and is used in most industries. In the automotive industry, welding transmission components are the most widespread application. Among numerous other applications is the welding of thin parts for electronic components. The laser beam may be **pulsed** (in milliseconds) with power levels up to 100 kW for applications such as the spot welding of thin materials. **Continuous** multi-kW laser systems are used for deep welds on thick sections.

Laser-beam welding produces welds of good quality with minimum shrinkage or distortion. Laser welds have good strength and generally are ductile and free of porosity. The process can be automated to be used on a variety of materials with thicknesses up to 25 mm; it is particularly effective on thin workpieces. 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.

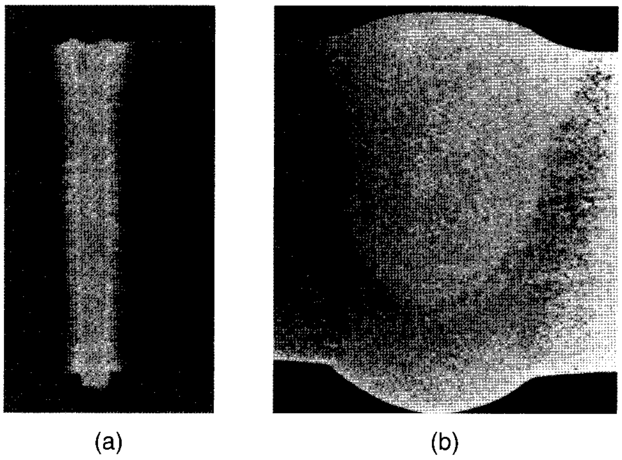

(a) (b)

FIGURE 30.14 Comparison of the sizes of weld beads: (a) laser-beam or electron-beam welding and (b) tungsten-arc welding. *Source*: Courtesy of American Welding Society.

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 as well as the skin; solid-state (YAG) lasers also are dangerous. (See Table 27.2 on types of lasers.)

The major advantages of LBW over EBW are the following:

- A vacuum is not required, and the beam can be transmitted through air.
- Laser beams can be shaped, manipulated, and focused optically (by means of fiber optics), so the process can be automated easily.
- The beams do not generate X-rays.
- The quality of the weld is better than in EBW; the weld has less tendency toward incomplete fusion, spatter, and porosity; and there is less distortion.

EXAMPLE 30.2 Laser Welding of Razor Blades

A close-up of the Gillette Sensor™ razor cartridge is shown in Fig. 30.15. Each of the two narrow, high-strength blades has 13 pinpoint welds—11 of which can be seen (as darker spots, about 0.5 mm in diameter) on each blade in the photograph. You can inspect the welds on actual blades with a magnifying glass or a microscope.

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 accurate and consistent weld quality.

Source: Courtesy of Lumonics Corporation, Industrial Products Division.

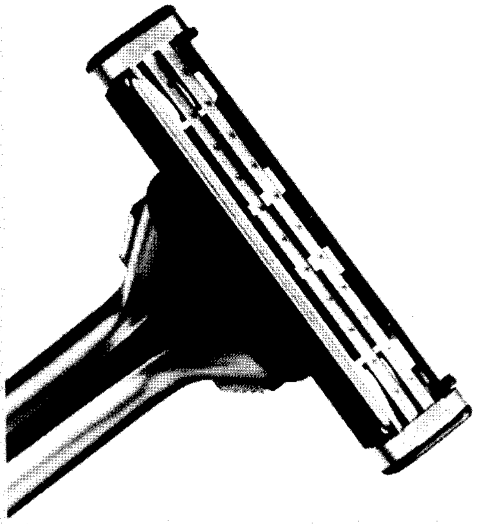


FIGURE 30.15 Detail of Gillette Sensor™ razor cartridge, showing laser spot welds.

30.8 Cutting

In addition to being cut by mechanical means, material can be cut into various contours with the use of 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 welding, but the heat source is now used to remove a narrow zone from a metal plate or sheet (Fig. 30.16a). This process is suitable particularly for steels. The basic reactions with steel are

$$\mathrm{Fe} + \mathrm{O} \longrightarrow \mathrm{FeO} + \mathrm{Heat}, \tag{30.6}$$

$$3\mathrm{Fe} + 2\mathrm{O_2} \longrightarrow \mathrm{Fe_3O_4} + \mathrm{Heat}, \tag{30.7}$$

and

$$4\mathrm{Fe} + 3\mathrm{O_2} \longrightarrow 2\,\mathrm{Fe_2O_3} + \mathrm{Heat}. \tag{30.8}$$

The greatest heat is generated by the second reaction, and it can produce a temperature rise to about 870°C. However, this temperature is not sufficiently high to cut steels; therefore, the workpiece is *preheated* with fuel gas, and oxygen is introduced later (see the nozzle cross section in Fig. 30.16a). The higher the carbon content of the steel, the higher is the preheating temperature required. Cutting takes place mainly by the oxidation (burning) of the steel; some melting also takes place. Cast irons and steel castings also can be cut by this method. The process generates a **kerf** similar to that produced by sawing with a saw blade or by wire electrical-discharge machining (see Fig. 27.12).

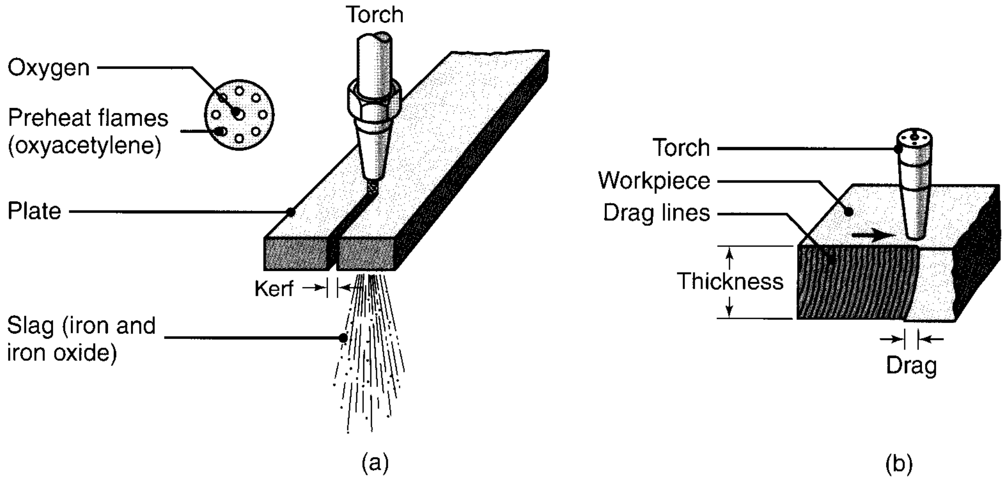


FIGURE 30.16 (a) Flame cutting of a steel plate with an oxyacetylene torch, and a cross section of the torch nozzle. (b) Cross section of a flame-cut plate, showing drag lines.

The maximum thickness that can be cut by OFC depends mainly on the gases used. With oxyacetylene gas, the maximum thickness is about 300 mm; with oxyhydrogen, it is about 600 mm. Kerf widths range from about 1.5 to 10 mm, with reasonably good control of tolerances. The flame leaves **drag lines** on the cut surface (Fig. 30.16b), resulting in a rougher surface than that produced by processes such as sawing, blanking, or other operations that use mechanical cutting tools. Distortion caused by uneven temperature distribution can be a problem in OFC.

Although long used for salvage and repair work, OFC can be used in manufacturing as well. Torches may be guided along specified paths either manually, mechanically, or automatically by machines using programmable controllers and robots. *Underwater cutting* is done with specially designed torches that produce a blanket of compressed air between the flame and the surrounding water.

Arc Cutting. *Arc-cutting* processes are based on the same principles as arc-welding processes. A variety of materials can be cut at high speeds by arc cutting. As in welding, these processes also leave a heat-affected zone that needs 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; thus, the metal being cut doesn't have to oxidize. This process is used especially for gouging and scarfing (removal of metal from a surface). However, the process is noisy, and the molten metal can be blown substantial distances and cause safety hazards.

Plasma-arc cutting (PAC) produces the highest temperatures. It is used for the rapid cutting of nonferrous and stainless-steel plates. The cutting productivity of this process is higher than that of oxyfuel–gas methods. PAC produces a good surface finish and narrow kerfs, and is the most popular cutting process utilizing programmable controllers employed in manufacturing today.

Electron beams and **lasers** also are used for very accurately cutting a wide variety of metals, as was 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. 30.17:

1. Base metal
2. Heat-affected zone
3. Weld metal.

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. A joint produced without a filler metal is called *autogenous*, and 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 the introduction of 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 and begins with the formation of columnar (dendritic) grains. (See Fig. 10.3.) These grains are relatively long and form parallel to the heat flow. Because metals are much better heat conductors than the surrounding air, the grains lie parallel to the plane of the two components being welded (Fig. 30.18a). In contrast, the grains in a shallow weld are shown in Figs. 30.18b and c.

Grain structure and grain size depend on the specific metal alloy, the particular welding process employed, and the type of filler metal. Because it began in a molten state, the weld metal basically has a *cast structure*, and since it has cooled slowly, it has coarse grains. Consequently, this structure generally has low strength, toughness, and ductility. However, the proper selection of filler-metal composition or of heat treatments following welding can improve the mechanical properties of the joint.

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. Preheating is important, particularly for metals having high thermal conductivity, such as aluminum and copper. Without preheating, the heat produced during welding dissipates rapidly through the rest of the parts being joined.

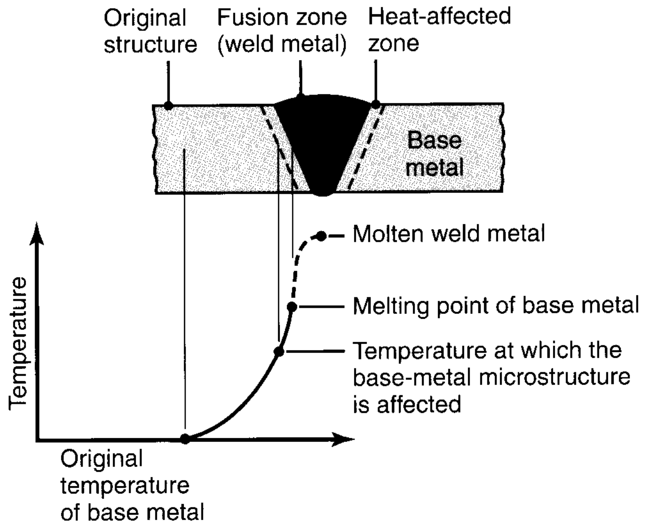


FIGURE 30.17 Characteristics of a typical fusion-weld zone in oxyfuel–gas and arc welding.

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 temperatures during welding. The portions of the base metal that are far enough away from the heat source do not undergo any microstructural changes during welding because of the far lower temperature to which they are subjected.

The properties and microstructure of the 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 the original grain size, grain orientation, and degree of prior cold

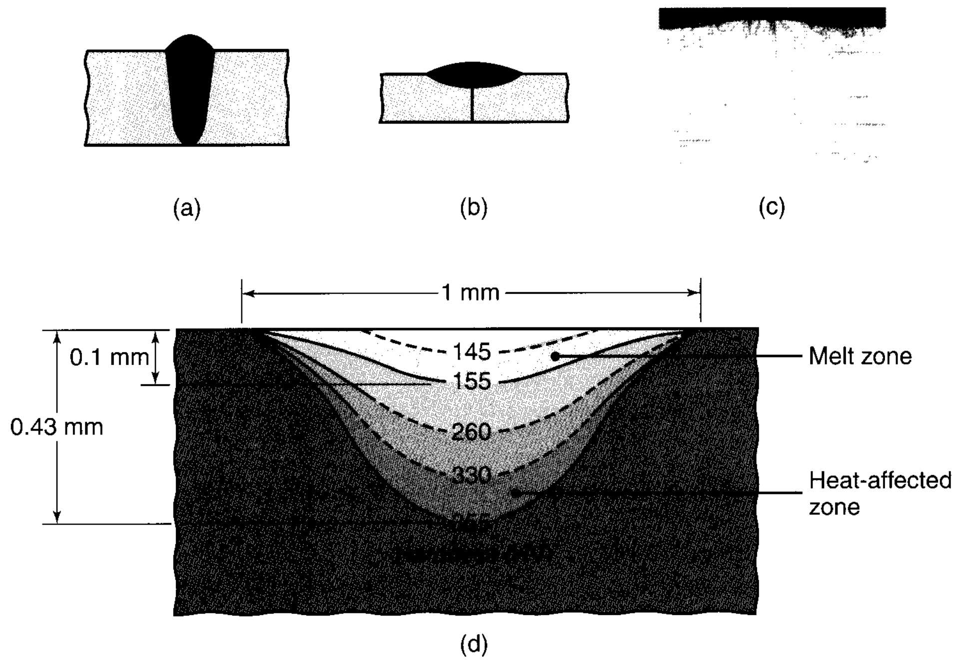


FIGURE 30.18 Grain structure in (a) a deep weld and (b) a shallow weld. Note that the grains in the solidified weld metal are perpendicular to their interface with the base metal. (c) Weld bead on a cold-rolled nickel strip produced by a laser beam. (d) Microhardness (HV) profile across a weld bead.

work), physical properties (such as the specific heat and thermal conductivity of the metals) influence the size and characteristics of the HAZ.

The strength and hardness of the HAZ (Fig. 30.18d) depend partly on how the original strength and hardness of the base metal was developed prior to the welding. As was described in Chapters 2 and 4, they may have been developed by (a) cold working, (b) solid-solution strengthening, (c) precipitation hardening, or (d) various heat treatments. The effects of these strengthening methods are complex, and the simplest to analyze are those in a base metal that has been cold worked, such as by cold rolling or cold forging.

The heat applied during welding *recrystallizes* the elongated grains of the cold-worked base metal. On the one hand, grains that are away from the weld metal will recrystallize into fine, equiaxed grains. On the other hand, grains close to the weld metal have been subjected to elevated temperatures for a longer time. Consequently, the grains will grow in size (grain growth), and this region will be softer and have lower strength. Such a joint will be weakest at its HAZ.

The effects of heat on the HAZ for joints made from dissimilar metals and for alloys strengthened by other methods are so complex as to be beyond the scope of this book. Details can be found in the more advanced references listed in 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 proper welding technologies or by poor operator training. The major discontinuities that affect weld quality are described here.

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 contain some porosity, which is 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 the 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 are important as well: With control of welding process parameters, the molten slag will float to the surface of the molten weld metal and thus will not become entrapped.

Slag inclusions can be prevented by the following practices:

- Cleaning the weld-bead surface by means of a wire brush (hand or power) or 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 and Penetration. *Incomplete fusion* produces poor weld beads, such as those shown in Fig. 30.19. A better weld can be obtained by the use of the following practices:

- Raising the temperature of the base metal.
- Cleaning the weld area before welding.

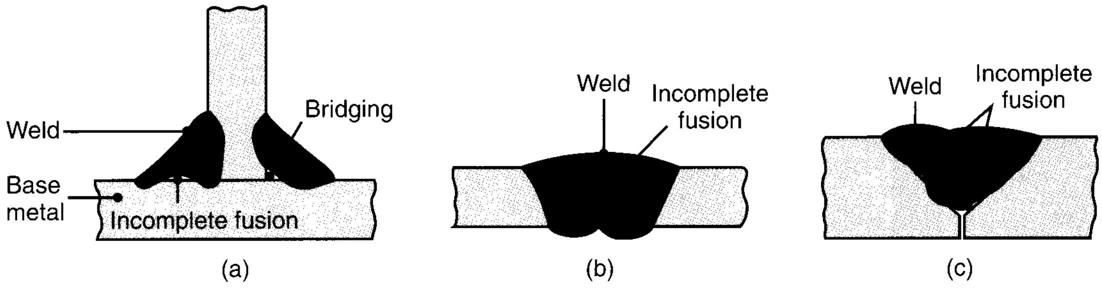


FIGURE 30.19 Examples of various discontinuities in fusion welds.

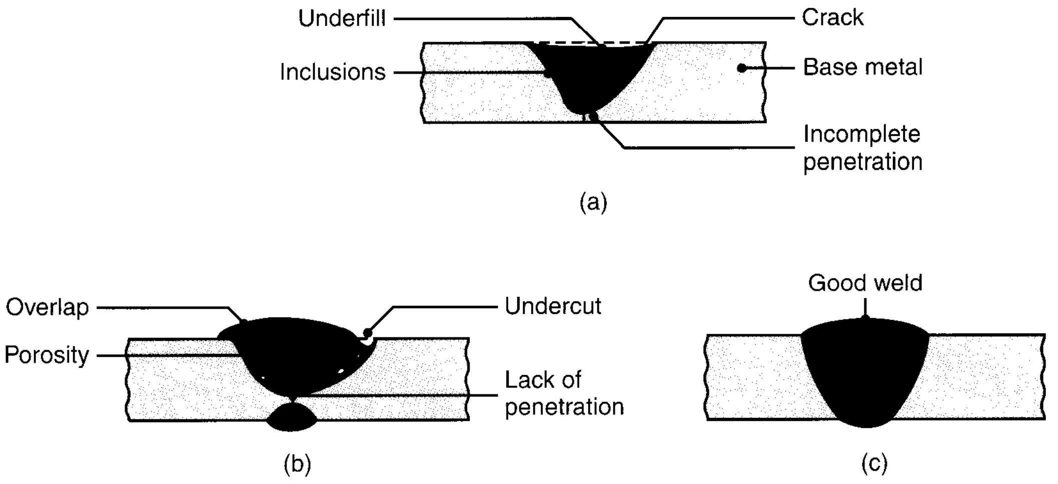


FIGURE 30.20 Examples of various defects in fusion welds.

- Modifying the joint design and changing the type of electrode used.
- Providing sufficient shielding gas.

Incomplete penetration occurs when the depth of the welded joint is insufficient. Penetration can be improved by the following practices:

- Increasing the heat input.
- Reducing the travel speed during the welding.
- Modifying the joint design.
- Ensuring that the surfaces to be joined fit each other 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. 30.20a).
- **Undercutting** results from the melting away of the base metal and the consequent generation of a groove in the shape of a sharp recess or notch (Fig. 30.20b). If it is deep or sharp, an undercut can act as a stress raiser and can reduce the fatigue strength of the joint; in such cases, it may lead to premature failure.
- **Overlap** is a surface discontinuity (Fig. 30.20b) usually caused by poor welding practice or by the selection of improper materials. A good weld is shown in Fig. 30.20c.

Cracks. Cracks may occur in various locations and directions in the weld area. Typical types of cracks are longitudinal, transverse, crater, underbead, and toe cracks (Fig. 30.21).

Cracks generally result from a combination of the following factors:

- Temperature gradients that cause thermal stresses in the weld zone.
- Variations in the composition of the weld zone that cause different rates of contraction during cooling.
- 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 when the weld metal begins to solidify.

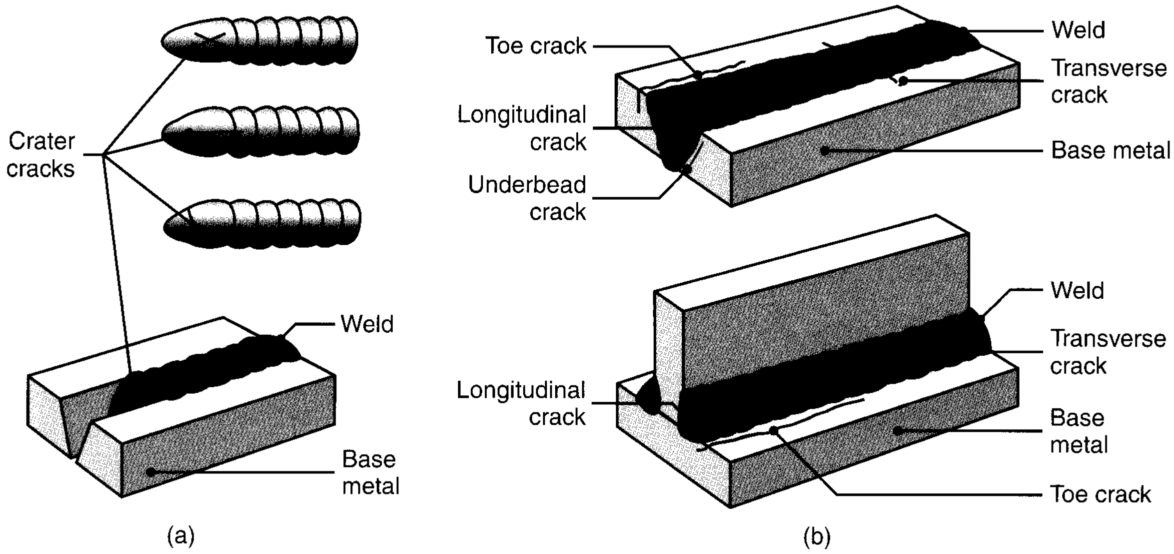


FIGURE 30.21 Types of cracks developed in welded joints. The cracks are caused by thermal stresses, similar to the development of hot tears in castings, as shown in Fig. 10.12.

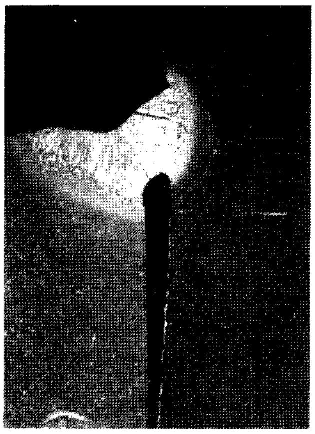


FIGURE 30.22 Crack in a weld bead. The two welded components were not allowed to contract freely after the weld was completed. *Source*: Courtesy of Packer Engineering.

- Hydrogen embrittlement (Section 2.10.2).
- Inability of the weld metal to contract during cooling (Fig. 30.22). This is a situation similar to *hot tears* that develop in castings (Fig 10.12) and is related to excessive restraint of the workpiece during the welding operation.

Cracks also are classified as **hot cracks**, which occur while the joint is still at elevated temperatures, and **cold cracks**, which develop after the weld metal has solidified.

The basic crack-prevention measures in welding are the following:

- Modify the joint design to minimize stresses developed from shrinkage during cooling.
- Change the parameters, procedures, and sequence of the welding operation.
- Preheat the components to be welded.
- Avoid rapid cooling of the welded components.

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 nonmetallic impurities and inclusions (stringers). This condition is particularly evident in rolled plates and in structural shapes. In welding such components, *lamellar tears* may develop because of shrinkage of the restrained components of the structure during cooling. Such 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 metal may spatter during welding and be deposited as small droplets on adjacent surfaces. In arc-welding processes, the electrode inadvertently may touch the parts being welded at places other than the weld zone. (Such encounters are called **arc strikes**.) The surface discontinuities thereby produced may be objectionable for reasons of appearance or of subsequent use of the welded part. If severe, these discontinuities adversely may affect the properties of the welded structure, particularly for notch-sensitive metals. Using proper welding techniques and procedures is important in avoiding surface damage.

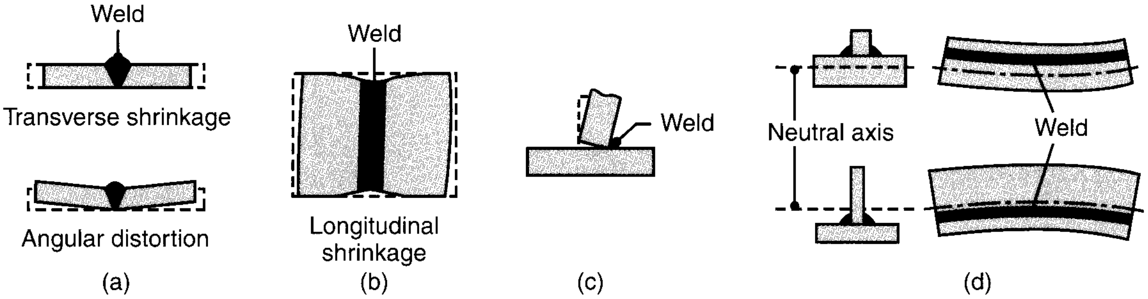


FIGURE 30.23 Distortion of parts after welding. Distortion is caused by differential thermal expansion and contraction of different regions of the welded assembly.

Residual Stresses. Because of localized heating and cooling during welding, the expansion and contraction of the weld area causes *residual stresses* in the workpiece. (See also Section 2.11.) Residual stresses can lead to the following defects:

- Distortion, warping, and buckling of the welded parts (Fig. 30.23).
- Stress-corrosion cracking (Section 2.10.2).
- Further distortion if a portion of the welded structure is subsequently removed, such as by machining or sawing.
- Reduced fatigue life of the welded structure.

The type and distribution of residual stresses in welds is described best by reference to Fig. 30.24a. 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, the heat from the weld zone dissipates laterally into the plates, while the weld area cools. Thus, the plates begin to expand longitudinally, while the welded length begins to contract (Fig. 30.23).

If the plate is not constrained, it will warp, as shown in Fig. 30.23a. However, if the plate is not free to warp, it will develop residual stresses that typically are distributed throughout the material, such as the stresses shown in Fig. 30.24. Note that the magnitude of the compressive residual stresses in the plates diminishes to zero at a point far away from the weld area. Because no external forces are acting on the welded plates, the tensile and compressive forces represented by these residual stresses must balance each other.

Events leading to the distortion of a welded structure are shown in Fig. 30.25. Before welding, the structure is stress free, as shown in Fig. 30.25a. The shape may be fairly rigid, and fixturing also may be present to support the structure. When the weld bead is placed, the molten metal fills the gap between the surfaces to be joined, and flows outward to form the weld bead. At this point, the weld is not under any stress. Afterward, the weld bead solidifies, and both the weld bead and the surrounding material cool to room temperature. As these materials cool, they tend to contract, but are constrained by the bulk of the weldment. The result is that the weldment distorts (Fig. 30.25c) and residual stresses develop.

The residual-stress distribution shown in Fig. 30.25 places the weld and the HAZ in a state of residual tension, which is harmful from a fatigue standpoint.

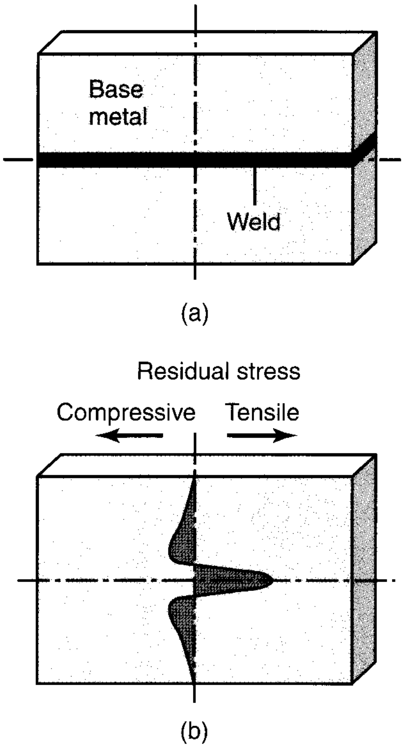


FIGURE 30.24 Residual stresses developed in (a) a straight-butt joint. Note that the residual stresses shown in (b) must be balanced internally. (See also Fig. 2.29.)

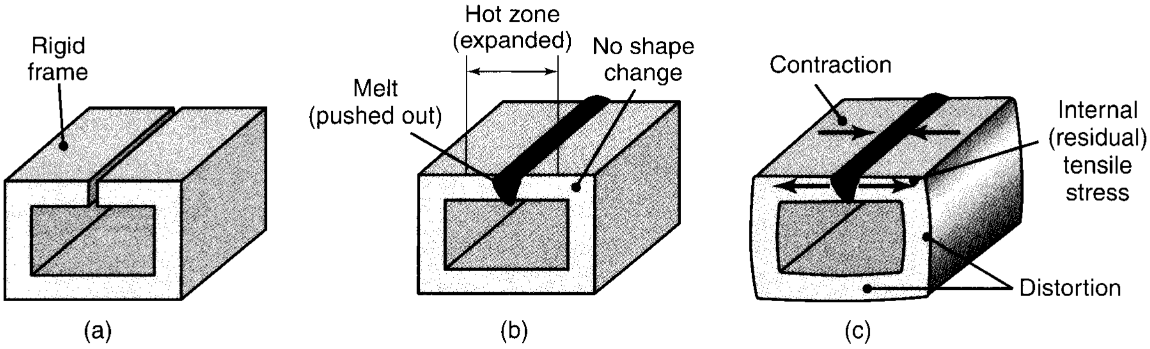


FIGURE 30.25 Distortion of a welded structure. *Source*: After J.A. Schey.

Many welded structures will use cold-worked materials (such as extruded or roll-formed shapes), and these are relatively strong and fatigue resistant. The weld itself may have porosity (see Fig. 30.20b), which can act as a stress riser and aid fatigue-crack growth, or there could be other cracks that can grow in fatigue. In general, the HAZ is less fatigue resistant than the base metal. Thus, the residual stresses developed can be very harmful, and it is not unusual to further treat welds in highly stressed or fatigue-susceptible applications.

In complex welded structures, residual-stress distributions are three dimensional and, consequently, difficult to analyze. The previous discussion involved two plates that were not restrained from movement. In other words, the plates were not an integral part of a larger structure. If, however, they are restrained, reaction stresses will be generated, because the plates are not free to expand or contract. This situation arises particularly in structures with high stiffness.

Stress Relieving of Welds. The problems caused by residual stresses (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 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 carefully in order to maintain acceptable strength and toughness in the welded structure. The workpieces may be heated in several ways, including (a) in a furnace, (b) electrically (resistively or inductively), or (c) by radiant lamps or a hot-air blast for thin sections. The temperature and time required for *stress relieving* depend on the type of material and on the magnitude of the residual stresses developed.

Other methods of stress relieving include peening, hammering, or surface rolling 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, in order to protect them against possible peening damage.

Residual stresses can also be relieved or reduced by plastically deforming the structure by a small amount. For instance, this technique can be used in welded pressure vessels by pressurizing the vessels internally (*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 could act as points of stress concentration.

In addition to being preheated for stress relieving, welds may be *heat treated* by various other techniques in order to modify other properties. These techniques include the annealing, normalizing, quenching, and tempering of steels and the solution treatment and aging of various alloys as described in Chapter 4.

30.9.2 Weldability

The *weldability* of a metal is usually 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; hence, generalizations are difficult. As noted previously, the material characteristics (such as alloying elements, impurities, inclusions, grain structure, and processing history) of both the base metal and the filler metal are important. For example, the weldability of steels decreases with increasing carbon content because of martensite formation (which is hard and brittle) and thus reduces the strength of the weld. Coated steel sheets 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 knowledge 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.

The preparation of surfaces for welding is important, as are the nature and properties of surface-oxide films and of adsorbed gases. The particular 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 postwelding techniques as stress relieving and heat treating.

Weldability of Ferrous Materials:

- *Plain-carbon steels:* Weldability is excellent for low-carbon steels, fair to good for medium-carbon steels, poor for high-carbon steels.
- *Low-alloy steels:* Weldability is similar to that of medium-carbon steels.
- *High-alloy steels:* Weldability generally is good under well-controlled conditions.
- *Stainless steels:* These generally are weldable by various processes.
- *Cast irons:* These generally are weldable, although their weldability varies greatly.

Weldability of Nonferrous Materials:

- *Aluminum alloys:* These are 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, these generally are weldable at a high rate of heat input. An inert shielding gas and lack of moisture are important.
- *Magnesium alloys:* These are weldable with the use of a protective shielding gas and fluxes.
- *Nickel alloys:* Weldability is similar to that of stainless steels. The lack of sulfur is undesirable.
- *Titanium alloys:* These are weldable with the proper use of shielding gases.
- *Tantalum:* Weldability is similar to that of titanium.

- *Tungsten:* Weldable under well-controlled conditions.
- *Molybdenum:* Weldability is similar to that of tungsten.
- *Niobium (columbium):* Possesses good weldability.

30.9.3 Testing of Welds

As in all manufacturing processes, the quality of a welded joint is established by testing. Several standardized tests and test procedures have been established. They are available from many 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 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 obtained by the procedures described in Section 2.2. These curves indicate the yield strength, Y, ultimate tensile strength, UTS, 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 (Figs. 30.26a and b) are prepared to simulate conditions to which actual welded joints are subjected. These specimens are subjected to tension so that the shear strength of the weld metal and the location of fracture can be determined.

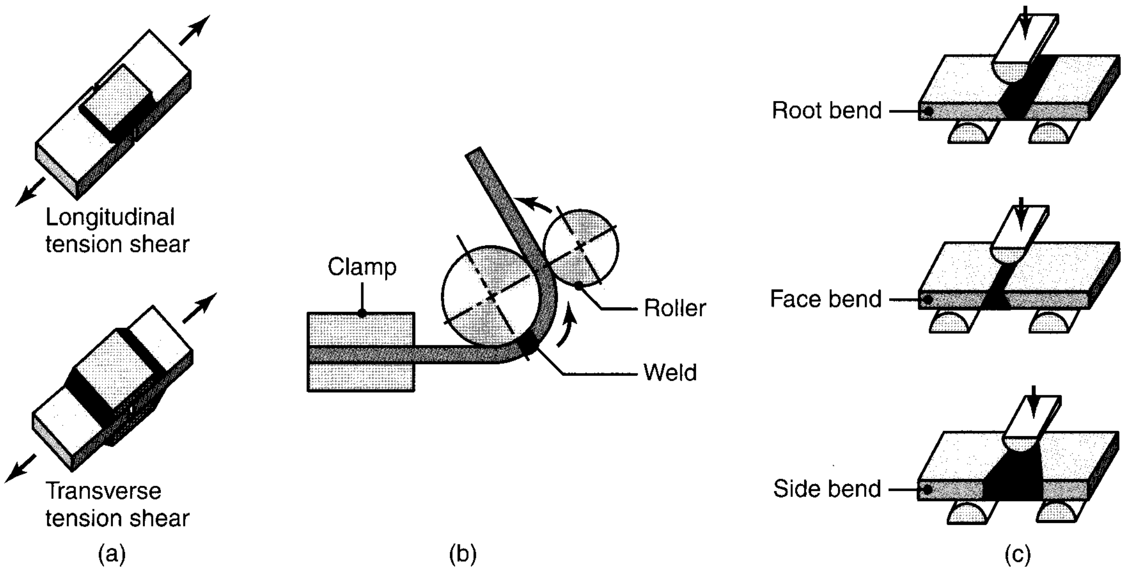


FIGURE 30.26 (a) Specimens for longitudinal tension-shear testing and for transfer tension-shear testing. (b) Wraparound bend-test method. (c) Three-point transverse bending of welded specimens.

- **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. 30.26c). In another test, the specimens are tested in *three-point transverse bending* (Fig. 30.26d; see also Fig. 2.11a). These tests help to determine the relative ductility and strength of welded joints.
- **Fracture toughness test:** Fracture toughness tests commonly utilize the impact-testing techniques described in Section 2.9. *Charpy V-notch* specimens are first prepared and then tested for toughness. Another toughness test is the *drop-weight test*, in which the energy is supplied by a falling weight.
- **Corrosion and creep tests:** In addition to undergoing mechanical tests, welded joints also may be tested for their resistance to corrosion and creep. Because of the difference in the composition and microstructure of the materials in the weld zone, *preferential corrosion* may take place in the zone. Creep tests are important in determining the behavior of welded joints and structures subjected to elevated temperatures.

Nondestructive Testing Techniques. Welded structures often have to be tested nondestructively, 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 (these tests are described in Section 36.10):

- **Visual**
- **Radiographic** (X-rays)
- **Magnetic-particle**
- **Liquid-penetrant**
- **Ultrasonic.**

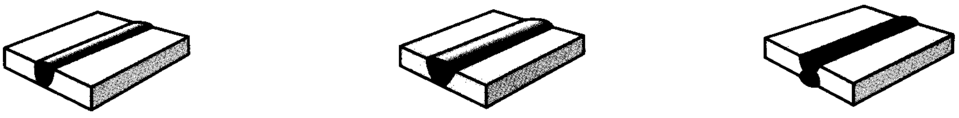
Testing for hardness distribution in the weld zone also may be a useful indicator of weld strength and microstructural changes.

30.10 Joint Design and Process Selection

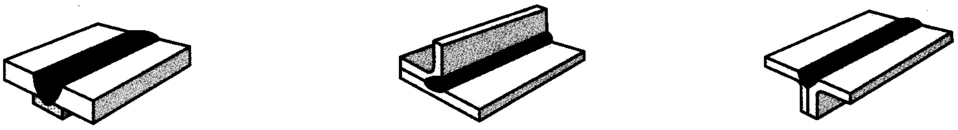
In describing individual welding processes, several examples were given concerning the types of welds and joints produced and their applications in numerous consumer and industrial products of various designs. Typical types of joints produced by welding, together with their terminology, are shown in Fig. 30.27. Standardized symbols commonly used in engineering drawings to describe the types of welds are shown in Fig. 30.28. These symbols identify the type of weld, the groove design, the weld size and length, the welding process, the sequence of operations, and other necessary information.

The general design guidelines for welding may be summarized as follows, with some examples given in Fig. 30.29 (various other types of joint design will be given in Chapters 31 and 32):

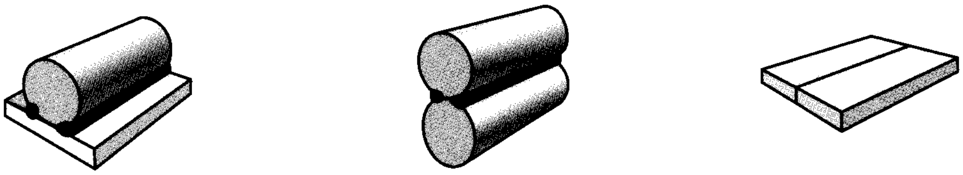
- Product design should minimize the number of welds because, unless automated, welding can be costly.
- Weld location should be selected so as to avoid excessive stresses or stress concentrations in the welded structure and for appearance.
- Weld location should be selected so as not to interfere with any subsequent processing of the joined components or with their intended use.



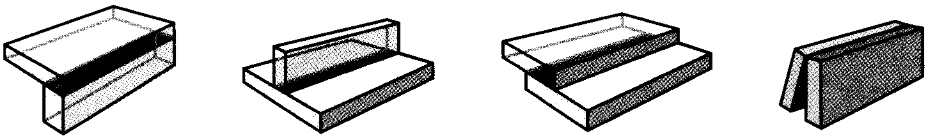
(a) Single square-groove weld (b) Single V-groove weld (c) Double V-groove weld



(d) Single V-groove weld (with backing) (e) Single-flare bevel-groove weld (f) Single-flare V-groove weld



(g) Double-flare bevel-groove weld (h) Double-flare V-groove weld (i) Butt joint



(j) Corner joint (k) T joint (l) Lap joint (m) Edge joint

FIGURE 30.27 Examples of welded joints and their terminology.

- Components should fit properly prior to welding. The method used to prepare edges, such as sawing, machining, or shearing, also can affect weld quality.
- The need for edge preparation should be avoided or minimized.
- Weld-bead size should be as small as possible, while maintaining the strength of the joint, to conserve weld metal and for better appearance.

Welding Process Selection. In addition to taking into account the process characteristics, capabilities, and material considerations described thus far in this chapter, the selection of a weld joint and an appropriate welding process involve the following considerations (see also Chapters 31 and 32).

- Configuration of the parts or structure to be joined, joint design, thickness and size of the components, and number of joints required.
- The methods used in manufacturing the components to be joined.
- Types of materials involved, which may be metallic or nonmetallic.
- Location, accessibility, and ease of joining.
- Application and service requirements, such a type of loading, any stresses generated, and the environment.
- Effects of distortion, warping, discoloration of appearance, and service.

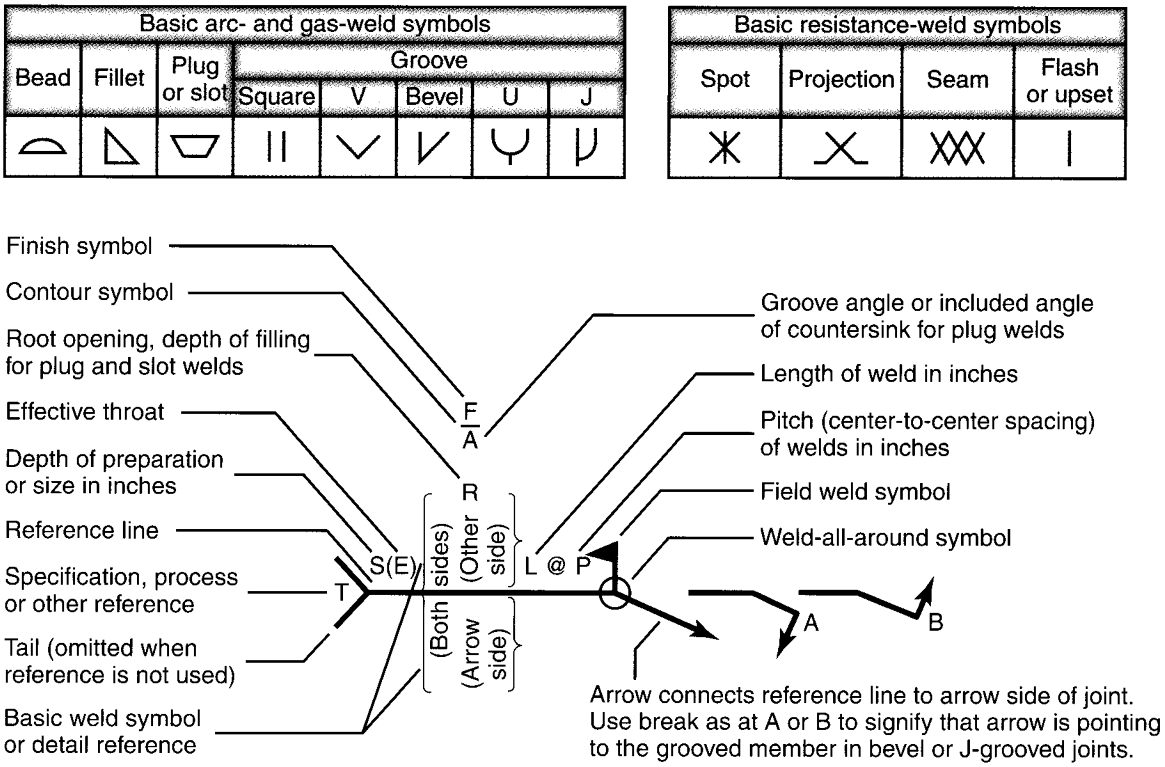


FIGURE 30.28 Standard identification and symbols for welds.

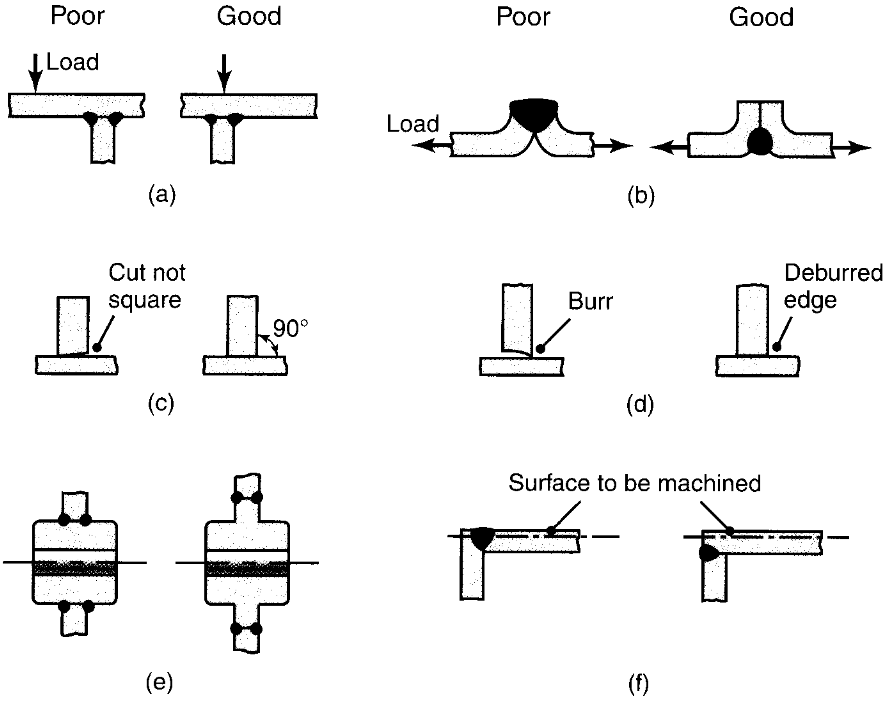


FIGURE 30.29 Some design guidelines for welds. *Source*: After J.G. Bralla.

- Costs involved in edge preparation, joining, and postprocessing (including machining, grinding, and finishing operations).
- Costs of equipment, materials, labor and skills required, and the joining operation.

Table VI.1 gave the various characteristics of individual welding processes—characteristics that would serve as an additional guide to process selection. Refering to this table, note that no single process has a high rating in all categories. For example,

- Arc welding, bolts, and riveting have high strength and reliability, but generally are not suitable for joining small parts.
- Resistance welding has strength and applications for both small and large parts. However, it is not easy to inspect visually for reliability, and resistance welding has lower tolerances and reliability than other processes.
- Fasteners are useful for large parts and can be easy to inspect visually, but they are costly and do not have much design variability.
- Adhesive bonding has high design variability. However, it has relatively low strength and is difficult to visually inspect for joint integrity.

EXAMPLE 30.3 Weld Design Selection

Three different types of weld designs are shown in Fig. 30.30. In Fig. 30.30a, the two vertical joints 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, by the alternative method, the appearance of the structure is improved and distortion is reduced.

In Fig. 30.30b, it can be shown that the design on the right can carry three times the moment M of the one on the left. Note that both designs require the same amount of weld metal and welding time. In Fig. 30.30c, the weld on the left requires about twice the amount of weld material than does the design on the right. Note that because more material must be machined, the design on the left will require more time for edge preparation, and more base metal will be wasted as a result.

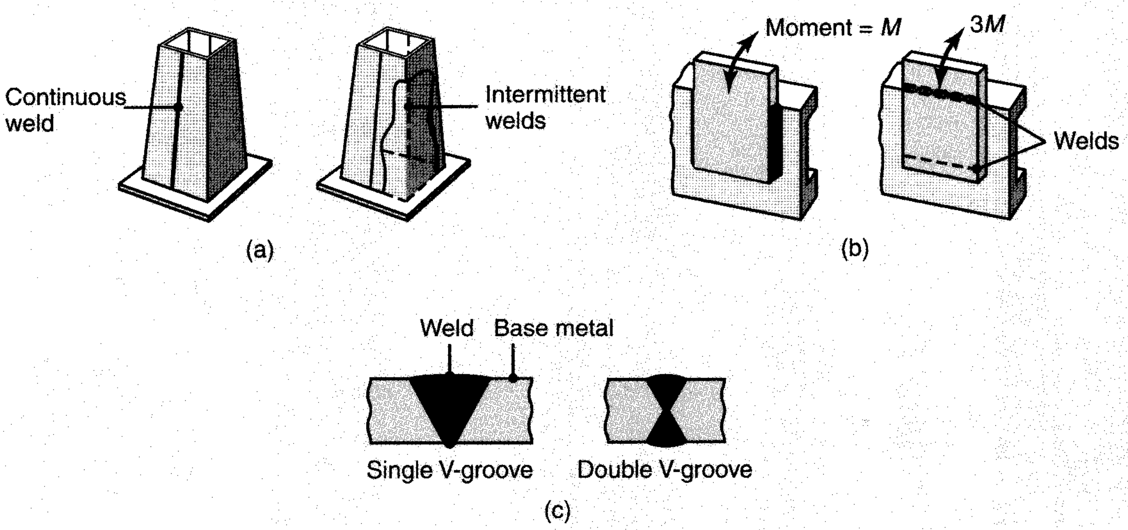


FIGURE 30.30 Examples of weld designs used in Example 30.3.

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 instead.
- In all of these processes, 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 rods may or may not be used in oxyfuel–gas and arc welding to fill the weld area.
- The 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, the strength required, and the change in product appearance caused by welding.
- A variety of welding equipment is available—much of which is now robotics and computer controlled with programmable features.
- The cutting of metals also can be done by processes whose principles are based on oxyfuel–gas and arc welding. The cutting of steels occurs mainly through oxidation (burning). The highest temperatures for cutting are obtained by plasma-arc cutting.
- The metallurgy of the welded joint is an important aspect of all welding processes, because it determines the strength and toughness of the joint. The welded joint consists of solidified metal and a heat-affected zone; each has a wide variation in microstructure and properties, depending on the metals joined and on the filler metals.

- The metallurgy of the welded joint is an important aspect of all welding processes, because it determines the strength and toughness of the joint. The welded joint consists of solidified metal and a heat-affected zone; each has a wide variation in microstructure and properties, depending on the metals joined and on the filler metals.

- Discontinuities such as porosity, inclusions, incomplete welds, tears, surface damage, and cracks can develop in the weld zone. Residual stresses and relieving them also are important considerations in welding.
- The weldability of metals and alloys depends greatly on their composition, the type of welding operation and process parameters employed, and the control of welding parameters.
- General guidelines are available to help in the initial selection of suitable and economical welding methods for a particular 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
Joining
Kerf
Keyhole technique
Laser-beam 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
Weld profile
Weld metal
Weldability
Welding gun

BIBLIOGRAPHY

ASM Handbook, Vol. 6: **Welding, Brazing, and Soldering,** ASM International, 1993.

Bowditch, W.A., and Bowditch, K.E., **Welding Technology Fundamentals**, Goodheart-Willcox, 1997.

Cary, H.B., and Helzer, S., **Modern Welding Technology,** 6th ed., Prentice Hall, 2004.

Croft, D., **Heat Treatment of Welded Structures**, Woodhead Publishing, 1996.

Davies, A.C., **The Science and Practice of Welding,** 10th ed. (2 vols.), Cambridge University Press, 1993.

Duley, W.W., **Laser Welding,** Wiley, 1999.

Evans, G.M., and Bailey, N., **Metallurgy of Basic Weld Metal,** Woodhead Publishing, 1997.

Hicks, J.G., **Welded Joint Design,** 2nd ed., Abington, 1997.

Houldcroft, P.T., **Welding and Cutting: A Guide to Fusion Welding and Associated Cutting Processes,** Industrial Press, 2001.

Jeffus, L.F., **Welding: Principles and Applications,** 6th ed., Delmar Publishers, 2007.

Kou, S., **Welding Metallurgy,** 2nd ed., Wiley, 2003.

Lippold, J.C., and Kotecki, D.J., **Welding Metallurgy and Weldability of Stainless Steels,** Wiley, 2005.

Minnick, W.H., **Gas Metal Arc Welding Handbook,** Goodheart-Willcox, 2000.

Mouser, J.D., **Welding Codes, Standards, and Specifications,** McGraw-Hill, 1997.

Schultz, H., **Electron Beam Welding,** Woodhead Publishing, 1994.

Steen, W.M., **Laser Material Processing,** 3rd ed., Springer, 2003.

Welding Handbook, 9th ed. (3 vols.), American Welding Society, 2007.

Welding Inspection Handbook, American Welding Society, 2000.

Weman, K., **Welding Process Handbook,** CRC Press, 2003.

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. Explain the basic principles of arc-welding processes.

30.4. Why is shielded metal-arc welding a commonly used process? Why is it also called stick welding?

30.5. Describe the functions and characteristics of electrodes. What functions do coatings have? How are electrodes classified?

30.6. What are the similarities and differences between consumable and nonconsumable electrodes?

30.7. Explain how cutting takes place when an oxyfuel–gas torch is used. How is underwater cutting done?

30.8. What is the purpose of flux? Why is it not needed in gas tungsten-arc welding?

30.9. What is meant by weld quality? Discuss the factors that influence it.

30.10. How is weldability defined?

30.11. Why are welding electrodes generally coated?

30.12. Describe the common types of discontinuities in welded joints.

30.13. What types of destructive tests are performed on welded joints?

QUALITATIVE PROBLEMS

30.14. Explain the reasons that so many different welding processes have been developed over the years.

30.15. Explain why some joints may have to be preheated prior to welding.

30.16. Describe the role of filler metals in welding.

30.17. What is the effect of the thermal conductivity of the workpiece on kerf width in oxyfuel–gas cutting? Explain.

30.18. Describe the differences between oxyfuel–gas cutting of ferrous and of nonferrous alloys. Which properties are significant?

30.19. Could you use oxyfuel–gas cutting for a stack of sheet metals? (*Note*: For stack cutting, see Fig. 24.25e.) Explain.

30.20. What are the advantages of electron-beam and laser-beam welding compared with arc welding?

30.21. Describe the methods by which discontinuities in welding can be avoided.

30.22. Explain the significance of the stiffness of the components being welded on both weld quality and part shape.

30.23. Comment on the factors that influence the size of the two weld beads shown in Fig. 30.14.

30.24. Which of the processes described in this chapter are not portable? Can they be made so? Explain.

30.25. Describe your observations concerning the contents of Table 30.1.

30.26. 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.

30.27. Comment on the factors involved in electrode selection in arc-welding processes.

30.28. 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.

30.29. Explain the significance of residual stresses in welded structures.

30.30. Rank the processes described in this chapter in terms of (a) cost and (b) weld quality.

30.31. Must the filler metal be made of the same composition as the base metal that is to be welded? Explain.

30.32. Describe your observations concerning Fig. 30.18.

30.33. If the materials to be welded are preheated, is the likelihood for porosity increased or decreased? Explain.

QUANTITATIVE PROBLEMS

30.34. Plot the hardness in Fig. 30.18d as a function of the distance from the top surface, and discuss your observations.

▶**30.35.** 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?

▶**30.36.** In Fig. 30.24b, 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.)

SYNTHESIS, DESIGN, AND PROJECTS

30.37. Comment on workpiece size and shape limitations for each of the processes described in this chapter.

30.38. Review the types of welded joints shown in Fig. 30.27 and give an application for each.

30.39. Comment on the design guidelines given in various sections of this chapter.

30.40. 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.

30.41. Discuss the need for, and the role of, work-holding devices in the welding operations described in this chapter.

30.42. 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.

30.43. Explain the factors that contribute to the differences in properties across a welded joint.

30.44. Explain why preheating the components to be welded is effective in reducing the likelihood of developing cracks.

30.45. Review the poor and good joint designs shown in Fig. 30.29, and explain why they are labeled so.

30.46. 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.

30.47. 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.

30.48. Comment on whether there are common factors that affect the weldability, castability, formability, and machinability of metals, as described in various chapter of this book. Explain with appropriate examples.

30.49. If you find a flaw in a welded joint during inspection, how would you go about determining whether or not the flaw is significant?

30.50. 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.

30.51. A common practice in repairing expensive broken or worn parts (such as 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.

CHAPTER 31

Solid-State Welding Processes

EXAMPLES:

CASE STUDY:

- This chapter describes an important family of joining processes in which the workpieces do not undergo a phase change and no filler metal is used; if heat is used, it is not externally applied, but instead is generated internally—for example, with friction.
- The chapter begins with a discussion of cold welding, followed by ultrasonic welding and the various forms of 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 also examines special capabilities of diffusion bonding combined with superplastic forming.
- Finally, economic considerations in welding are discussed.

31.1 Introduction

This chapter describes **solid-state welding** processes, in which joining takes place without fusion at the interface of the two parts to be welded. Unlike the situation with the fusion-welding processes described in Chapter 30, in solid-state welding no liquid or molten phase is present in the joint. The principle of solid-state welding is demonstrated best with the following example: If two clean surfaces are brought into close contact with each other under sufficient pressure, they form bonds and produce a joint. To form a strong bond, it is essential that the interface be free of oxide films, residues, metalworking fluids, other contaminants, and even adsorbed layers of gas.

Solid-state bonding involves one or more of the following phenomena:

- **Diffusion:** The transfer of atoms across an interface; thus, applying external heat improves the strength of the bond between the two surfaces being joined, as occurs in *diffusion bonding*. Heat may be generated internally by friction (as utilized in *friction welding*), through electrical-resistance heating (as in *resistance-welding* processes, such as *spot welding*), and externally by induction heating (as in *butt-welding* tubes).
- **Pressure:** The higher the pressure, the stronger is the interface (as in *roll bonding* and *explosion welding*), where plastic deformation also occurs. Pressure

and resistance heating may be combined, as in *flash welding, stud welding*, and *resistance projection welding*.

- **Relative interfacial movements:** When movements of the contacting surfaces (*faying surfaces*) occur (as in *ultrasonic welding*), even very small amplitudes will disturb the mating surfaces, break up any oxide films, and generate new, clean surfaces—thus improving the strength of the bond.

Most of the joining processes outlined here are now automated by *robotics, vision systems, sensors*, and *adaptive* and *computer controls* (see Part IX) for cost reduction, consistency, reliability of weld quality, and higher productivity. The costs involved in the joining process are outlined in Section 31.8.

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 necessary that at least one (but preferably both) of the mating parts be ductile. Prior to welding, the interface is degreased, wire brushed, and wiped to remove oxide smudges. Cold welding can be used to join small workpieces made of soft, ductile metals. Applications include wire stock and electrical connections.

During the 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 occurs in the bonding of aluminum and steel, where a brittle intermetallic compound is formed at the interface. 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); this process is called *roll bonding* or *roll welding* (ROW). Developed in the 1960s, roll bonding is used for manufacturing some U.S. coins (see Example 31.1). The process can be carried out at elevated temperatures (*hot roll bonding*). Surface preparation is important for interfacial strength.

Typical examples are the *cladding* of (a) pure aluminum over precipitation-hardened aluminum-alloy sheet (Alclad) and (b) stainless steel over mild steel (for corrosion resistance). A common application of roll bonding is the production of bimetallic strips for thermostats and similar controls using two layers of materials with different thermal-expansion coefficients. Bonding in only selected regions in the interface can be achieved by depositing a parting agent, such as graphite or ceramic, called *stop-off* (see Section 31.7).

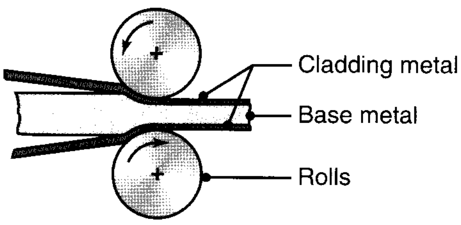


FIGURE 31.1 Schematic illustration of the roll bonding, or cladding, process.

EXAMPLE 31.1 Roll Bonding of the U.S. Quarter

The technique used for manufacturing composite U.S. quarters is the roll bonding of two outer layers of 75% Cu–25% Ni (cupronickel), where each layer is 1.2 mm thick, with an inner layer of pure copper 5.1 mm thick. To obtain good bond strength, the faying surfaces are cleaned chemically and wire brushed. First, the strips are rolled to a thickness of 2.29 mm; a second rolling operation reduces the thickness to 1.36 mm. The strips thus undergo a total reduction in thickness of 82%.

Because volume constancy is maintained in plastic deformation, there is a major increase in the surface area between the layers, and it causes the generation of clean interfacial surfaces. This extension in surface area under the high pressure of the rolls, combined with the solid solubility of nickel in copper (see Section 4.2.1), produces a strong bond.

31.3 Ultrasonic Welding

In *ultrasonic welding* (USW), the faying surfaces of the two components are subjected to a static 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. 26.24a.) The frequency of oscillation is generally in the range from 10 to 75 kHz, although a lower or higher frequency can be employed. Proper coupling between the transducer and the tip (called—by analogy with *electrode*—a **sonotrode**, from the word *sonic*) is important for efficient operation.

The shearing stresses cause plastic deformation at the interface of the two components, breaking up oxide films and contaminants and thus allowing good contact and producing a strong solid-state bond. The temperature generated in the weld zone is usually in the range from one-third to one-half of the melting point (absolute scale) of the metals joined. Consequently, neither melting nor fusion takes place.

In certain situations, however, the temperature generated can be sufficiently high to cause metallurgical changes in the weld zone. Also, the mechanism responsible for the joining of *thermoplastics* by ultrasonic welding is different from that for metals, and melting does take place at the interface, because plastics have much lower melting temperatures. (See Table 7.2.)

The ultrasonic-welding process is versatile and reliable. It can be used with a wide variety of metallic and nonmetallic materials, including dissimilar metals (*bimetallic strips*). It is used extensively for the joining of plastics, for packaging with foils, and (in the automotive and consumer electronics industries) for the lap welding of sheet, foil, and thin wire. The welding tip can be replaced with *rotating disks* (Fig. 31.2b) for the 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). Moderate skill is required to operate the equipment.

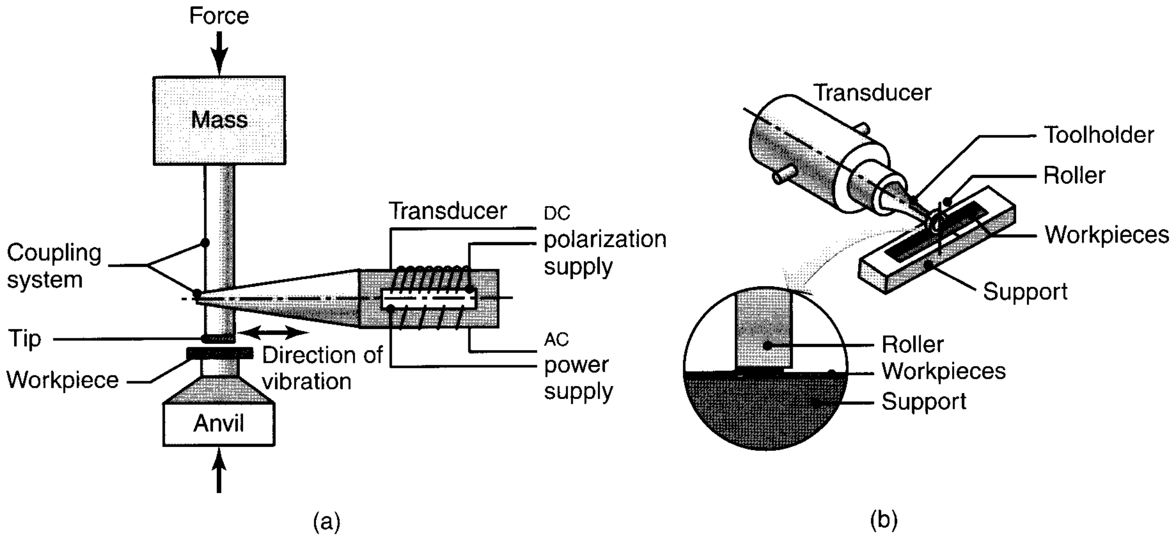


FIGURE 31.2 (a) Components of an ultrasonic-welding machine for making lap welds. The lateral vibrations of the tool tip cause plastic deformation and bonding at the interface of the workpieces. (b) Ultrasonic seam welding using a roller as the *sonotrode*.

31.4 Friction Welding

In the joining processes described thus far, the energy required for welding (typically chemical, electrical, or ultrasonic energy) is supplied from external sources. In *friction welding* (FRW), the heat required for welding is generated through (as the name implies) friction at the interface of the two components being joined. You can demonstrate the significant rise in temperature caused by friction by rubbing your hands together or by sliding down a rope rapidly.

In friction welding, developed in the 1940s, one of the workpiece components remains stationary while the other is placed in a chuck or collet and rotated at a high constant speed. The two members to be joined are then brought into contact under an axial force (Fig. 31.3). The surface speed of the rotating parts may be as high as 900 m/min. 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 removed by the radially outward movement of the hot metal at the interface.

The rotating member must be clamped securely to the chuck or collet to resist both torque and axial forces without slipping. The pressure at the interface and the resulting friction produce sufficient heat for a strong joint to form.

The weld zone usually is confined to a narrow region; its size depends on the following parameters:

- Amount of heat generated.
- Thermal conductivity of the materials.
- Mechanical properties of the materials being joined at elevated temperatures.

The shape of the welded joint depends on the rotational speed and on the axial pressure applied (Fig. 31.4). These factors must be controlled to obtain a uniform,

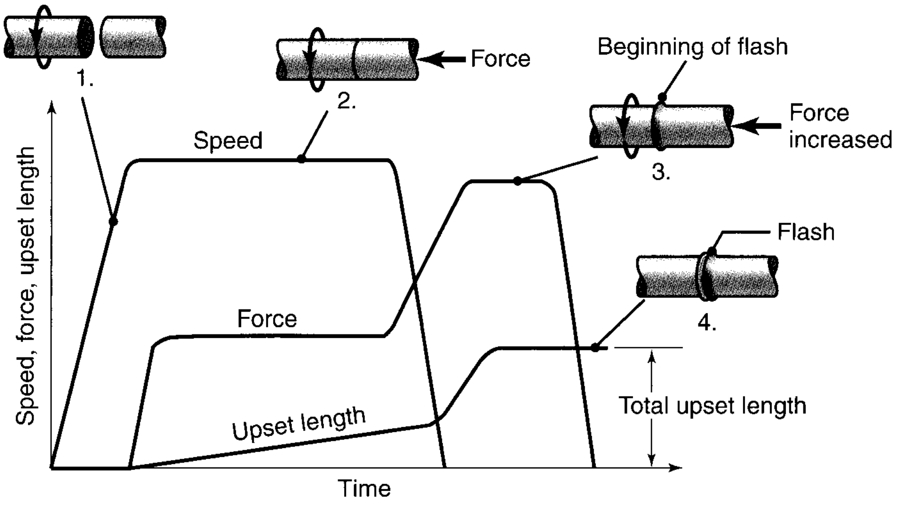
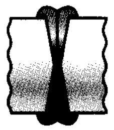


FIGURE 31.3 Sequence of operations in the friction-welding process: (1) The part on the left is rotated at high speed; (2) The part on the right is brought into contact with the part on the left under an axial force; (3) The axial force is increased, and the part on the left stops rotating; flash begins 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.

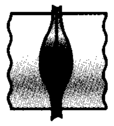
strong joint. The radially outward movement of the hot metal at the interface pushes oxides and other contaminants out of the interface.

Friction welding can be used to join a wide variety of materials, provided that one of the components has some rotational symmetry. Solid or tubular parts can be joined by this method 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.

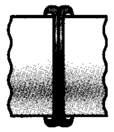
The surface speed of the rotating member may be as high as 15 m/s. Because of the combined heat and pressure, the interface in frictional welding develops a flash by plastic deformation (upsetting) of the heated zone. This flash (if objectionable) can easily be removed by machining or grinding. Friction-welding machines are fully automated, and the operator skill required is minimal—once individual cycle times for the complete operation are set properly.



(a) High pressure or low speed



(b) Low pressure or high speed



(c) Optimum

FIGURE 31.4 Shape of the fusion zones in friction welding as a function of the axial force applied and the rotational speed.

Inertia Friction Welding. This process is a modification of friction welding, although the two terms have been used interchangeably. The energy required for frictional heating in inertia friction welding is supplied by the kinetic energy of a flywheel. The flywheel is accelerated to the proper speed, the two members are brought into contact, and an axial force is applied. As friction at the interface slows the flywheel, the axial force is increased. 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 (the levels depend on the workpiece size and properties). In one application of inertia friction welding, 10-mm-diameter shafts are welded to automotive turbocharger impellers at a rate of one joint every 15 seconds.

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. In *linear friction welding*, the components do not have to be circular or tubular in their cross section. The process is capable of welding square or rectangular components (as well as round parts) made of metals or plastics. In this process, one part is moved across the face of the other part by a balanced reciprocating mechanism.

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. Various other metal parts, with rectangular cross sections as large as 50 × 20 mm have been welded successfully.

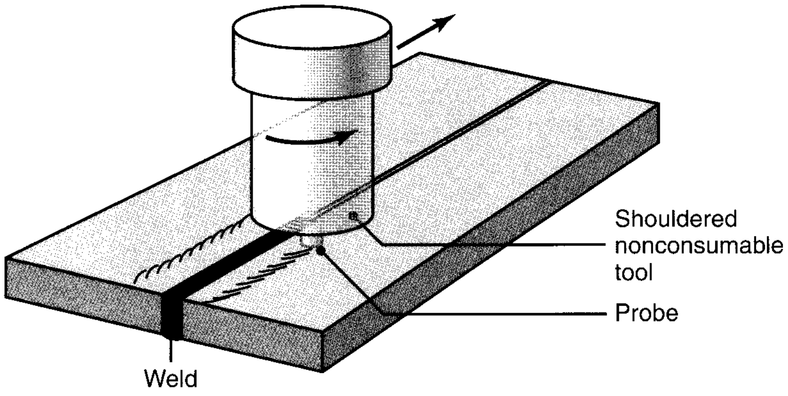


FIGURE 31.5 The principle of the friction-stir-welding process. Aluminum-alloy plates up to 75 mm thick have been welded by this process.

Friction Stir Welding. In conventional friction welding, heating of an interface is achieved through friction by rubbing two contacting surfaces. In the *friction-stir-welding* (FSW) process, developed in 1991, a third body is rubbed against the two surfaces to be joined. A rotating nonconsumable probe, typically 5 to 6 mm in diameter and 5 mm high, is plunged into the joint (Fig. 31.5). The contact pressure causes frictional heating, raising the temperature to between 230° and

260°C. The probe at the tip of the rotating tool forces mixing (or stirring) of the material in the joint.

Materials such as aluminum, copper, steel, and titanium have been welded successfully, and developments are taking place to extend FSW applications to polymers and composite materials. The process is now being applied to aerospace, automotive, shipbuilding, and military vehicles, using sheet or plates. With developments in rotating-tool design, other possible applications include inducing microstructural changes, refining grain in materials, and improving localized toughness in castings.

The welding equipment can be a conventional, vertical-spindle milling machine (Fig. 24.15b), and the process is relatively easy to implement. The thickness of the welded material can be as little as 1 mm and as much as 50 mm welded in a single pass. Welds produced by friction stir welding have high quality, minimal pores, and a uniform material structure. The welds are produced with low heat input and therefore low distortion and little microstructural changes. No shielding gas or surface cleaning is required.

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 not requiring consumable electrodes, shielding gases, or flux.

The heat generated in resistance welding is given by the general expression

$$H = I^2Rt, \tag{31.1}$$

where

H = Heat generated in joules (watt-seconds)
I = Current (in amperes)
R = Resistance (in ohms)
t = 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 denotes the energy losses through conduction and radiation. The equation then becomes

$$H = I^2RtK, \tag{31.2}$$

where the value of K is less than unity.

The *total resistance* is the sum of the following properties (see Fig. 31.6):

- **a.** Resistances of the electrodes;
- **b.** Electrode–workpiece contact resistance;
- **c.** Resistances of the individual parts to be welded;
- **d.** Contact resistance between the two workpieces to be joined (*faying surfaces*).

The actual temperature rise at the joint depends on the specific heat and the thermal conductivity of the metals to be joined. For example, metals such as aluminum and copper have high thermal conductivity, so they require high heat concentrations. Similar or dissimilar metals can be joined by resistance welding. The magnitude of the current in resistance-welding operations may be as high as 100,000 A, but the voltage is typically only 0.5 to 10 V.

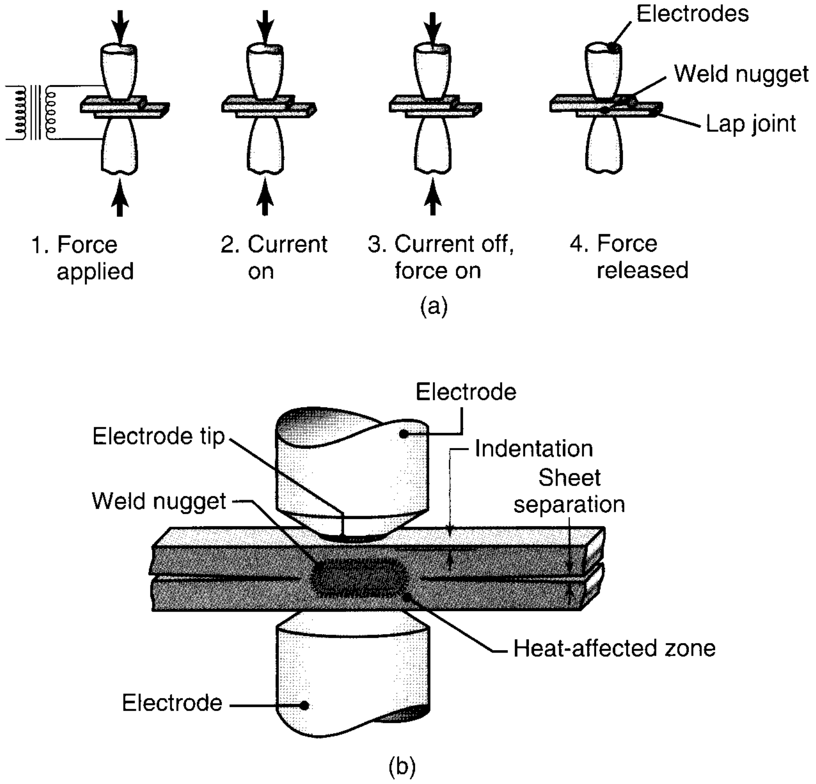


FIGURE 31.6 (a) Sequence of events in resistance spot welding. (b) Cross section of a spot weld, showing the weld nugget and the indentation of the electrode on the sheet surfaces. This is one of the most commonly used processes in sheet-metal fabrication and in automotive-body assembly.

The strength of the bond depends on surface roughness and on the cleanliness of the mating surfaces. Oil films, paint, and thick oxide layers should therefore be removed before welding. The presence of uniform, thin layers of oxide and of other contaminants is not as critical.

Developed in the early 1900s, resistance-welding processes require specialized machinery. Much of it is now operated by programmable computer control. Generally, the machinery is not portable, and the process is suitable primarily for use in manufacturing plants and machine shops. The operator skill required is minimal, particularly with modern machinery.

31.5.1 Resistance Spot Welding

In *resistance spot welding* (RSW), the tips of two opposing solid, cylindrical electrodes touch a lap 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. Accurate control and timing of the alternating electric current and of the pressure are essential in resistance welding. In the automotive industry, for example, the number of cycles ranges up to about 30 at a frequency of 60 Hz. (See also *high-frequency resistance welding* in Section 31.5.3.)

The weld nugget (Fig. 31.6b) is generally 6 to 10 mm in diameter. The surface of the spot weld has a slightly discolored indentation. Currents range from 3000 to

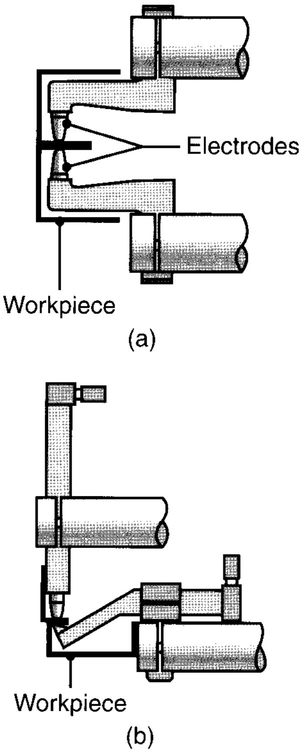
40,000 A. The current level depends on the materials being welded and on their thicknesses. For example, the current is typically 10,000 A for steels and 13,000 A for aluminum. Electrodes generally are made of copper alloys and must have sufficient electrical conductivity and hot strength to maintain their shape.

Spot welding is the simplest and most commonly used resistance-welding process. Welding may be performed by means of single (most common) or multiple pairs of electrodes (as many as a hundred or more), and the required pressure is supplied through mechanical or pneumatic means. **Rocker-arm-type** spot-welding machines normally are used for smaller parts; **press-type** machines are used for larger workpieces. The shape and surface condition of the electrode tip and the accessibility of the site are important factors in spot welding. A variety of electrode shapes are used to spot-weld 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 large sheet-metal structures. Modern spot-welding equipment is computer controlled for optimum timing of current and pressure; its spot-welding guns are manipulated by programmable robots (Fig. 31.8c). Automobile bodies can have as many as 10,000 spot welds; they are welded at high rates with the use of multiple electrodes (see Fig. I.10 in the General Introduction).

Testing Spot Welds. Spot-welded joints may be tested for weld-nugget strength by means of the following techniques (Fig. 31.9):

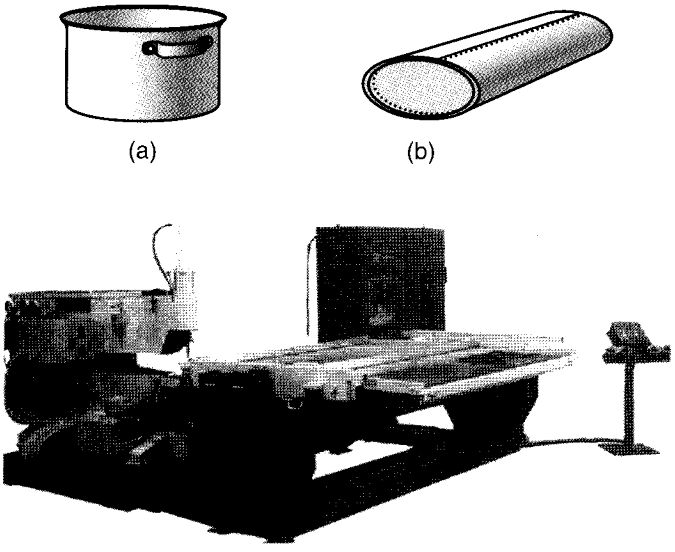
- Tension-shear
- Cross-tension
- Twist
- Peel.



(a)

(b)

FIGURE 31.7 Two electrode designs for easy access to the components to be welded.


(a) (b)

(c)

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. *Source*: Courtesy of Taylor–Winfield Corporation.

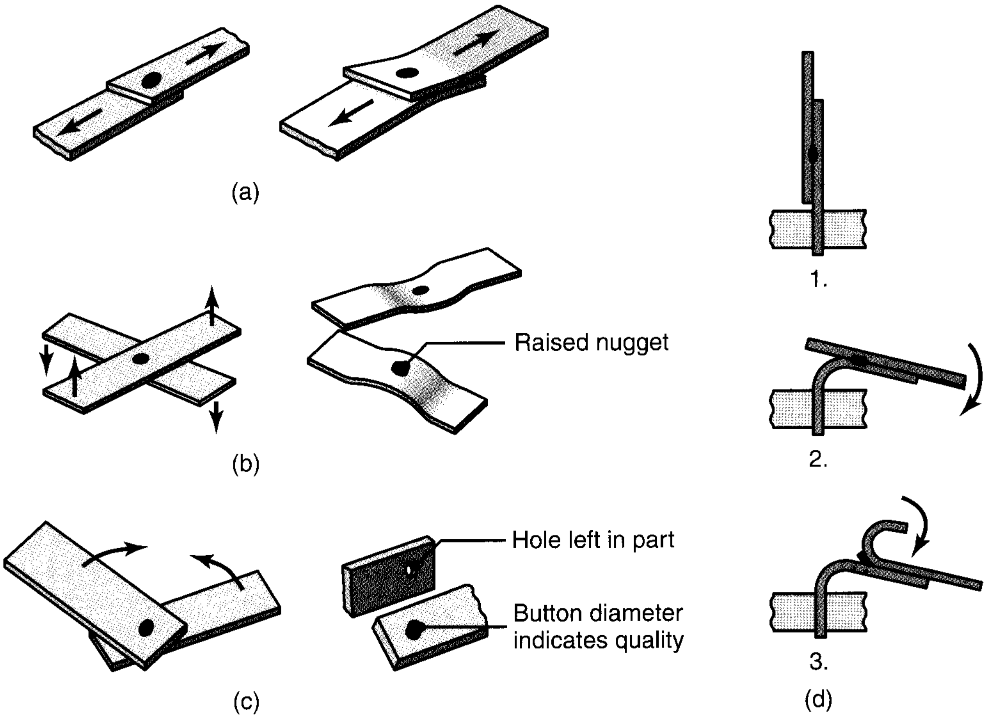


FIGURE 31.9 Test methods for spot welds: (a) tension-shear test, (b) cross-tension test, (c) twist test, (d) peel test (see also Fig. 32.9).

Because they are easy to perform and are inexpensive, tension-shear tests are commonly used in fabricating facilities. The cross-tension and twist tests are capable of revealing 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.

EXAMPLE 31.2 Heat Generated in Spot Welding

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 second by means of electrodes 5 mm in diameter. Estimate the heat generated and its distribution in the weld zone if the effective resistance in the operation is 200 $\mu\Omega$.

Solution According to Eq. (31.1),

$$\text{Heat} = (5000)^2(0.0002)(0.1) = 500\text{ J}.$$

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 (0.008 g/mm^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

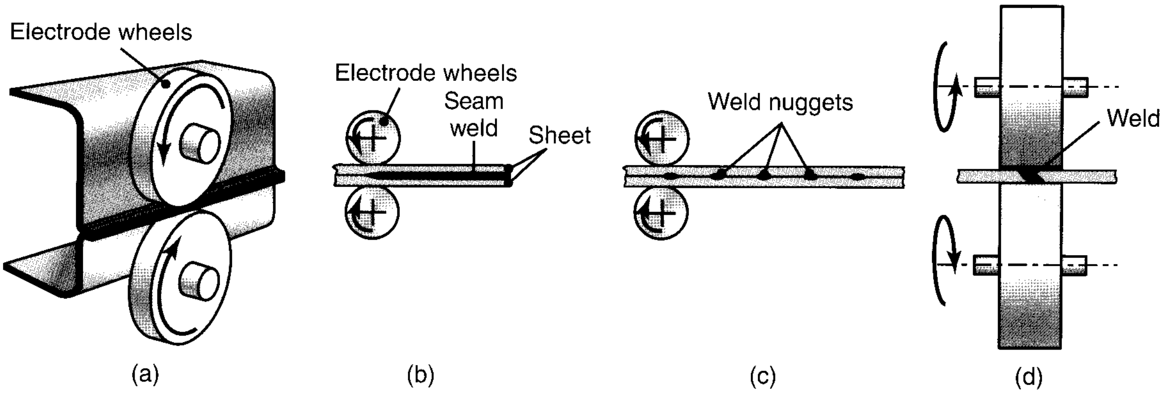


FIGURE 31.10 (a) Seam-welding process in which rotating rolls act as electrodes. (b) Overlapping spots in a seam weld. (c) Roll spot welds and (d) Mash seam welding.

the current reaches a sufficiently high level in the AC cycle. With a high enough frequency or slow enough traverse speed, these spot welds actually overlap into a continuous seam and produce a joint that is liquid tight and gastight (Fig. 31.10b).

In **roll spot welding**, current to the rollers is applied only intermittently, resulting in 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, and the welded seam thickness is only 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 (see Section 16.2.2).

The RSEW process is used to make the longitudinal (side) seam of cans (for household products) mufflers, gasoline tanks, and other containers. The typical welding speed is 1.5 m/min for thin sheets.

31.5.3 High-frequency Resistance Welding

High-frequency resistance welding (HFRW) is similar to seam welding, except that high-frequency current (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 then are pressed together by passing the tube through a pair of squeeze rolls. Any flash formed is then trimmed off.

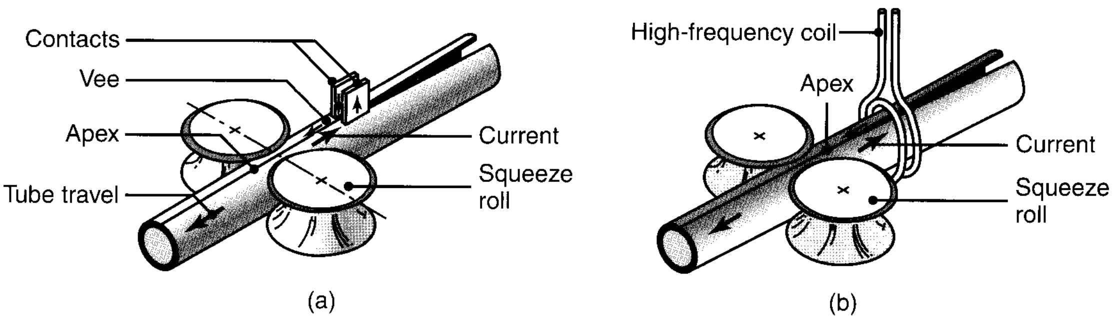


FIGURE 31.11 Two methods of high-frequency continuous butt welding of tubes.

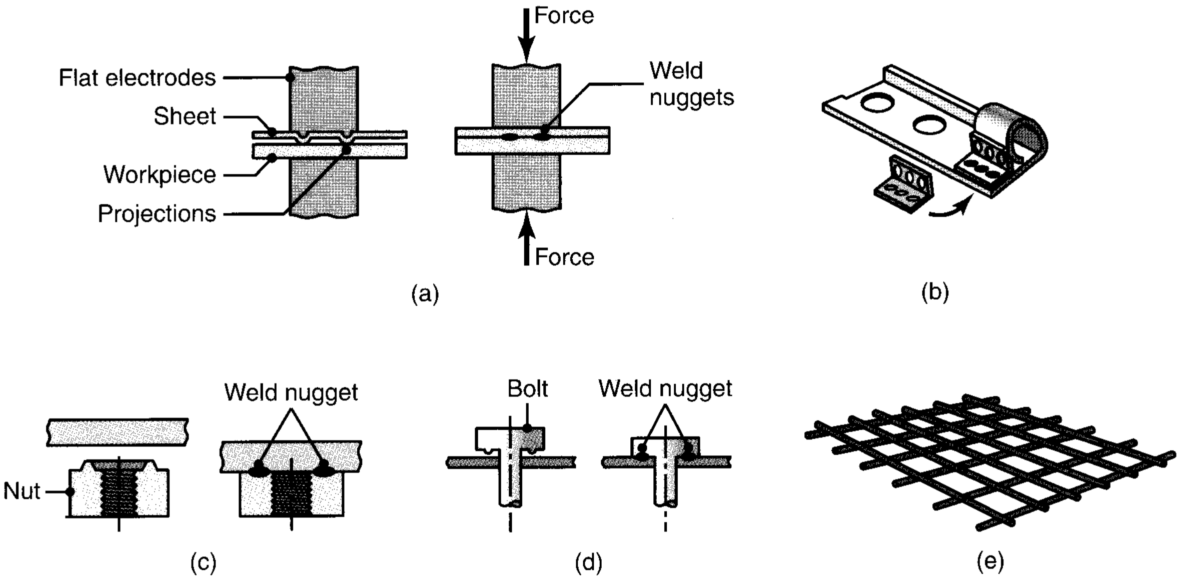


FIGURE 31.12 (a) Schematic illustration of resistance projection welding. (b) A welded bracket. (c) and (d) Projection welding of nuts or threaded bosses and studs. (e) Resistance-projection-welded grills.

Structural sections (such as I-beams) can be fabricated by HFRW by welding the webs and flanges made from long, flat pieces. Spiral pipe and tubing, finned tubes for heat exchangers, and wheel rims also may be made by this technique. In another method, called **high-frequency induction welding** (HFIW), the roll-formed tube is 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. 16.36) 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. The electrodes (typically made of copper-based alloys) are large and flat, and water cooled to keep their temperature low. Weld nuggets similar to those in spot welding are formed as the electrodes exert pressure to soften and compress the projections.

Spot-welding equipment can be used for resistance projection welding by modifying the electrodes. Although the embossing of the workpieces adds expense, the process produces a number of welds in one pass, extends electrode life, and is capable of welding metals of different thicknesses, such as a sheet welded over a plate. Nuts and bolts can be welded to sheets and plates by this process (Figs. 31.12c and d), with projections that are produced by machining or forging. Joining a network of rods and wires (such as the ones making up metal baskets, grills (Fig. 31.12e) oven racks, and shopping carts) also is considered resistance projection welding, because of the many small contact areas between crossing wires (grids).

31.5.5 Flash Welding

In *flash welding* (FW), also called **flash butt welding**, heat is generated very rapidly from the arc as the ends of the two members begin to make contact and develop an electrical resistance at the joint (Fig. 31.13a). After the proper temperature is

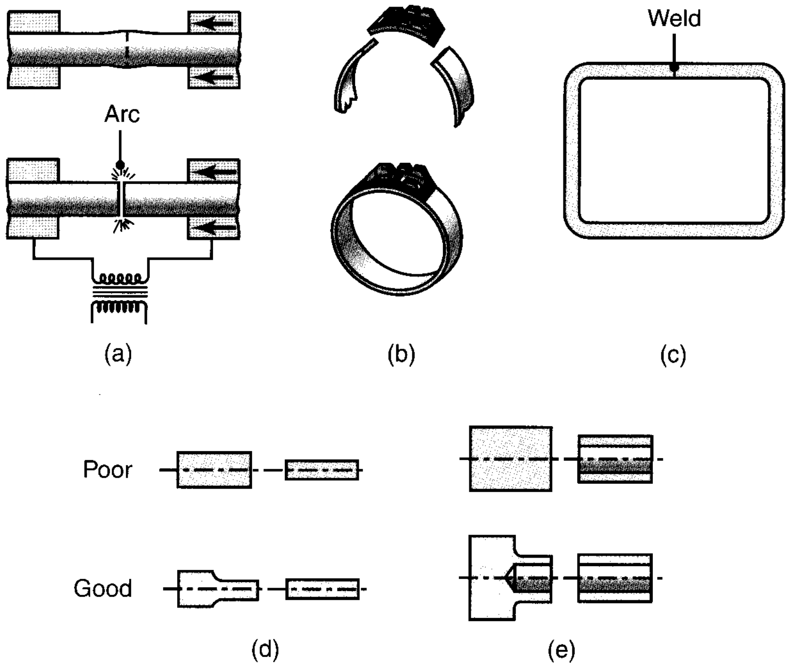


FIGURE 31.13 (a) Flash-welding process for end-to-end welding of solid rods or tubular parts. (b) and (c) Typical parts made by flash welding. (d) and (e) Some design guidelines for flash welding.

reached and the interface begins to soften, an axial force is applied at a controlled rate and a weld is formed by plastic deformation of the joint. The mechanism is called *hot upsetting*, and 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—hence 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; therefore, the quality of the weld is good. However, a significant amount of material may be burned off during the welding process. The joint may be machined later to improve its appearance. The machines for flash welding usually are automated and large and have a variety of power supplies ranging from 10 to 1500 kVA.

The flash-welding process is suitable for end-to-end or edge-to-edge joining of sheets of similar or dissimilar metals 0.2 to 25 mm thick and for end-joining bars 1 to 75 mm in diameter. Thinner sections have a tendency to buckle under the axial force applied during welding. Rings made by forming processes (such as those shown in Fig. 16.22) also can be flash butt welded. In addition, the process is used to repair broken band-saw blades with the use of fixtures that are mounted on the band-saw frame.

The flash-welding process can be automated for reproducible welding operations. Typical applications are the joining of pipe and of tubular shapes for metal furniture and windows. The process is also used for welding the ends of sheets or coils of wire in continuously operating rolling mills (Chapter 13) and in the feeding of wire-drawing equipment (Chapter 15). Once the appropriate process parameters are established, the required operator skill is minimal. Some design guidelines for mating surfaces in flash welding are shown in Figs. 31.13d and e. Note the importance of uniform cross sections at the joint.

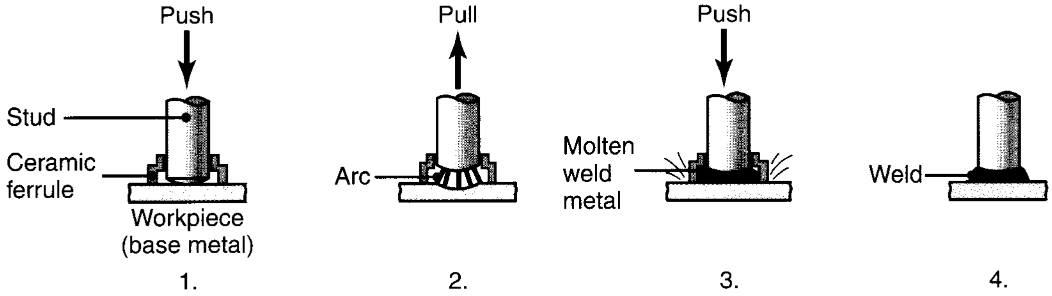


FIGURE 31.14 The sequence of operations in stud welding commonly used for welding bars, threaded rods, and various fasteners onto metal plates.

31.5.6 Stud Welding

Stud welding (SW) is also called *stud arc welding* and is similar to flash welding. The stud (which may be a small part or, more commonly, a threaded rod, hanger, or handle) serves as one of the electrodes while being joined to another component, which is usually a flat plate (Fig. 31.14). Polarity for aluminum is usually direct-current electrode positive (DCEP), and for steel it is direct-current electrode negative (DCEN).

In order to concentrate the heat generated, 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 also is available. Typical applications of stud welding include automobile bodies, electrical panels, and shipbuilding; the process is also used 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 very short—on the order of 1 to 6 milliseconds. The choice between this process and stud arc welding depends on such factors as the types of metals to be joined, the workpiece thickness and cross section, the stud diameter, and the shape of the joint.

31.5.7 Percussion Welding

The resistance-welding processes already described usually employ an electrical transformer to meet the power requirements. Alternatively, the 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 to develop localized high heat at the joint. The process is useful where heating of the components adjacent to the joint is to be avoided, as 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 for household products have been resistance seam welded (with a lap joint up the side of the can) for many years. Beginning in about 1987, 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. Thus, some material is saved. Multiplied by the

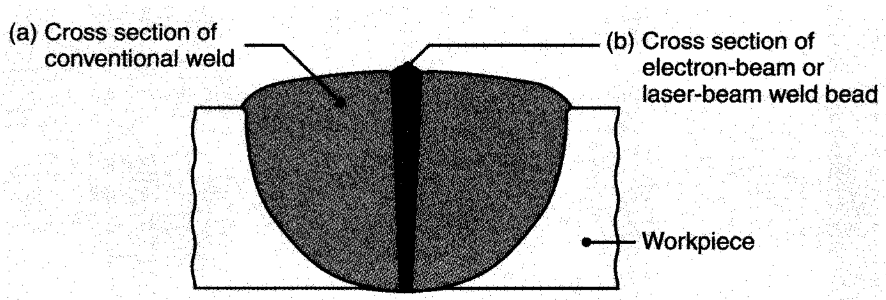


FIGURE 31.15 The relative sizes of the weld beads obtained by tungsten-arc and by electron-beam or laser-beam welding.

billions of cans made each year, this amount becomes a very significant savings.

- Because laser welds have a very narrow zone (Fig. 31.15; see also Fig. 30.14), the unprinted area on the can surface (the printing margin) is greatly reduced. As a result, the can's appearance and its customer acceptance are improved.
- The resistance lap-welded joint can be subject to corrosion by the contents of the can (e.g., tomato juice). This effect may change the taste and can cause a potential liability risk. 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 that has been 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.

This 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 of the interface. As a result, the bond strength from explosion welding is very high.

The explosive may be a flexible plastic sheet or cord or in granulated or liquid form, which is cast or pressed onto the flyer plate. The detonation speed is usually in the range from 2400 to 3600 m/s; it depends on the type of explosive, the thickness of the explosive layer, and the packing density of the layer. There is a minimum denotation speed necessary for welding to occur in this process. Detonation is carried out with a standard commercial blasting cap.

This process is suitable particularly for cladding a plate or a slab with a dissimilar metal. Plates as large as 6 × 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. Explosion welding is inherently dangerous, so it requires safe handling by well-trained and experienced personnel.

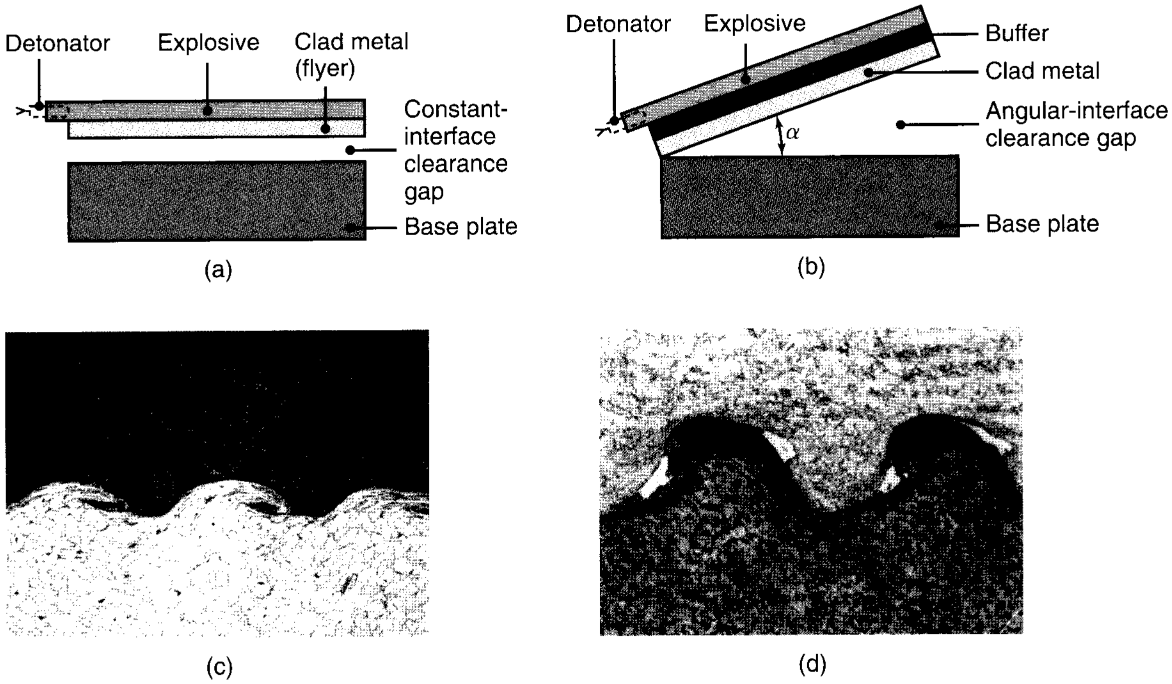


FIGURE 31.16 Schematic illustration of the explosion-welding process: (a) constant-interface clearance gap and (b) angular-interface clearance gap. (c) Cross section of explosion-welded joint: titanium (top) and low-carbon steel (bottom). (d) Iron–nickel alloy (top) and low-carbon steel (bottom).

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 the interface) and secondarily from plastic deformation of the faying surfaces. This 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. (See also Sections 1.7 and 1.8.)

The bonded 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) how clean the faying surfaces are. 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 electroplating 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 to be joined. The parts usually are heated in a furnace or by electrical resistance. High-pressure autoclaves also are used for bonding complex parts.

Although this process was developed in the 1970s as a modern welding technology, the principle of diffusion bonding dates back centuries to when goldsmiths bonded gold over copper to create a product called **filled gold**. First, a thin layer of gold foil is produced and placed over copper, and a weight is placed on top of the foil. Finally, the assembly is placed in a furnace and left until a strong bond is obtained; hence, the process is also called *hot-pressure welding* (HPW).

Diffusion bonding generally is most suitable for joining dissimilar metals. It also is 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 also an important mechanism of sintering in powder metallurgy (Section 17.4). Because diffusion involves migration of the atoms across the joint, the process is slower than other welding processes.

Although diffusion welding is used for fabricating complex parts in low quantities for the aerospace, nuclear, and electronics industries, it has been automated to make it suitable and economical for moderate-volume production. Unless the process is highly automated, considerable operator training and skill are required. Equipment cost is related approximately to the diffusion-bonded area and is in the range of \$3 to \$6/mm^2.

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 the conservation of expensive strategic materials and the reduction of manufacturing costs. The military aircraft illustrated in Fig. 31.17 has more than 100 diffusion-bonded parts, some of which are shown.

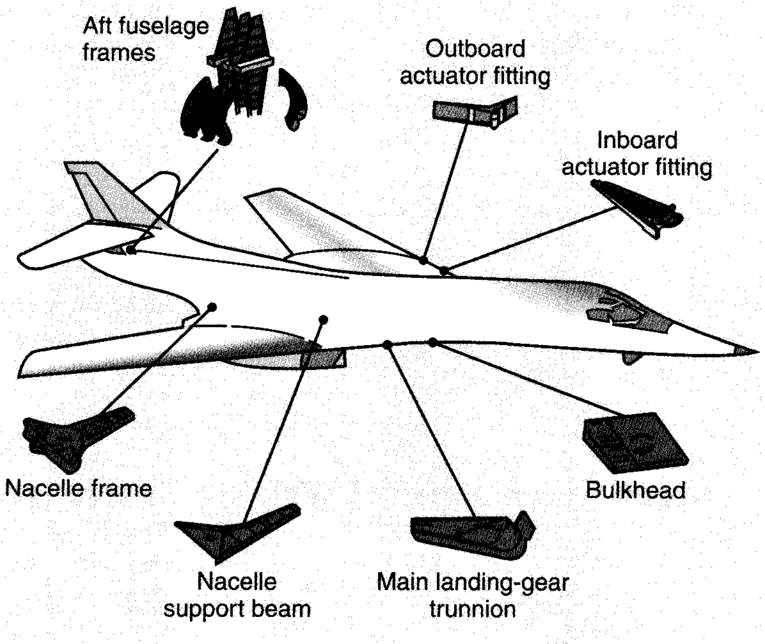


FIGURE 31.17 Aerospace diffusion bonding applications.

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 (usually) are diffusion bonded and formed are shown in Fig. 31.18. After the diffusion bonding of selected locations on the sheets, the unbonded (*stop-off*) regions are expanded in a mold by air or fluid pressure. These structures are thin and have high stiffness-to-weight ratios; hence, they are particularly useful in aircraft and aerospace applications.

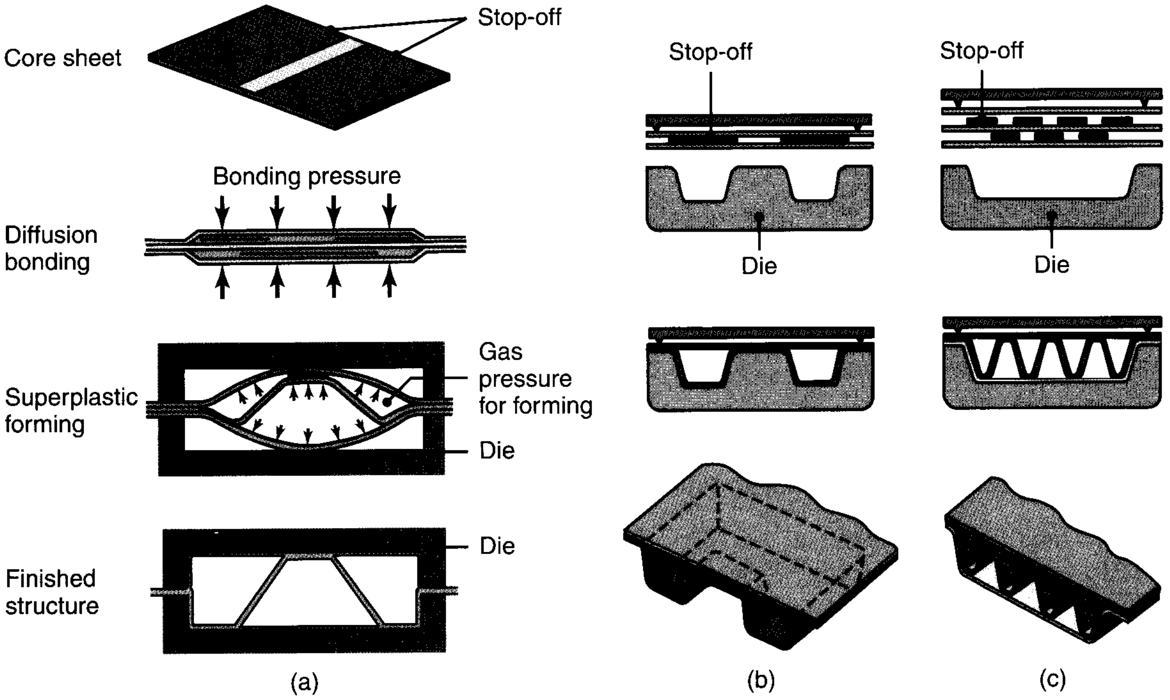


FIGURE 31.18 The sequence of operations in the fabrication of a structure by the diffusion bonding and superplastic forming of three originally flat sheets. See also Fig. 16.48. *Sources*: (a) After D. Stephen and S.J. Swadling. (b) and (c) Courtesy of Rockwell International Corp.

Diffusion bonding–superplastic forming improves productivity by eliminating the number of parts in a structure, mechanical fasteners, labor, and cost. It produces parts with good dimensional accuracy and low residual stresses. First developed in the 1970s, this technology is now well advanced for titanium structures (typically using Ti-6Al-4V and 7475-T6) and various other alloys for aerospace applications.

31.8 Economics of Welding Operations

The characteristics, advantages, and limitations of the welding processes described thus far have included brief discussions regarding welding costs. The relative costs of some selected processes are shown in Tables 30.1 and VI.1. As in all other manufacturing operations, costs in welding and joining processes can vary widely, depending on such factors as the 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 particular joining process.

The general *welding and joining costs* for some common operations (all described throughout Chapters 30 through 32) can be summarized as follows:

- *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 automate.

Equipment costs for welding can be summarized as follows (note that these costs can vary widely):

- *High* ($100,000 to $200,000): electron-beam and laser-beam welding.
- *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 the operator skill, welding time, and preparation required. However, much depends on the automation of the equipment employed, including the use of robotics and computer controls programmed to follow a prescribed path (*seam tracking*) during welding. It has been observed that in systems with robotic controls, the number reaches 80%, whereas in manual welding operations (see Table 30.1), the actual welding time spent by the operator is only about 30% of the overall time.

Labor costs may be summarized as follows:

- *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 Friction Welding of Pistons

There has been a sustained effort among heavy-truck manufacturers to design and manufacture diesel engines with reduced emissions. A number of technologies have become more prevalent since the 1980s, reflecting the need for *green design* (see Sections I.4 and 40.4). Exhaust-gas recirculation (the reintroduction of a portion of the spent exhaust gases into the intake stream of the engine) has become standard and is known to reduce nitrous-oxide emissions. Unfortunately, this strategy leads to less efficient combustion and lower component durability because of the abrasive-wear particles (see Section 33.5) and acids that are recirculated into the engine. To maintain and even improve engine efficiency, engine manufacturers have increased cylinder pressures and operating temperatures, which together lead to an even more demanding environment for engine components.

In the U.S. market, the traditional aluminum pistons in diesel engines were found to be unable to function reliably in modern engine designs. The problems identified with pistons were a tendency to "mushroom" and fracture under the high firing pressures in the cylinder, inadequate cooling of the piston, and scuffing (wear) at the pin that joins the piston to the connecting rod. A solution, shown in Fig. 31.19, is a Monosteel® piston, which has the following design attributes:

- The piston is produced from steel, which has higher strength and better high-temperature mechanical properties than the aluminum alloys previously used (see Section 6.2).
- A two-piece design allows the incorporation of an oil gallery, permitting circulation of cooling oil in the piston. One of the main advantages of the Monosteel® design is the use of a very large gallery, resulting in effective heat transfer from the piston. This design has been shown to reduce piston temperatures in the rim by around 40°C compared with earlier piston designs.
- The piston's steel skirt is much more rigid than the aluminum skirt, resulting in smaller deformation and allowing for designs with tighter clearances. This feature results in a more stable piston with less oil consumption (thus further reducing harmful exhaust emissions).

Monosteel® pistons are produced from two forged components, which are machined prior to welding. The process used to join these components is inertia friction welding (Section 31.4), which has the following advantages in this application:

- The process leads to well-controlled, reliable, and repeatable high-quality welds.
- Friction welds are continuous and do not involve any porosity, thereby producing a high-strength weld that seals the oil gallery.
- The welding process is fairly straightforward to optimize, the main process variables being

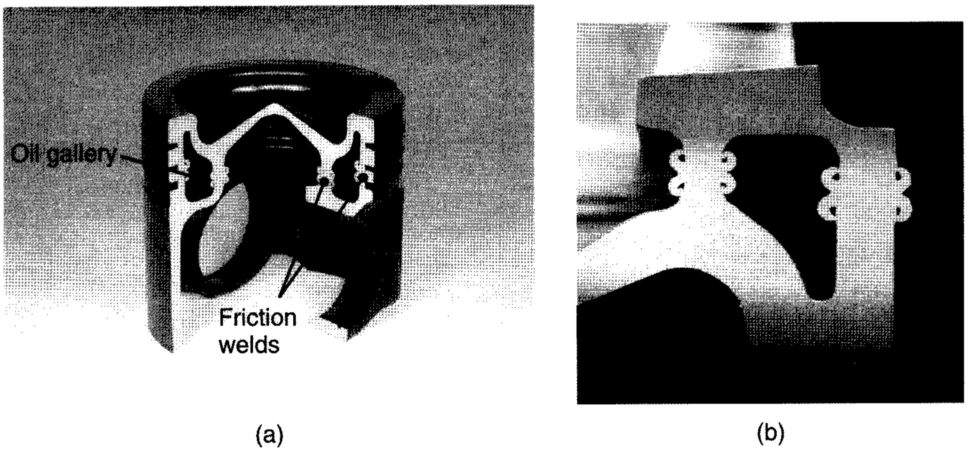


FIGURE 31.19 The Monosteel® piston. (a) Cutaway view of the piston, showing the oil gallery and friction-welded sections; (b) detail of the friction welds before the external flash is removed by machining; note that this photo is a reverse of the one on the left.

energy (or spindle speed for a given flywheel) and contact pressure.

The friction-welding process does not require any operator intervention or expertise, as it is entirely machine controlled. Although the capital investment is significant compared with that of other welding technologies, weld quality (and the ability to weld in this application) is significantly more favorable. The Monosteel® piston shown in Fig. 31.19 was produced on a 230-metric-ton force capacity inertia friction welder using a peripheral velocity of 7.5 m/s and a contact pressure of 140 MPa (see Fig. 31.3). As can be seen, the weld zone contains the optimum flash shape (see Fig. 31.4), which is removed from the exterior piston surface by a turning operation (Section 23.2), after which the piston skirt is ground (Section 26.3). Production takes place at relatively high rates; 40–60-second cycle times are typical, but can be higher or lower depending on piston size. The Monosteel® piston has been applied to multiple engine platforms and has been in high-volume production since 2003.

Source: Courtesy of D. Adams, Manufacturing Technology, Inc., and K. Westbrooke, Federal Mogul, Inc.

SUMMARY

- In addition to the traditional joining processes of oxyfuel–gas and arc welding, a number of other joining processes that are based on producing a strong joint under pressure and/or heat 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 electrical resistance or furnaces), or it may be generated internally (as in friction welding).
- Among important developments is the combining of diffusion-bonding and superplastic-forming processes. Productivity is improved, as is the capability to make complex parts economically.
- As in all manufacturing operations, certain hazards are inherent in welding operations. Some relate to the machinery and equipment used, others to the nature of the process. Proper safety precautions always must be taken in work areas where welding is done.

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
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. 6: **Welding, Brazing, and Soldering**, ASM International, 1993.

Bowditch, W.A., and Bowditch, K.E., **Welding Technology Fundamentals**, Goodheart–Willcox, 1997.

Cary, H.B., **Modern Welding Technology**, 6th ed., Prentice Hall, 2004.

Evans, G.M., and Bailey, N., **Metallurgy of Basic Weld Metal**, Woodhead Publishing, 1997.

Hicks, J.G., **Welded Design: Theory and Practice**, William Andrews, 2005.

Houldcroft, T., **Welding and Cutting: A Guide to Fusion Welding and Associated Cutting Processes**, Industrial Press, 2001.

Jeffus, L.F., **Welding: Principles and Applications**, 6th ed., Delmar Publishers, 2007.

Lagoda, T., **Life Estimation of Welded Joints**, Springer, 2008.

Mathers, G., **Welding of Aluminum and its Alloys**, CRC Press, 2002.

Mishra, R.S., and Mahoney, M.W., **Friction Stir Welding and Processing, American Society for Metals**, 2007.

Mouser, J.D., **Welding Codes, Standards, and Specifications**, McGraw-Hill, 1997.

Nicholas, M.G., **Joining Processes: Introduction to Brazing and Diffusion Bonding**, Chapman & Hall, 1998.

Weman, K., **Welding Process Handbook**, CRC Press, 2003.

Zhang, H., and Senkara, J., **Resistance Welding: Fundamentals and Applications**, CRC Press, 2005.

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. Describe the advantages and limitations of explosion welding.

31.7. Describe the principle of resistance-welding processes.

31.8. What type of products are suitable for stud welding? Why?

31.9. What is the advantage of linear friction welding over inertia friction welding?

31.10. Describe how high-frequency butt welding operates.

QUALITATIVE PROBLEMS

31.11. Explain the reasons that the processes described in this chapter were developed.

31.12. Explain the similarities and differences between the joining processes described in this chapter and those described in Chapter 30.

31.13. Describe your observations concerning Fig. 31.16c and d.

31.14. Would you be concerned about the size of weld beads, such as those shown in Fig. 31.15? Explain.

31.15. What advantages does friction welding have over other methods described in this and in the preceding chapter?

31.16. Describe the significance of faying surfaces.

31.17. Discuss the factors that influence the strength of (a) a diffusion-bonded and (b) a cold-welded component.

31.18. What are the sources of heat for the processes described in this chapter?

31.19. Can the roll-bonding process be applied to a variety of part configurations? Explain.

31.20. Why is diffusion bonding, when combined with the superplastic forming of sheet metals, an attractive fabrication process? Does it have any limitations?

31.21. List and explain the factors involved in the strength of weld beads.

31.22. Give some of the reasons that spot welding is used commonly in automotive bodies and in large appliances.

31.23. Explain the significance of the magnitude of the pressure applied through the electrodes during a spot-welding operation.

31.24. Give some applications for (a) flash welding, (b) stud welding, and (c) percussion welding.

31.25. Discuss the need for, and role of, work-holding devices in the welding operations described in this chapter.

31.26. 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.

31.27. Could the process shown in Fig. 31.11 also be applicable to part shapes other than round? Explain, and give specific examples.

31.28. In spot-weld tests, what would be the reason for weld failure to occur at the locations shown in Fig. 31.9?

QUANTITATIVE PROBLEMS

31.29. The energy required in ultrasonic welding is found to be related to the product of workpiece thickness and hardness. Explain why this relationship exists.

31.30. 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 $\mu\Omega$.

31.31. Calculate the temperature rise in Problem 31.30, assuming that the heat generated is confined to the volume of material directly between the two round electrodes and the temperature is distributed uniformly.

31.32. Calculate the range of allowable currents in Problem 31.30 if the temperature should be between 0.7 and 0.8 times the melting temperature of copper. Repeat this problem for carbon steel.

SYNTHESIS, DESIGN, AND PROJECTS

31.33. Comment on workpiece size and shape limitations (if any) for each of the processes described in this chapter.

31.34. Explain how you would fabricate the structures shown in Fig. 31.18 by methods other than diffusion bonding and superplastic forming.

31.35. Describe part shapes that cannot be joined by the processes described in this chapter. Gives specific examples.

31.36. Comment on the feasibility of applying explosion welding in a factory environment.

31.37. Discuss your observations concerning the welding design guidelines illustrated in Figs. 31.13d and e.

31.38. Referring to Fig. 14.11b, 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.

31.39. Explain how the projection-welded parts shown in Fig. 31.12 could be made by any of the processes described in this book.

31.40. Using a magnifier, inspect the cross sections of coins such as the U.S. dime and nickel, and comment on your observations.

31.41. Describe the methods you would use for removing the flash from welds, such as those shown in Fig. 31.3. How would you automate these methods for a high-production facility?

31.42. 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.

31.43. 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.

31.44. 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?

31.45. 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.

Brazing, Soldering, Adhesive-Bonding, and Mechanical-Fastening Processes

CHAPTER 32

- This last chapter of Part VI describes various joining, bonding, and fastening processes that involve mechanisms unlike those in the preceding two chapters.
- Brazing and soldering are different from welding in that no diffusion takes place at the interface, thus bond strength depends on adhesive forces.
- Brazing and soldering are differentiated by the temperature at which filler metals melt: Brazing takes place above 450°C and produces stronger joints than soldering, while soldering involves lower temperatures and is widely applied in the electronics industry.
- Adhesive bonding is versatile, and a wide variety of adhesives is available for numerous applications.
- Mechanical joining processes are then described, such as using bolts, nuts, rivets, snap fasteners, or shrink fits in assembly.
- The chapter ends with a discussion of economic considerations in joining operations.

32.1 Introduction

In most of the joining processes described in Chapters 30 and 31, the mating surfaces of the components are heated to elevated temperatures by various external or internal means, to cause fusion and bonding at the joint. But what if we want to join a pair of materials that cannot withstand high temperatures, such as electronic components? What if the parts to be joined are fragile, intricate, or made of two or more materials with very different characteristics, properties, sizes, thicknesses, and cross sections?

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 obtained. Brazing and soldering are distinguished arbitrarily by temperature. Temperatures for soldering are lower than those for brazing, and the strength of a soldered joint is much lower.

The chapter also describes the principles and types of *adhesive bonding*. 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 metallic and nonmetallic materials. The 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, present various challenges.

Although all of the joints described thus far are of a permanent nature, in many applications joined components have to be taken apart for replacement, maintenance, repair, or adjustment. But how, for example, do we take apart a product without destroying the joint? If joints are required that are not permanent, but still must be as strong as welded joints, the obvious solution is to use *mechanical fastening*, such as fastening with bolts, screws, nuts, or a variety of other fasteners.

32.2 Brazing

Brazing is a joining process in which a *filler metal* is placed between the faying surfaces to be joined (or at their periphery) and the temperature is raised sufficiently to melt the filler metal, but not the components (the base metal)—as would be the case in fusion welding. Thus, brazing is a liquid–solid-state bonding process. Upon cooling and solidification of the filler metal, a strong joint is obtained (Fig. 32.1). Filler metals used for brazing typically melt above 450°C, which is below the melting point (*solidus temperature*) of the metals to be joined (see, for example, Fig. 4.5). Brazing is derived from the word *brass*, an archaic word meaning "to harden," and the process was first used as far back as 3000 to 2000 B.C.

In a typical brazing operation, a filler (*braze*) metal wire is placed along the periphery of the components to be joined, as shown in Fig. 32.2a. Heat is then applied

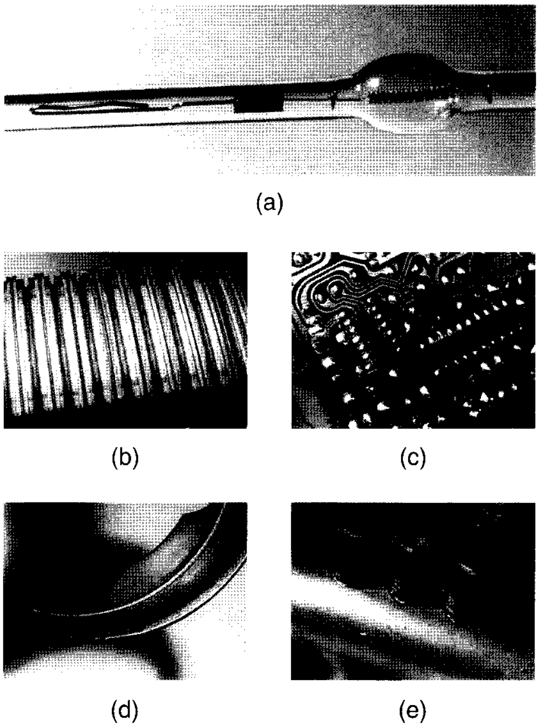


FIGURE 32.1 Examples of brazed and soldered parts. (a) Resistance-brazed light-bulb filament; (b) brazed radiator heat exchanger; (c) soldered circuit board; (d) brazed ring housing; and (e) brazed heat exchanger. *Source*: Courtesy of Edison Welding Institute.

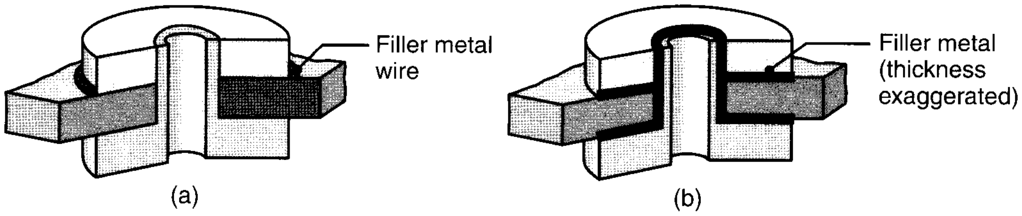


FIGURE 32.2 An example of furnace brazing (a) before and (b) after brazing. The filler metal is a shaped wire and moves into the interfaces by *capillary action* with the application of heat.

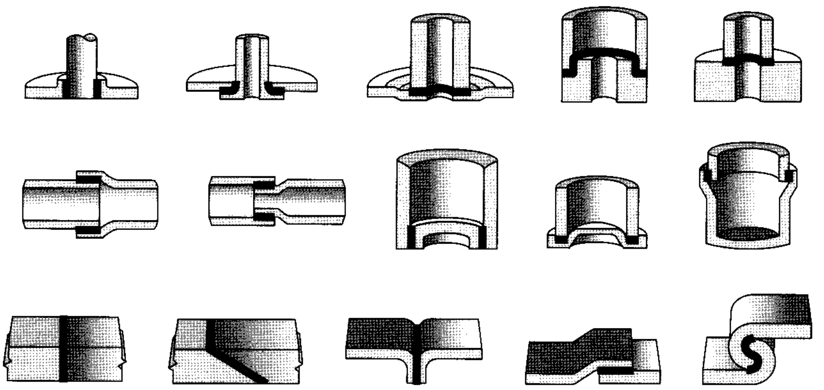


FIGURE 32.3 Joint designs commonly used in brazing operations. The clearance between the two parts being brazed is an important factor in joint strength. If the clearance is too small, the molten braze metal will not penetrate the interface fully. If it is too large, there will be insufficient capillary action for the molten metal to fill the interface.

by various external means, melting the braze metal and, by *capillary* action, filling the closely fitting space (*joint clearance*) at the interfaces (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); the major difference is that the base metal does not melt. The main application of braze welding is in repair work, typically on parts made of cast steels and irons. Because of the wider gaps between the components being welded (as in oxyfuel–gas welding), more braze metal is used than in conventional brazing.

In general, dissimilar metals can be assembled with good joint strength. Examples of joints made are shown in Fig. 32.3. Intricate, lightweight shapes can be joined rapidly and with little distortion.

Filler Metals. Several filler metals (*braze metals*) are available with a range of brazing temperatures (Table 32.1). Note that, unlike those for other welding operations, filler metals for brazing generally have a composition that is significantly different from those of the metals to be joined. They are available in a variety of shapes, such as wire, rod, ring, shim stock, and filings. The selection of the type of filler metal and its composition are important in order 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* at the joint (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 joint can change as a result of subsequent processing or during the service life of a brazed part. For example, when titanium is brazed with pure tin as the filler metal, it is possible for the tin to diffuse completely into the

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	700–925
Ferrous and nonferrous (except aluminum and magnesium)	Silver and copper alloys, copper–phosphorus	620–1150
Iron-, nickel-, and cobalt-based alloys	Gold	900–1100
Stainless steels, nickel- and cobalt-based alloys	Nickel–silver	925–1200

titanium base metal when it is subjected to subsequent aging or to heat treatment. Consequently, the joint no longer exists.

Fluxes. The use of a *flux* is essential in brazing; a flux 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 the capillary action.

It is essential that the surfaces to be brazed be clean and free from rust, oil, and other contaminants in order (a) for proper wetting and spreading of the molten filler metal in the joint and (b) to develop maximum bond strength. Sand blasting also may be used to improve the surface finish of the faying surfaces for brazing. Because they are corrosive, fluxes must be removed after brazing, usually by washing with hot water.

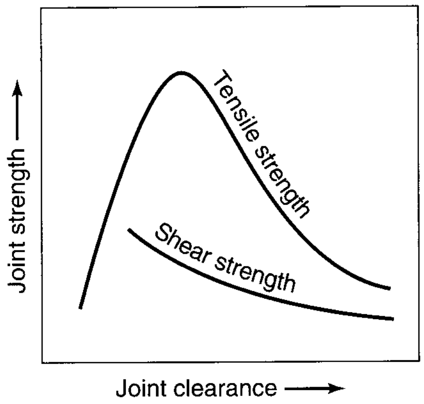


FIGURE 32.4 The effect of joint clearance on the tensile and shear strength of brazed joints. Note that, unlike tensile strength, shear strength continually decreases as the clearance increases.

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. Joint clearances typically range from 0.025 to 0.2 mm. As shown in Fig. 32.4, the smaller the gap, the higher is the *shear strength* of the joint. The shear strength of brazed joints can reach 800 MPa by using brazing alloys containing silver (*silver solder*). Note that there is an optimum gap for achieving maximum *tensile strength* of the joint.

Because clearances are very small, roughness of the mating surfaces becomes important. The surfaces to be brazed must be cleaned chemically or mechanically to ensure full capillary action; thus, the use of a flux is also important.

32.2.1 Brazing Methods

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 in the joint. Suitable part thicknesses are typically in the range from 0.25 to 6 mm. Torch brazing is difficult to control and requires skilled labor; however, it can be automated as a production process by using multiple torches. Torch brazing can also be used for repair work.

Furnace Brazing. The parts in *furnace brazing* (FB) are first cleaned and preloaded with brazing metal in appropriate configurations; then the assembly is 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. Skilled labor is not required, and complex shapes can be brazed because the whole assembly is heated uniformly in the furnace.

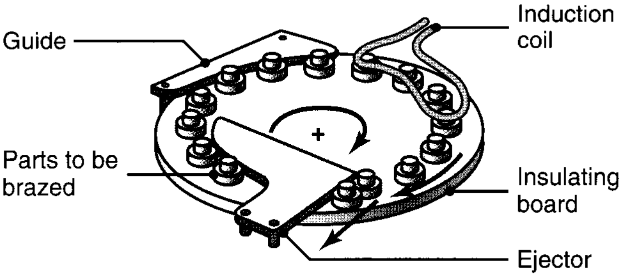


FIGURE 32.5 Schematic illustration of a continuous induction-brazing setup for increased productivity.

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. 4.26). Unless a protective (neutral) atmosphere is utilized, fluxes generally are used. Part thicknesses usually are less than 3 mm. Induction brazing is particularly suitable for brazing parts continuously (Fig. 32.5).

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 typically with thicknesses of 0.1 to 12 mm either are preloaded with filler metal or supplied externally with the metal during brazing. As in induction brazing, the process is rapid, heating zones can be confined to very small areas, and the process can be automated to produce reliable and uniform quality.

Dip Brazing. *Dip brazing* (DB) is carried out by dipping the assemblies to be brazed into either a molten filler-metal bath or a molten salt bath (Section 4.12) at a temperature just above the melting point of the filler metal. Thus, all component surfaces are coated with the filler metal. Consequently, 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. Molten salt baths, which also act as fluxes, are used for complex assemblies of various thicknesses. Depending on the size of the parts and the bath size, as many as 1000 joints can be made at one time by dip brazing.

Infrared Brazing. The heat source in *infrared brazing* (IRB) is a high-intensity quartz lamp. The process is particularly suitable for brazing very thin components, usually less than 1 mm thick, including honeycomb structures (Section 16.12). The radiant energy is focused on the joint, and brazing can be carried out in a vacuum. *Microwave heating* also can be used.

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 minutes to 24 hours. This process is used for strong lap or butt joints and for difficult joining operations. 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 (see also Sections 27.6 and 27.7).

Braze Welding. The joint in *braze welding* is prepared as it is in fusion welding, described in Chapter 30. While an oxyacetylene torch with an oxidizing flame is used, filler metal is deposited at the joint (hence the term *welding*) rather than drawn in by capillary action. As a result, considerably more filler metal is used than in brazing. However, temperatures in braze welding generally are lower than in fusion welding, and thus part distortion is minimal. The use of a flux is essential in this process. The principal use of braze welding is for maintenance and repair work, such as work 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 design 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 work-holding devices may be required to hold the parts together during brazing; some will allow for thermal expansion and contraction during the brazing operation.

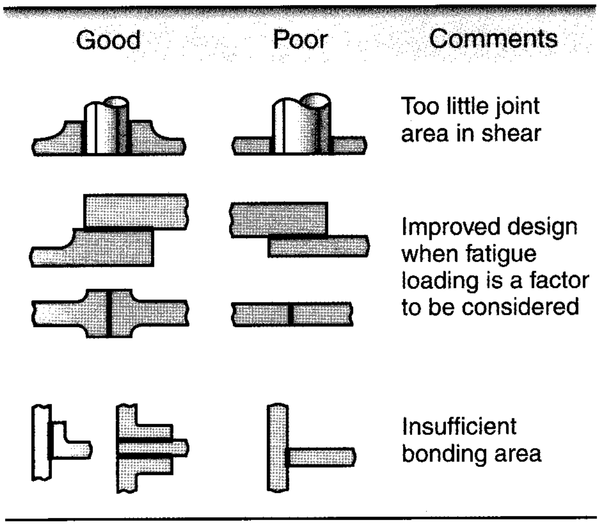


FIGURE 32.6 Examples of good and poor design for brazing. *Source*: American Welding Society.

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. Two important characteristics of solders are low surface tension and high wetting capability. Heat sources for soldering are usually 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.

32.3.1 Types of Solders and Fluxes

Solders melt at a temperature that is the *eutectic point* of the solder alloy (see, for example, Fig. 4.7). Solders traditionally have been tin–lead alloys in various proportions. For example, a solder of 61.9% Sn–38.1% Pb composition 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 solder compositions are tin–zinc, lead–silver, cadmium–silver, and zinc–aluminum alloys (Table 32.2).

Because of the *toxicity* of lead and its adverse effects on the environment, **lead-free solders** are being developed continuously and are coming into wider use. Among the various candidate materials are silver, indium, and bismuth eutectic alloys in

TABLE 32.2

Types of Solders and Their 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

combination with tin. Three typical compositions are 96.5% Sn–3.5% Ag, 42% Sn–58% Bi, and 48% Sn–52% In. However, none of these combinations are suitable for every soldering application.

Fluxes are used in soldering and for the same purposes as they are in welding and brazing, as described in Section 32.2. 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.

32.3.2 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 metallic elements to induce interfacial bonding. This method is similar to that used for joining carbides and ceramics (see Section 32.6.3). A common example of the method is *tinplate*, which is steel sheet coated with tin, thus making it very easy to solder. Tinplate is a common material used in making cans for food.

32.3.3 Soldering Techniques

The following soldering techniques are somewhat similar to brazing methods:

a. **Torch soldering** (TS).
b. **Furnace soldering** (FS).
c. **Iron soldering** (INS) (with the use of a soldering iron).
d. **Induction soldering** (IS).
e. **Resistance soldering** (RS).
f. **Dip soldering** (DS).
g. **Infrared soldering** (IRS).

Other soldering techniques, for special applications, are:

h. **Ultrasonic soldering** (in which a transducer subjects the molten solder to ultrasonic cavitation. This action removes the oxide films from the surfaces to be joined and thus eliminates the need for a flux—hence the term *fluxless soldering*).
i. **Reflow (paste) soldering** (RS).
j. **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 greater detail.

Reflow Soldering. *Solder pastes* are solder–metal particles held together by flux, binding, and wetting agents. The pastes are semisolid in consistency, have high viscosity, and thus are capable of maintaining their shape for relatively long periods.

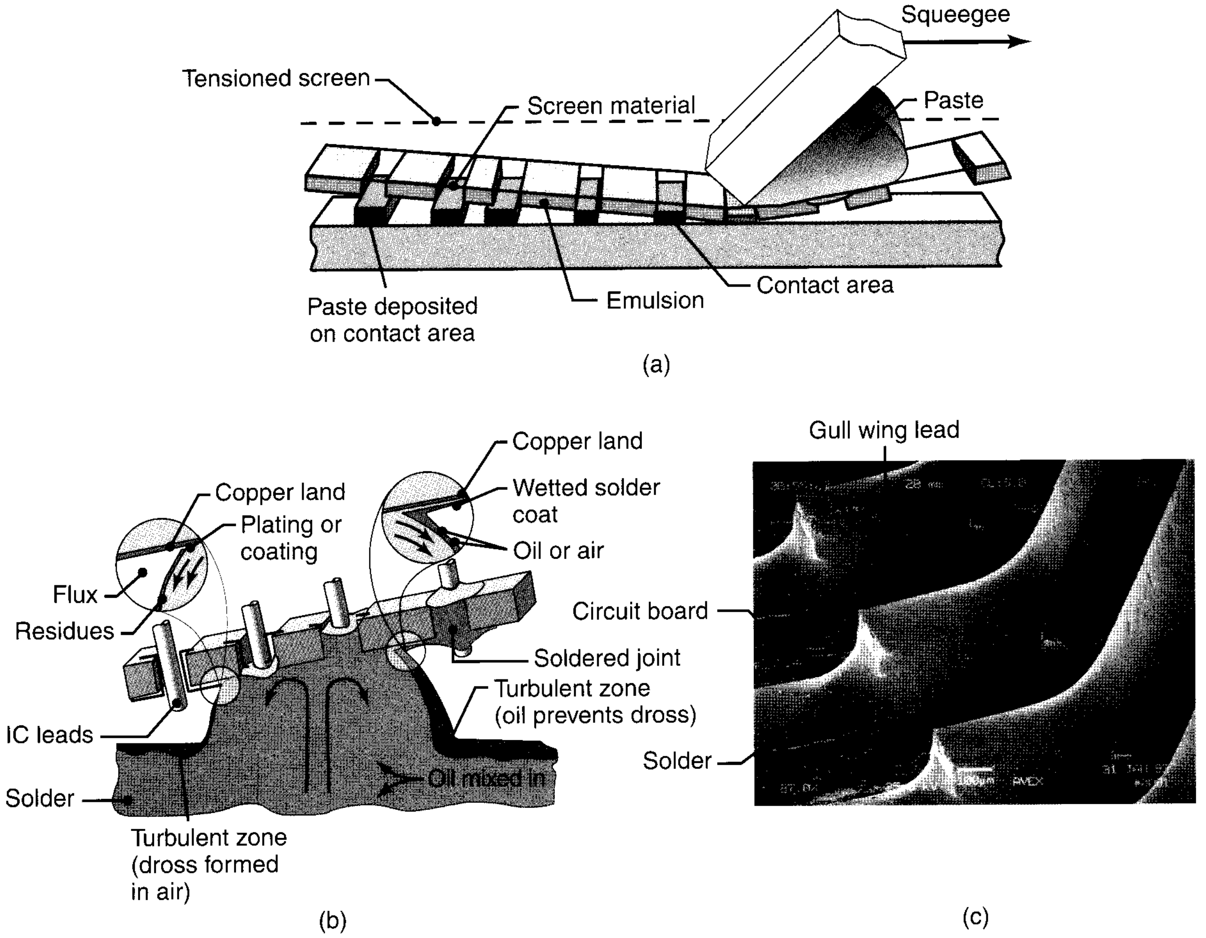


FIGURE 32.7 (a) Screening solder paste onto a printed circuit board in reflow soldering. (b) Schematic illustration of the wave-soldering process. (c) SEM image of wave-soldered joint on surface-mount device. *Source*: (a) After V. Solberg.

The paste is placed directly onto the joint, or on flat objects for finer detail, and it can be applied via a *screening* or *stenciling* process, as shown in Fig. 32.7a. Stenciling is commonly used during the attachment of electrical components to printed circuit boards. An additional benefit of reflow soldering is that the surface tension of the paste helps keep surface-mount packages aligned on their pads; this feature improves the reliability of the solder joints. (See also Section 28.11.)

Once the paste has been placed and the joint assembled, it is heated in a furnace and soldering takes place. In reflow soldering, the product is heated in a controlled manner, so that the following sequence of events occurs:

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 carefully.
4. The solder particles are melted, and they wet the joint.
5. The assembly is cooled at a low rate to prevent thermal shock and fracture of the solder joint.

Although this process appears to be straightforward, there are several process variables for each stage, and good control over temperatures and exposures must be maintained at each stage in order to ensure proper joint strength.

Wave Soldering. *Wave soldering* is a common technique for attaching circuit components to their boards (see Section 28.11). To understand the principle of wave soldering, it is important to note that molten solder does not wet all surfaces. The solder will not stick to most polymer surfaces, and it is easy to remove while molten. Also, as can be observed with a simple handheld soldering iron, the solder wets metal surfaces and forms a good bond only when the metal is preheated to a certain temperature. Thus, wave soldering requires separate fluxing and preheating operations before it can be completed.

A typical wave-soldering operation is illustrated in Fig. 32.7b. A standing laminar wave of molten solder is generated by a pump. Preheated and prefluxed circuit boards are then conveyed over the wave. The solder wets the exposed metal surfaces, but it does not remain attached to the polymer package for the integrated circuits, and it does not adhere to the polymer-coated circuit boards. An *air knife* (basically a high-velocity jet of hot air) 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 commence. Bonding usually is accomplished by (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) performing 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

The 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, during which they may be subjected to significant temperature variations and to vibration. In recognition of this requirement, it is essential that the solder joints used to attach such devices to circuit boards be sufficiently strong and reliable and also that the solder joints be applied extremely rapidly with 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. Further space savings are achieved by mounting integrated circuits into surface-mount packages, which allow 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 desired to solder all of the joints via a reliable automated process. It is important to recognize that, for efficiency of assembly, all of the in-line circuits should be restricted to insertion from one side of the board. Indeed, there is no performance requirement that would dictate otherwise, and this restriction greatly simplifies manufacturing.

The basic steps in soldering the connections on such a board are as follows (see Figs. 32.7b and c):

1. Apply solder paste to one side.
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 the adhesive, attach the surface-mount devices onto the secondary side.
6. Cure the adhesive.
7. Perform a wave-soldering operation on the secondary 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 and produce a more uniform paste thickness.) Surface-mount circuit components are then placed on the board, and the board is heated in a furnace to around 200°C to reflow the solder and form strong connections between the surface mount and the circuit board.

At this point, 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 then cured, the board is flipped, and wave soldering is done.

The wave-soldering operation simultaneously joins the surface-mount components to the secondary side and solders the leads of the in-line components from the board's primary side. The board is then cleaned and inspected prior to the performance of electronic quality checks.

32.3.4 Soldering Applications and Design Guidelines

Soldering is used extensively in the electronics industry. Note, however, that because soldering temperatures are relatively low, a soldered joint has very limited utility at elevated temperatures. Moreover, since solders generally do not have much strength, the process cannot be used for load-bearing (structural) members. Joint strength can be improved significantly by *mechanical interlocking* of the joint, as illustrated in Fig. 32.8.

Soldering can be used to join various metals of different thicknesses. Copper and precious metals such as silver and gold are easy to solder. Aluminum and stainless steels are difficult to solder because of their strong, thin oxide film. However, these and other metals can be soldered with the aid of special fluxes that modify surfaces. Although manual operations require skill and are time consuming, soldering speeds can be high with the use of automated equipment.

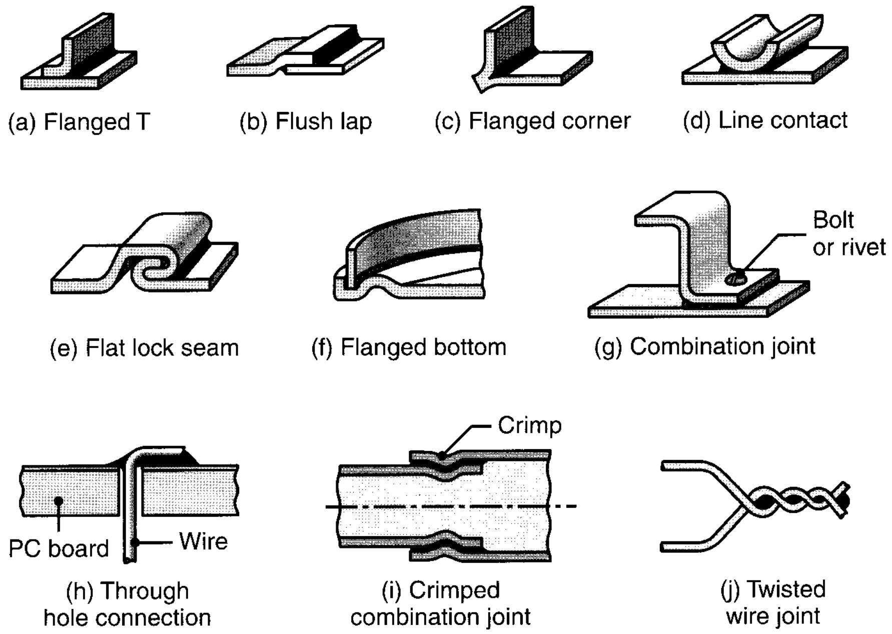


FIGURE 32.8 Joint designs commonly used for soldering.

Design guidelines for soldering are similar to those for brazing. 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

Numerous parts and components can be joined and assembled by **adhesives** rather than by one or more of the joining methods described thus far. 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 wood layers 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 various forms: liquid, paste, solution, emulsion, powder, tape, and film. When applied, adhesives typically are about 0.1 mm thick.

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.

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	Good	Poor	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.

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). The three basic types of adhesives are the following:

a. **Natural adhesives**—such as starch, dextrin (a gummy substance obtained from starch), soya flour, and animal products.
b. **Inorganic adhesives**—such as sodium silicate and magnesium oxychloride.
c. **Synthetic organic adhesives**—which may be thermoplastics (used for nonstructural and some structural bonding) or thermosetting polymers (used primarily for structural bonding).

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	"Crazy glue™"
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; rubber-like	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

Because of their strength, synthetic organic adhesives are the most important adhesives in manufacturing processes, particularly for load-bearing applications. They are classified as follows:

- **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).
- **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 may be classified on the basis of their specific chemistries:

- **Epoxy-based systems**: These systems have high strength and high-temperature properties to as high as 200°C. Typical applications include automotive brake linings and bonding agents for sand molds for casting.
- **Acrylics**: These adhesives are suitable for applications with substrates that are not clean.
- **Anaerobic systems**: The 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 40s.
- **Urethanes**: These adhesives have high toughness and flexibility at room temperature, and they 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, which are followed in affordability by polyurethanes, acrylics, silicones, and cyanoacrylates. Adhesives for high-temperature applications in a range up to about 260°C (such as polyimides and polybenzimidazoles) are generally the most expensive. Most adhesives have an optimum temperature (ranging from about room temperature to about 200°C) for maximum shear strength.

32.4.2 Electrically Conducting Adhesives

Although the majority of adhesive bonding applications require mechanical strength, a relatively recent advance is the development and application of electrically conducting adhesives to 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.

In electrically conducting adhesives, the polymer is the matrix and contains conducting metals (fillers) in forms such as flakes and particles (see also Section 7.3 on *electrically conducting polymers*). There is a minimum proportion of fillers necessary to make the adhesive electrically conducting; typically, it is 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 conducting particle contact geometry can be controlled to impart isotropic or anisotropic electrical conductivity to the adhesive. Metals used are typically silver, nickel, copper, and gold, as well as carbon. More recent developments include polymeric particles (such as polystyrene) coated with thin metallic films of silver or gold. Matrix materials are generally epoxies, although thermoplastics also are used and are available as film or as paste. Applications of electrically conducting adhesives include calculators, remote controls, and control panels. In addition, there are high-density uses in electronic assemblies, liquid-crystal displays, pocket TVs, and electronic games.

32.4.3 Surface Preparation, Process Capabilities, and Applications

Surface preparation is very important in adhesive bonding. Joint strength depends greatly on the absence of dirt, dust, oil, and various other contaminants. This dependence can be observed when one is attempting to apply an adhesive tape over a dusty or oily surface. Contaminants also affect the wetting ability of the adhesive and prevent even spreading of the adhesive over the interface. Thick, weak, or loose oxide films on workpiece 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 to improve adhesion or to introduce mechanical locking. However, the roughness must not be too high, because air may be trapped, in which case the joint strength is reduced. Various compounds and primers are available that modify surfaces to improve adhesive-bond strength. Liquid adhesives may be applied by brushes, sprayers, or rollers.

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 also be combined with mechanical joining methods (Section 32.5) to further improve the strength of the bond. Joint design and bonding methods require care and skill. Special equipment is usually required, such as fixtures, presses, tooling, and autoclaves and ovens for curing.

Adhesive joints are designed to withstand shear, compressive, and tensile forces, but they should not be subjected to *peeling* (Fig. 32.9). Note, for example, how easily you can peel adhesive tape from a surface, yet be unable to slide it along

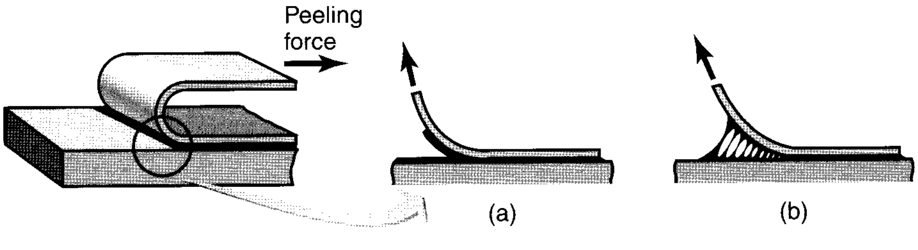


FIGURE 32.9 Characteristic behavior of (a) brittle and (b) tough adhesives in a peeling test. This test is similar to the peeling of adhesive tape from a solid surface.

the surface. During peeling, the behavior of an adhesive may be either brittle or ductile and tough—requiring high forces to peel the adhesive.

Applications. Major industries that use adhesive bonding extensively are the aerospace, automotive, appliances, and building products industries. Applications include automotive brake-lining assemblies, laminated windshield glass, appliances, helicopter blades, honeycomb structures, and aircraft bodies and control surfaces.

An important consideration in the use of adhesives in production is *curing time*, which can range from a few seconds (at high temperatures) to several hours (at room temperature), particularly for thermosetting adhesives. Thus, production rates can be low compared with those of other joining processes. Furthermore, adhesive bonds for structural applications rarely are suitable for service above 250°C.

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 nondestructive-testing methods for adhesives.

The major advantages of adhesive bonding are as follows:

- The interfacial bond has sufficient strength for structural applications, but is also used for nonstructural purposes, such as sealing, insulation, the prevention of electrochemical corrosion between dissimilar metals, and the reduction of vibration and of noise (by means of internal damping at the joints).
- Adhesive bonding distributes the load at an interface and thereby eliminates localized stresses that usually result from joining the components with mechanical fasteners, such as bolts and screws. Moreover, structural integrity of the sections is maintained (because no holes are required).
- The external appearance of the bonded components is unaffected.
- Very thin and fragile components can be bonded without significant increase in their weight.
- Porous materials and materials of very different properties and sizes can be joined.
- Because adhesive bonding usually is carried out at a temperature between room temperature and about 200°C, there is no significant distortion of the components or change in their original properties. Avoiding distortion is important, particularly for materials that are heat sensitive.

The major limitations of adhesive bonding are the following:

- There is a limited range of service temperatures.
- Bonding time can be long.
- There is a need for great care in surface preparation.
- Bonded joints are difficult to test nondestructively, particularly for large structures.
- The limited reliability of adhesively bonded structures during their service life and under hostile environmental conditions (such as *degradation* by temperature, oxidation, stress corrosion, radiation, or dissolution) may be a significant concern.

The cost of adhesive bonding depends on the particular operation. In many cases, however, the overall economics of the process make adhesive bonding an attractive joining process. Sometimes it is the only one that is feasible or practical. The cost of equipment varies greatly, depending on the size and type of operation.

32.4.4 Design for Adhesive Bonding

- Designs for adhesive bonding should ensure that joints are subjected only to compressive, tensile, and shear forces and not to peeling or cleavage.

- Several joint designs for adhesive bonding are shown in Figs. 32.10 through 32.12. They vary considerably in strength; hence, selection of the appropriate design is important and should include considerations such as the type of loading and the environment.

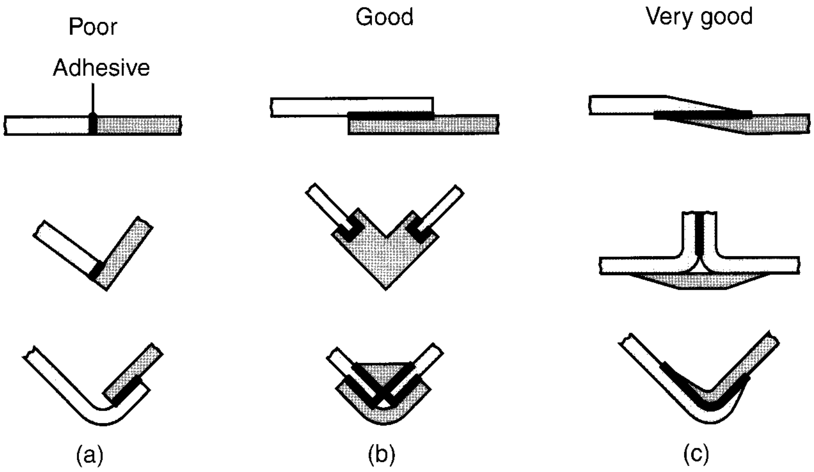


FIGURE 32.10 Various joint designs in adhesive bonding. Note that good designs require large contact areas between the members to be joined.

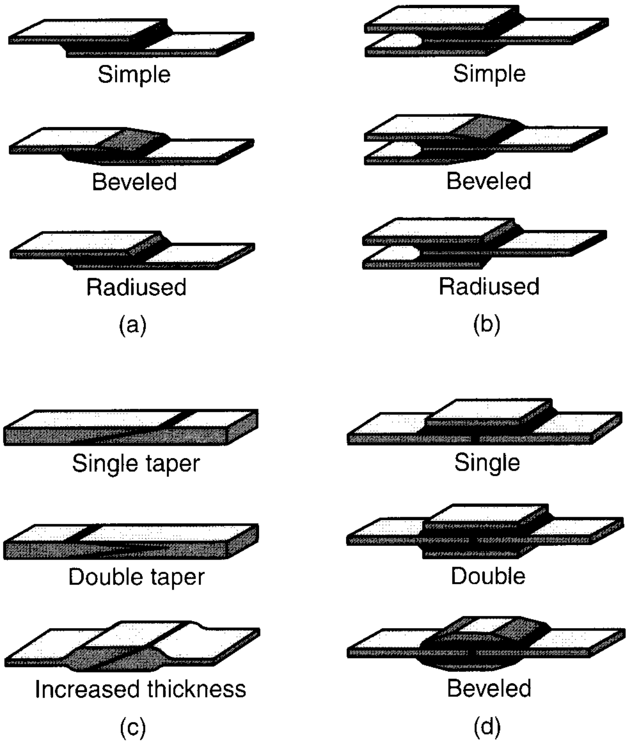


FIGURE 32.11 Desirable configurations for adhesively bonded joints: (a) single lap, (b) double lap, (c) scarf, and (d) strap.

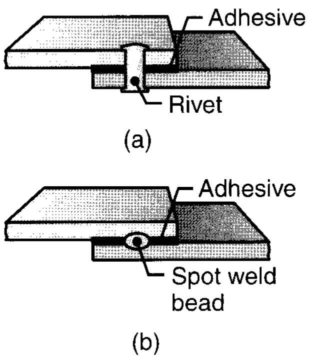


FIGURE 32.12 Two examples of combination joints, for purposes of improved strength, air tightness, and resistance to crevice corrosion.

- Butt joints require large bonding surfaces. Simple lap joints tend to distort under tension because of the force couple at the joint. (See Fig. 31.9a.)
- The coefficients of thermal expansion of the components to be bonded should preferably be close to each other in order to avoid internal stresses during adhesive bonding. Also, situations in which thermal cycling can cause differential movement across the joint should be avoided.

CASE STUDY 32.1 Light Curing Acrylic Adhesives for Medical Products

Cobe Cardiovascular, Inc., is a leading manufacturer of blood collection and processing systems, as well as extracorporeal 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 the use of solvents, and Cobe particularly wanted to eliminate its use of methylene chloride for environmental and occupational safety reasons. Towards this goal, the company began to redesign most of its assemblies to accept light-curing (ultraviolet or visible) adhesives. Most of the company's devices were made of transparent plastics. Consequently, its engineers required 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 (Fig. 32.13). The reservoir is a one-time use, disposable device; its purpose is to collect and hold the blood during open-heart and 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 reser-

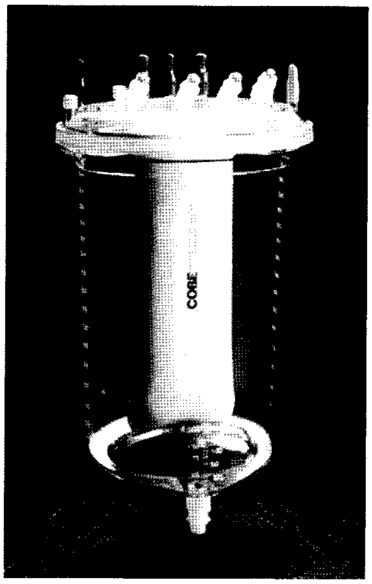


FIGURE 32.13 The Cobe Laboratories blood reservoir. The lid is bonded to the bowl with an airtight adhesive joint and tongue-in-groove joint. *Source*: Courtesy of Cobe Laboratories.

voir consists of a clear, polycarbonate lid joined to a polycarbonate bucket. The joint is a tongue-and-groove configuration; the goal was to create a strong, elastic joint that could withstand repeated stresses with no chance of leakage.

Light-cured acrylic adhesives offer a range of performance properties that make them well suited for this application. First and foremost, they achieve high bond strength to the thermoplastics typically used to form 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 the high bond strength after sterilization. Fortunately, disposable medical devices are typically subjected to very few sterilization cycles during manufacturing. Also, these adhesives can endure a limited number of cycles of gamma irradiation, electron beam irradiation, autoclaving, ethylene oxide, or chemical immersion.

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; for example, Loctite® 3211 yields at elongations in excess of 200%. This flexibility is critical, because the bonded joints are typically subjected to large amounts of bending and flexing when the devices are pressurized during qualification testing and use. If an adhesive is too rigid, it will fail this type of testing, even if it offers higher shear strength than a comparable and more flexible adhesive. Finally, light-cured acrylics are widely available in formulations that meet international quality standard certification (ISO; see Section 36.6), which means that, when processed properly, they will not cause biocompatibility problems in the final assembly.

While these performance features are attractive, the adhesive also must meet certain processing characteristics during manufacturing. Light-cured acrylic adhesives have found wide use in medical-device assembly/joining operations, because their processing characteristics are compatible with the high-speed automated manufacturing processes employed. These adhesives are available in a wide variety of viscosities and are dispensed easily through either pressure–time or positive-displacement dispensing systems. Once dispensed on the part, they can remain in contact with even highly stressed plastic parts for several minutes or longer without causing stress cracking or degradation of the plastic. For example, Liquid Loctite® 3211 can remain in contact with polycarbonate that has been bent to induce stresses up to 17 MPa for more than 15 minutes without stress cracking. Finally, the adhesive can be converted completely from a liquid to a solid state in seconds when exposed to light of the proper intensity and wavelength.

Since Loctite® 3211 absorbs light in the visible as well as the ultraviolet range, it can be used successfully on plastics that contain UV blockers, such as many grades of polycarbonate. The ability to have a long open time when parts can be positioned, yet cure the adhesive on demand, is a unique benefit to light-curing adhesives, dramatically reducing scrap costs. The equipment used to irradiate the part with high-intensity light typically requires a space of $1 \times 2\ \text{m}^2$ on a production line, which generally is much less than that required for the ovens used by heat-cured adhesives or the racking shelves required for slower curing adhesives. Since floor space carries a cost premium in clean-room environments, this is a significant benefit.

It also is important that the joint be designed properly to maximize performance. If the enclosure is bonded with a joint consisting of two flat faces in intimate contact, peel stresses (see Fig. 32.9) will be acting on the bond when the vessel is pressurized. Peel stresses are the most difficult type for an adhesive joint to withstand, due to the fact that the entire load is concentrated on the leading edge of the joint. The tongue-and-groove design that the company used addressed this concern, with the groove acting as a reservoir for holding 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 and tensile 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 quickly can 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 with the use of this adhesive, the environmental concerns and the issues associated with solvent bonding were eliminated, with the accompanying benefit 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. It also provided the structural strength needed and thus maintained a competitive edge for the company in the marketplace.

Source: Courtesy of P.J. Courtney, Loctite Corporation.

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 sometime during the product's service life or life cycle. Numerous products, such as mechanical pencils, watches, computers, appliances, engines, and bicycles, have components that are fastened mechanically. *Mechanical fastening* may be preferred over other 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, and adjustable components and fixtures.
- Lower overall cost of manufacturing the product.

The most common method of mechanical fastening is by the use of bolts, nuts, screws, pins, and a variety of other **fasteners**. Also known as **mechanical assembly**, mechanical fastening generally requires that the components have *holes* through which the fasteners are inserted. These joints may be subjected to both shear and tensile stresses and should 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 in a solid body can be produced by several processes, 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 the product during casting, forging, extrusion, and powder metallurgy. For improved accuracy and surface finish, many of these hole-making operations may be followed by finishing operations, such as shaving, deburring, reaming, and honing, as described in various sections of Part IV.

Because of the fundamental differences in their characteristics, each of the hole-making processes produces holes with different surface finishes, surface properties, and dimensional accuracy and characteristics. The most significant influence of a hole in a solid body is its tendency to reduce the component's fatigue life by stress concentration. For holes, fatigue life can be improved best by inducing compressive residual stresses on the cylindrical surface of the hole. These stresses usually are developed by pushing a round rod (*drift pin*) through the hole and expanding it by a very small amount. This operation plastically deforms the surface layers of the hole in a manner similar to that seen in shot peening or in roller burnishing (Section 34.2).

Threaded Fasteners. Bolts, screws, and nuts are among the most commonly used *threaded fasteners*. Numerous standards and specifications (including thread dimensions, dimensional tolerances, pitch, strength, and the quality of the materials used to make these fasteners) are described in the references at the end of this chapter. Bolts and screws may be secured with nuts, 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 in which fastening does not require a tapped hole or a nut.

If the joint is to be subjected to vibration (such as in aircraft, engines, and machinery), several specially designed nuts and lock washers are available. They increase the frictional resistance in the torsional direction and so inhibit any vibrational loosening of the fasteners.

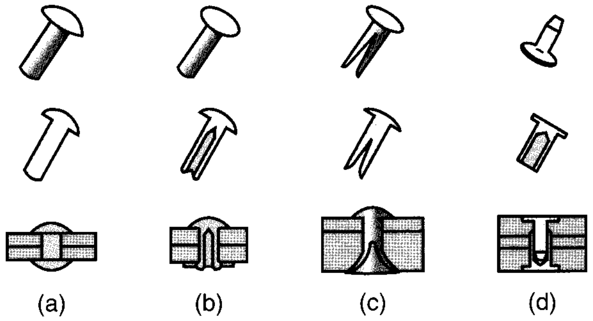


FIGURE 32.14 Examples of rivets: (a) solid, (b) tubular, (c) split or bifurcated, and (d) compression.

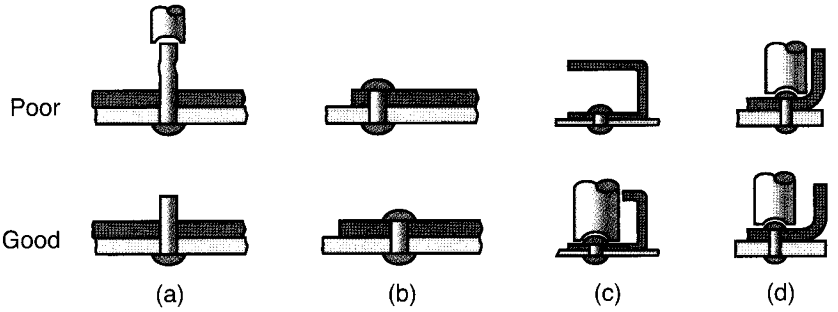


FIGURE 32.15 Design guidelines for riveting. (a) Exposed shank is too long; the result is buckling instead of upsetting. (b) Rivets should be placed sufficiently far from edges to avoid stress concentrations. (c) Joined sections should allow ample clearance for the riveting tools. (d) Section curvature should not interfere with the riveting process. *Source*: After J.G. Bralla.

Rivets. The most common method of permanent or semipermanent mechanical joining is by *riveting* (Fig. 32.14). Hundreds of thousands of rivets may be used in the construction and assembly of one large commercial aircraft. Riveting may be done either at room temperature or at elevated temperatures. Rivets may be solid or tubular. Installing a solid rivet takes two steps: placing the rivet in the hole (usually punched or drilled) and deforming the end of its shank by upsetting it (*heading*; see Fig. 14.11). A hollow rivet is installed by flaring its smaller end (see Section 16.6). Explosives can be placed within the rivet cavity and detonated to expand the end of the rivet. The riveting operation also may be performed by hand or by mechanized means, including the use of programmable robots. Some design guidelines for riveting are illustrated in Fig. 32.15.

32.5.1 Other Fastening Methods

Numerous other techniques are used in joining and assembly applications. The most common types are described here.

Metal Stitching and Stapling. The process of *metal stitching and stapling* (Fig. 32.16) is much like that of the ordinary stapling of papers. 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*, the fastener material must be sufficiently thin and ductile to withstand the large localized deformation during sharp bending.

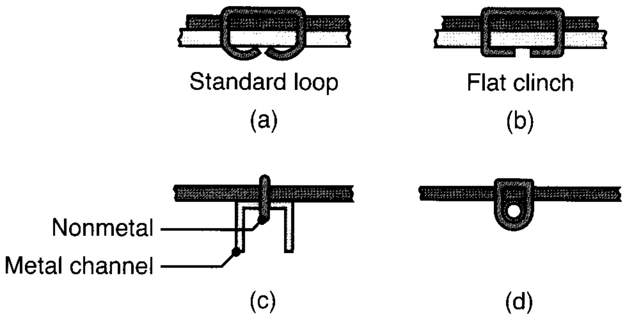


FIGURE 32.16 Typical examples of metal stitching.

Seaming. *Seaming* is based on the simple principle of folding two thin pieces of material together, much like the joining of two pieces of paper by folding them at the corners (Fig. 32.17). Common examples of seaming are seen at the tops of beverage cans (see last illustration in Fig. 16.30), in containers for food and household products, and in heating and air-conditioning ducts. In seaming, the materials should be capable of undergoing bending and folding at very small radii; otherwise, they will crack. The performance and reliability of seams may be improved as well as making them impermeable by the addition of adhesives or polymeric coatings and seals or by soldering.

Crimping. The *crimping* 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. Crimping can be done on both tubular and flat parts, provided that the materials are sufficiently thin and ductile to withstand the large localized deformations. Metal caps are fastened to glass bottles by crimping just as some connectors are crimped to electrical wiring.

Spring and Snap-in Fasteners. Several types of spring and snap-in fasteners are shown in Fig. 32.19. Such fasteners are used widely in automotive bodies and household appliances. They are economical, and they permit easy and rapid component assembly.

Shrink and Press Fits. Components also may be assembled by shrink fitting and press fitting. *Shrink fitting* is based on the thermal contractions of two components. Typical applications are assembling die components and mounting gears and cams onto shafts. In *press fitting*, one component is forced over another; when the components are designed properly, this process results in high joint strength.

Shape-memory Alloys. The characteristics of these materials were described in Section 6.13. Recall their use as fasteners because of their unique capability to recover their shape upon heating. Several advanced applications include their use as coupling in the assembly of titanium-alloy tubing for aircraft.

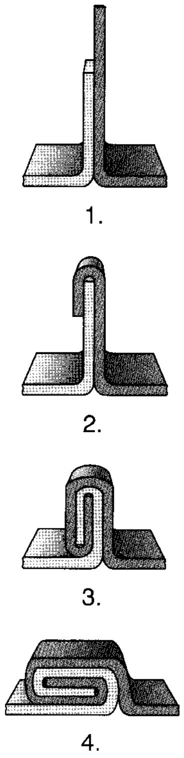


FIGURE 32.17 Stages in forming a double-lock seam.

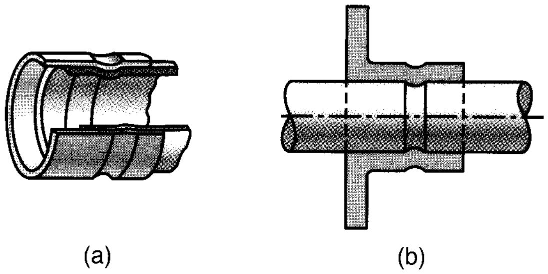

(a) (b)

FIGURE 32.18 Two examples of mechanical joining by crimping.

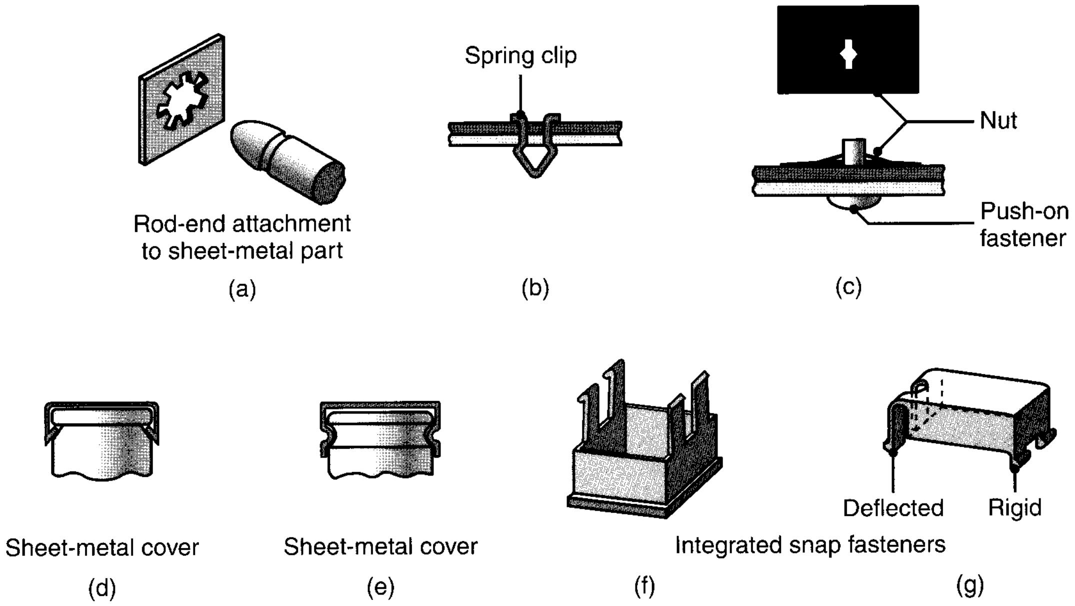


FIGURE 32.19 Examples of spring and snap-in fasteners used to facilitate assembly.

32.5.2 Design for Mechanical Fastening

The design of mechanical joints requires a consideration of the type of loading to which the structure will be subjected and of the size and spacing of holes. *Compatibility* of the fastener material with that of the components to be joined is important. Incompatibility may lead to *galvanic corrosion*, also known as *crevice corrosion* (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; this combination causes rapid corrosion and loss of joint strength. Aluminum or zinc fasteners on copper products will react in a similar manner.

Other general design guidelines for mechanical joining include the following (see also Section 37.10):

- It is generally less costly to use fewer, but larger, fasteners than to use a large number of small ones.
- Part assembly should be accomplished with a minimum number of fasteners.
- The fit between parts to be joined should be as loose as possible to reduce costs and to facilitate the assembly process.
- Fasteners of standard size should be used whenever possible.
- Holes should not be too close to edges or corners, to avoid the possibility of tearing the material when it is subjected to external forces.

32.6 Joining Plastics, Ceramics, and Glasses

Plastics can be joined by many of the methods already 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 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 and ensures a good bond with the application of pressure.

Because of the low thermal conductivity of thermoplastics, the heat source may burn or char the surfaces of the components if applied at too high a rate. Such burning or charring can cause difficulties in obtaining sufficiently deep fusion for proper joint strength. *Oxidation* also can be a problem in joining some polymers (such as polyethylene), because it causes degradation. Typically, an inert shielding gas (such as nitrogen) is used to prevent oxidation.

External heat sources may be chosen from among the following (the choice depends on the compatibility of the polymers to be joined):

- *Hot air, inert gases*, or a filler material of the same type is also used.
- In a process known as *hot-tool welding* or *hot-plate welding*, heated tools and dies are pressed against the surfaces to be joined and heat them by the interdiffusion of molecular chains. This process commonly is used in butt-welded pipes and (end-to-end) tubing.
- *Infrared radiation* (from high-intensity quartz heat lamps) is focused into a narrow beam onto the surface to be joined.
- *Radio waves* are particularly useful for thin films; frequencies are in the range from 100 to 500 Hz.
- *Dielectric heating* at frequencies of up to 100 MHz are effecive for the through heating of polymers such 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—a process known as *resistive-implant welding*. Alternatively, in *induction welding*, these elements at the interface may be subjected to radio-frequency exposure. In both cases, because they are left in the weld zone, the elements at the interface must be compatible with the use of the joined product.
- *Lasers* emitting defocused beams at low power prevent degradation of the polymer.

Internal heat sources are developed by the following means:

- *Ultrasonic welding* is the most commonly used process for thermoplastics, particularly amorphous polymers such as acrylonitrile-butadiene-styrene (ABS) and high-impact polystyrene; frequencies are in the range from 20 to 40 kHz. The ultrasonic welding process illustrated in Fig. 31.2 is still applicable, but note that the tool can apply vertical motion, causing a released-compression loading. Due to the high hysteresis of polymers in a loading cycle, the heat for welding is developed in the polymer and not at the interface.
- *Friction welding* (also called *spin welding* for polymers) and *linear friction welding* (also called *vibration welding*) are particularly useful for joining polymers with a high degree of crystallinity, such as acetal, polyethylene, nylons, and polypropylene.
- *Orbital welding* is similar to friction welding, except 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 by means of adhesives. Plastics (such as PVC, polyethylene, polypropylene, acrylics, and ABS) can be joined in this manner. For example, specially designed portable fusion-sealing systems are used to allow in-field joining of plastic pipe (usually made of polyethylene and used for natural-gas delivery).

Coextruded multiple food wrappings consist of different types of films, which are 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 during 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 the 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 used to apply the adhesive to the connecting sleeve and pipe surfaces (a step much like that using primers in painting), and then the pieces are pushed together.

Adhesive bonding of polyethylene, polypropylene, and polytetrafluoroethylene (Teflon) can be difficult, because adhesives do not bond readily to them. The surfaces of parts made of these materials usually have to be treated chemically to improve bonding. The use of 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. The use of self-tapping metal screws is a common practice. *Integrated snap fasteners* have gained wide acceptance for simplifying assembly operations; fastener geometries are shown in Figs. 32.19f and g. Because the fastener can be molded directly at the same time as the plastic, it adds very little to the cost of the assembly. This technique is very cost effective, because it reduces assembly time and minimizes the number of parts required.

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 together and holding them together until sufficient joint strength is developed.

Electromagnetic Bonding. Thermoplastics also may be joined by magnetic means by embedding tiny particles on the order of 1 μm in the polymer. A high-frequency field then causes induction heating of the polymer and melts 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 other molded-in *inserts.*
- *Mechanical fasteners*, particularly those using 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.*

32.6.3 Joining Ceramics and Glasses

A wide variety and numerous types of ceramics and glasses are now available with unique and important properties. Ceramics and glasses are used as products, as components of products, or as tools, molds, and dies. These materials 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 materials can be joined by adhesive bonding. A typical example is assembling broken ceramic pieces with a two-component epoxy, which is dispensed from two separate tubes and is mixed just prior to 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 when it comes to stiffness, hardness, brittleness, resistance to high temperatures, and chemical inertness. Thus, joining them to each other or to other metallic or nonmetallic materials requires special considerations, and several highly specialized joining processes have been developed.

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*, as described in Section 34.5. In this technique, known as the *Mo–Mn process*, first the ceramic part is coated with a slurry of oxides of molybdenum and manganese. Next, the part is fired, forming a glassy layer on its surface. Then this layer is plated with nickel, and since the part now has a metallic surface, it can be brazed to a metal surface by means of an appropriate filler metal.

Tungsten carbide and titanium carbide can be brazed easily 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. Common applications include brazing cubic boron nitride or diamond tips to carbide inserts (Figs. 22.10 and 22.11) and carbide tips to masonry drills (Fig. 23.21). Depending on their particular structure, ceramics and metals also can be joined by *diffusion bonding*. It may be necessary to place a metallic layer at the joint to make it stronger.

In addition, ceramic components can be joined or assembled together during their primary shaping process; a common example is attaching handles to coffee mugs prior to firing them. Thus, the 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 be bonded easily to each other. This is commonly done by first heating and softening the surface to be joined, then pressing the two pieces together, and finally cooling them. Bonding glass to metals is also possible, because of the diffusion of metal ions into the amorphous surface structure of 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 (described in 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 is as follows:

- *Highest:* brazing, bolts, nuts, and other fasteners.
- *Intermediate:* riveting and adhesive bonding.
- *Lowest:* seaming and crimping.

The variety of processes involved and the general costs are as follows:

Brazing

- *Manual brazing:* The basic equipment costs about \$300, but it can run over \$50,000 for automated systems.
- *Furnace brazing:* Costs vary widely, ranging from about \$2000 for simple batch furnaces to \$300,000+ 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:* The cost of equipment varies widely, from \$2000 to more than \$200,000; the more expensive equipment includes various computer-control features.
- *Infrared brazing:* Equipment cost ranges from \$500 to \$30,000.
- *Diffusion brazing:* The cost of equipment ranges from \$50,000 to \$300,000.

Soldering. The cost of soldering equipment depends on its complexity and on the level of automation. The cost ranges 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 the interfaces include brazing and soldering. These processes instead utilize filler material that requires some temperature rise in the joint. They can be used to join dissimilar metals of intricate shapes and various thicknesses.
- Adhesive bonding has gained increased acceptance in major industries, such as the aerospace and automotive industries. In addition to possessing good bond strength, adhesives have other favorable characteristics, such as the ability to seal, to insulate, to prevent electrochemical corrosion between dissimilar metals, and to reduce vibration and noise by means of internal damping in the bond.
- Surface preparation and joint design are important factors in adhesive bonding.
- Mechanical fastening is one of the oldest and most common joining methods. Bolts, screws, and nuts are common fasteners for machine components and structures that are likely to be taken apart for maintenance, for ease of transportation, or for various other reasons.
- Rivets are semipermanent or permanent fasteners used in buildings, bridges, and transportation equipment. A wide variety of other fasteners and fastening techniques is available for numerous permanent or semipermanent applications.
- Thermoplastics can be joined by fusion-welding techniques, by adhesive bonding, or by mechanical fastening. Thermosets usually are 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 interface and 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.

Adams, R.D., Comyn, J., and Wake, W.C., **Structural Adhesive Joints in Engineering**, Chapman & Hall, 1997.

Bath, J. (ed.), **Lead-Free Soldering**, Springer 2007.

Bickford, J.H., **Introduction to the Design and Behavior of Bolted Joints**, 4th ed., Marcel Dekker, 2007.

Brockmann, W., Geiss, P.L., Klingen, J., and Schreoeder, K.B., **Adhesive Bonding: Materials, Applications and Technology**, Wiley, 2009.

Hamrock, B.J., Jacobson, B., and Schmid, S.R., **Fundamentals of Machine Elements**, 2nd ed., McGraw-Hill, 2005.

Handbook of Plastics Joining: A Practical Guide, William Andrew, Inc., 1996.

Humpston, G., and Jacobson, D.M., **Principles of Soldering**, ASM International, 2004.

Hwang, J.S., **Modern Solder Technology for Competitive Electronics Manufacturing**, McGraw-Hill, 1996.

Jacobson, D.M., and Humpston, G., **Principles of Brazing**, ASM International, 2005.

Manko, H.H., **Soldering Handbook for Printed Circuits and Surface Mounting**, Van Nostrand Reinhold, 1995.

———, **Solders and Soldering–Materials, Design, Production, and Analysis for Reliable Bonding**, McGraw-Hill, 1992.

Nicholas, M.G., **Joining Processes: Introduction to Brazing and Diffusion Bonding**, Chapman & Hall, 1998.

Parmley, R.O. (ed.), **Standard Handbook of Fastening and Joining**, 3rd ed., McGraw-Hill, 1997.

Peaslee, R.L., **Brazing Footprints: Case Studies in High-Temperature Brazing**, Wall Colmonoy Corp., 2003.

Petrie, E.M., **Handbook of Adhesives and Sealants**, McGraw-Hill, 2nd ed., 2006.

Roberts, P., **Industrial Brazing Practice**, CRC Press, 2004.

Rotheiser, J., **Joining of Plastics: Handbook for Designers and Engineers**, Hanser Gardner, 2004.

Satas, D. (ed.), **Handbook of Pressure Sensitive Adhesive Technology**, 3rd ed., Satas & Associates, 1999.

Schwartz, M.M., **Brazing**, 2nd ed., ASM International, 2003.

———, **Joining of Composite-Matrix Materials**, ASM International, 1994.

Speck, J.A., **Mechanical Fastening, Joining, and Assembly**, Marcel Dekker, 1997.

Woodgate, R.W., **Handbook of Machine Soldering**, Wiley, 1996.

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. Soldering is generally applied to thinner components. Explain why.

32.5. Explain the reasons that a variety of mechanical joining methods have been developed over the years.

32.6. Describe the similarities and differences between the functions of a bolt and those of a rivet.

32.7. What precautions should be taken in the mechanical joining of dissimilar metals?

32.8. What difficulties are involved in joining plastics? Why?

32.9. What are the principles of (a) wave soldering and (b) reflow soldering?

32.10. What is a peel test? Why is it useful?

QUALITATIVE PROBLEMS

32.11. Describe some applications in manufacturing for single-sided and some for double-sided adhesive tapes.

32.12. Comment on your observations concerning the joints shown in Figs. 32.3, 32.6, 32.10, and 32.11.

32.13. Give examples of combination joints other than those shown in Fig. 32.12.

32.14. Discuss the need for fixtures for holding workpieces in the joining processes described in this chapter.

32.15. Explain why adhesively bonded joints tend to be weak in peeling.

32.16. It is common practice to tin-plate electrical terminals to facilitate soldering. Why is it tin that is used?

32.17. How important is a close fit for two parts that are to be brazed? Explain.

32.18. 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.

QUANTITATIVE PROBLEMS

▶**32.19.** Refer to the simple butt and lap joints shown at the top row of Fig. 32.10a. (a) Assuming that the area of the butt joint is 5 × 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 that its area is 15 × 15 mm.

▶**32.20.** In Fig. 32.12a, assume that the cross section of the lap joint is 20 × 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.

32.21. 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.

32.22. 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.

SYNTHESIS, DESIGN, AND PROJECTS

32.23. Examine various household products and describe how their components are joined and assembled. Explain why those particular processes were used and not others.

32.24. Name several products that have been assembled by (a) seaming, (b) stitching, and (c) soldering.

32.25. Suggest methods of attaching a round bar (made of a thermosetting plastic) perpendicularly to a flat metal plate. Discuss their advantages and limitations.

32.26. 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.

32.27. Prepare a list of design guidelines for joining by the processes described in this chapter. Would these guidelines be common to most processes? Explain.

32.28. 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.

32.29. Answer Problem 32.28, 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.

32.30. Comment on workpiece size and shape limitations, if any, for each of the processes described in this chapter.

32.31. Describe part shapes that cannot be joined by the processes covered in this chapter. Give specific examples.

32.32. 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.

32.33. 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.

32.34. 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.

32.35. Perform a literature search to determine the properties and types of adhesives used to affix artificial hips onto the human femur.

Surface Technology

PART VII

Our first visual or tactile contact with the objects around us is through their *surfaces*. We can see or feel surface roughness, waviness, reflectivity, and other features, such as scratches, nicks, cracks, and discoloration. 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—influences such as the effect of surface preparation on fatigue life and the sensitivity of brittle materials to surface scratches and defects.

Machinery and accessories have numerous members that slide against each other: slideways, bearings, tools and dies for cutting and forming, and pistons and cylinders. Close examination will reveal that (a) some of these surfaces are smooth while others are rough, (b) some are lubricated while others are dry, (c) some are subjected to heavy loads while others support light loads, (d) some are subjected to elevated temperatures while others are at room temperature, and (e) some surfaces slide against each other at high relative speeds while others move slowly.

In addition to possessing geometric features, a surface constitutes a thin layer on the bulk material. A surface's physical, chemical, metallurgical, and mechanical properties depend not only on the material and its processing history, but also on the environment to which the surface is exposed. The term *surface integrity* is used to describe the chemical, mechanical, and metallurgical characteristics of a surface.

Because of the various mechanical, physical, thermal, and chemical effects that result from its processing history, the surface of a manufactured part usually possesses properties and behavior that are significantly different from those of its bulk. Although the bulk material generally determines the component's overall mechanical properties, the component's surfaces directly influence the part's performance in the following areas (Fig. VII.1):

- Friction and wear of tools, molds, dies, and of the products made.
- Effectiveness of lubricants during the manufacturing process and throughout the part's service life.
- Appearance and geometric features of the part and their role in subsequent operations, such as welding, soldering, adhesive bonding, painting, and coating.
- Resistance to corrosion.

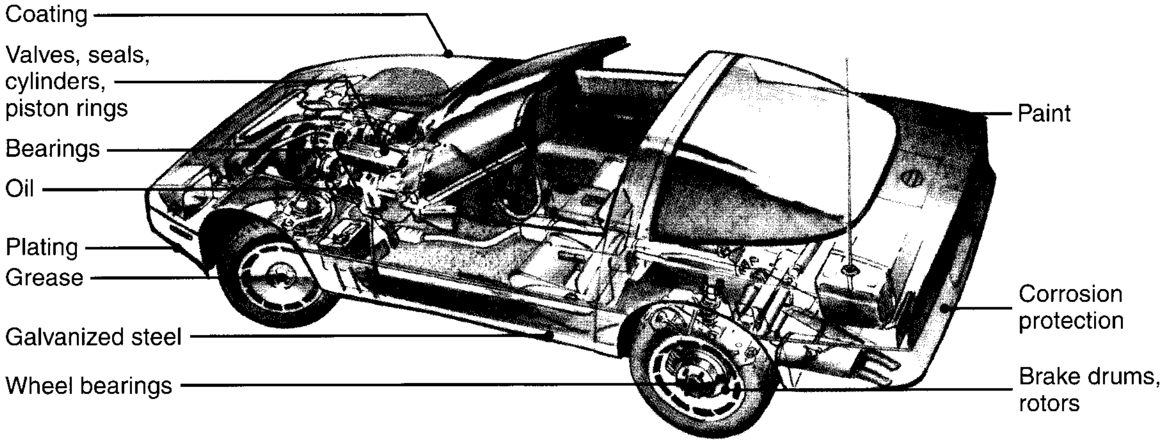


FIGURE VII.1 Components in a typical automobile that are related to the topics described in Part VII.

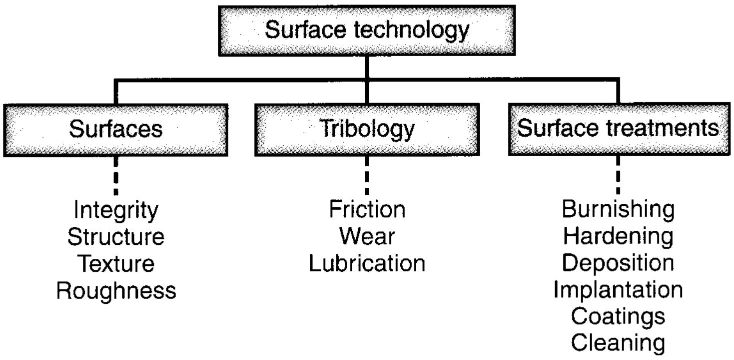


FIGURE VII.2 An outline of topics covered in Part VII.

- Crack initiation as a result of surface defects such as roughness, scratches, seams, and heat-affected zones, which can lead to weakening and premature failure of the part, through fatigue, for instance.
- Thermal and electrical conductivity of contacting bodies. For example, rough surfaces have higher thermal and electrical resistances than smooth surfaces.

Following the outline shown in Fig. VII.2, this part of the book will present surface characteristics in terms of their structure and topography. The material and process variables that influence the friction and wear of materials will then be described. Several mechanical, thermal, electrical, and chemical methods can be used to modify surfaces for improved frictional behavior, effectiveness of lubricants, resistance to wear and corrosion, and surface finish and appearance.

Surface Roughness and Measurement; Friction, Wear, and Lubrication

CHAPTER 33

- This chapter describes the features of surfaces that have a direct effect on both the selection of manufacturing processes and the service life of the parts produced.
- Surface features such as roughness, texture, and lay are discussed, as well as approaches used to quantitatively describe and measure surfaces.
- The chapter also examines the nature of friction, its role in manufacturing, and the factors involved in its magnitude.
- Wear and lubrication are then discussed, along with various approaches to minimizing wear.
- The chapter ends with a summary of commonly used lubricants, their additives, and their selection for a particular manufacturing process.

33.1 Introduction

This chapter begins with a description of the nature of surfaces, which are distinct entities and have properties significantly different from those of the bulk; this is particularly true for metals, because of various surface-oxide layers. Several defects can exist on a surface, depending on the manner in which the surface was generated. These defects (as well as various surface textures) can have a major influence on the surface integrity of workpieces, tools, and dies.

Two common methods of surface-roughness measurement in engineering practice are described in this chapter, including the instrumentation involved and a brief description of surface-roughness requirements in engineering design. Because of their increasing importance in precision manufacturing and nanofabrication, three-dimensional surface measurements are discussed and illustrated as well.

The chapter also describes those aspects of *friction*, *wear*, and *lubrication*—collectively known as **tribology**—which are relevant to manufacturing processes and operations and to the service life of products. We first describe friction and wear for metallic and nonmetallic materials and how they are influenced by various material and process variables. An understanding of these relationships is necessary for the proper selection of tool and die materials and of appropriate metalworking fluids for a particular operation. Wear has a major economic impact, as it is estimated that in the United States alone the total cost of replacing worn parts is more than $100 billion per year.

The chapter then describes the fundamentals of metalworking fluids, including the types, characteristics, and application of commonly used liquid and solid lubricants and the lubrication practices employed. The importance of biological and environmental considerations in the use, application, recycling, and ultimate disposal of metalworking fluids is also discussed.

33.2 Surface Structure and Integrity

Upon close examination of the surface of a piece of metal, it is found that it generally consists of several layers (Fig. 33.1):

1. The *bulk* metal (also known as the metal *substrate*) has a structure that depends on the composition and processing history of the metal.
2. Above the bulk metal is a layer that usually has been deformed plastically and work hardened to a greater extent than the bulk during the manufacturing process. The depth and properties of the work-hardened layer (*surface structure*) depend on such factors as the processing method used and how much frictional sliding the surface encountered. For example, if the surface is produced by machining with a dull and worn tool or the surface is ground by a dull grinding wheel, the work-hardened layer will be relatively thick and usually will also have residual stresses.
3. Unless the metal is processed and kept in an inert (oxygen-free) environment or is a noble metal (such as gold or platinum), an *oxide layer* forms over the work-hardened layer. The oxide on a metal surface is generally much harder than the base metal; hence, it is more abrasive. As a result, it has important effects on friction, wear, and lubrication. For example,
   - *Iron* has an 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 (which is exposed to the environment).
   - *Aluminum* has a dense, amorphous (without any crystalline structure) 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 then 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*, as described in Section 3.8).

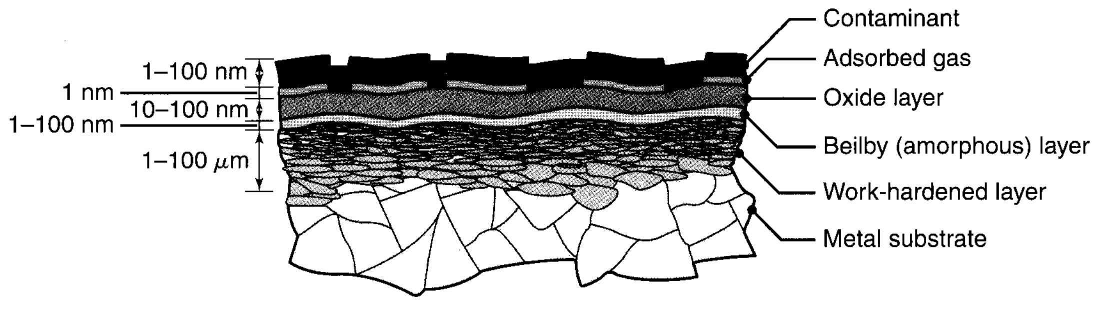


FIGURE 33.1 Schematic illustration of a cross section of the surface structure of a metal. The thickness of the individual layers depends on both processing conditions and the processing environment. *Source*: After E. Rabinowicz and B. Bhushan.

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, grease, lubricant residues, cleaning-compound residues, and pollutants from the environment.

Surfaces have properties that generally are very different from those of the substrate material. The factors that pertain to the surface structures of the metals just described are also factors in the surface structure of plastics and ceramics. The surface texture of these materials depends (as with metals) on the method of production.

Surface Integrity. *Surface integrity* describes not only the topological (geometric) 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 manufacturing operations, because it influences such properties as fatigue strength, resistance to corrosion, and service life (see, for example, Fig. 2.29).

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 factors, such as (a) defects in the original material, (b) the method by which the surface is produced, and (c) improper control of the process parameters (which can result in excessive stresses, temperatures, or surface deformation).

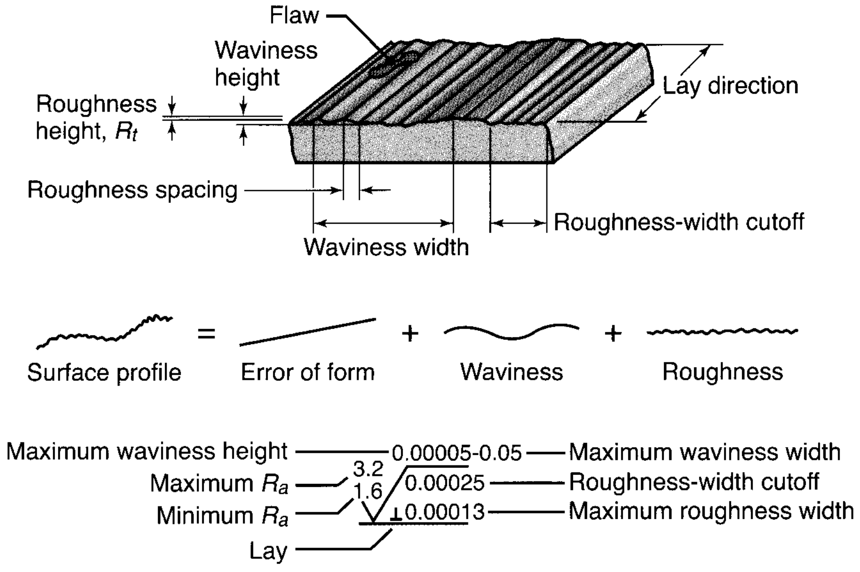
The following are general definitions of the major **surface defects** (listed in alphabetical order) found in practice:

- **Cracks** may be external or internal; cracks that require a magnification of 10× or higher to be seen by the naked eye are called **microcracks.**
- **Craters** are shallow depressions.
- The **heat-affected zone** is the portion of a metal which is subjected to thermal cycling without melting, such as that shown in Fig. 30.17.
- **Inclusions** are small, nonmetallic elements or compounds in the material.
- **Intergranular attack** is the weakening of grain boundaries through liquid-metal embrittlement and corrosion.
- **Laps, folds,** and **seams** are surface defects resulting from the overlapping of material during processing.
- **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 chemical or physical attack.
- **Residual stresses** (either tension or compression) on the surface are caused by nonuniform deformation and a nonuniform temperature distribution.
- **Splatter** is small resolidified molten metal particles deposited on a surface, as during welding.
- **Surface plastic deformation** is a severe surface deformation caused by high stresses due to factors such as friction, tool and die geometry, worn tools, and processing methods.

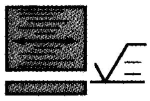
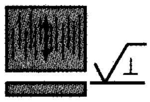
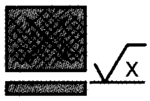
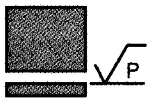
33.3 Surface Texture and Roughness

Regardless of the method of production, all surfaces have their own characteristics, which collectively are referred to as *surface texture.* Although the description of surface texture as a geometrical property is complex, the following 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, or inclusions.
- **Lay (directionality)** is the direction of the predominant surface pattern, usually visible to the naked eye.



(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 microinches. (b) Common surface lay symbols.

- **Roughness** is defined as closely spaced, irregular deviations on a small scale; it is expressed in terms of its height, width, and distance along the surface.
- **Waviness** is a recurrent deviation from a flat surface; it is measured and described in terms of the space between adjacent crests of the waves (*waviness width*) and height between the crests and valleys of the waves (*waviness height*).

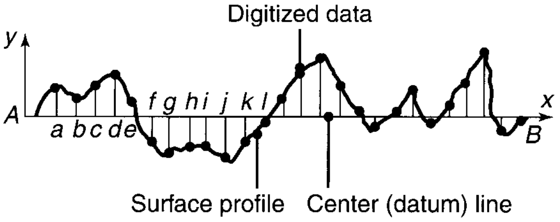


FIGURE 33.3 Coordinates used for surface-roughness measurement defined by Eqs. (33.1) and (33.2).

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, and is defined as

$$R_a = \frac{a + b + c + d + \cdots}{n}, \qquad (33.1)$$

where all ordinates a, b, c, ..., are absolute values and n is the number of readings.

The **root-mean-square roughness** (R_q, formerly identified as RMS) is defined as

$$R_q = \sqrt{\frac{a^2 + b^2 + c^2 + d^2 + \cdots}{n}}. \qquad (33.2)$$

The datum line AB in Fig. 33.3 is located so 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) also can be used and 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.

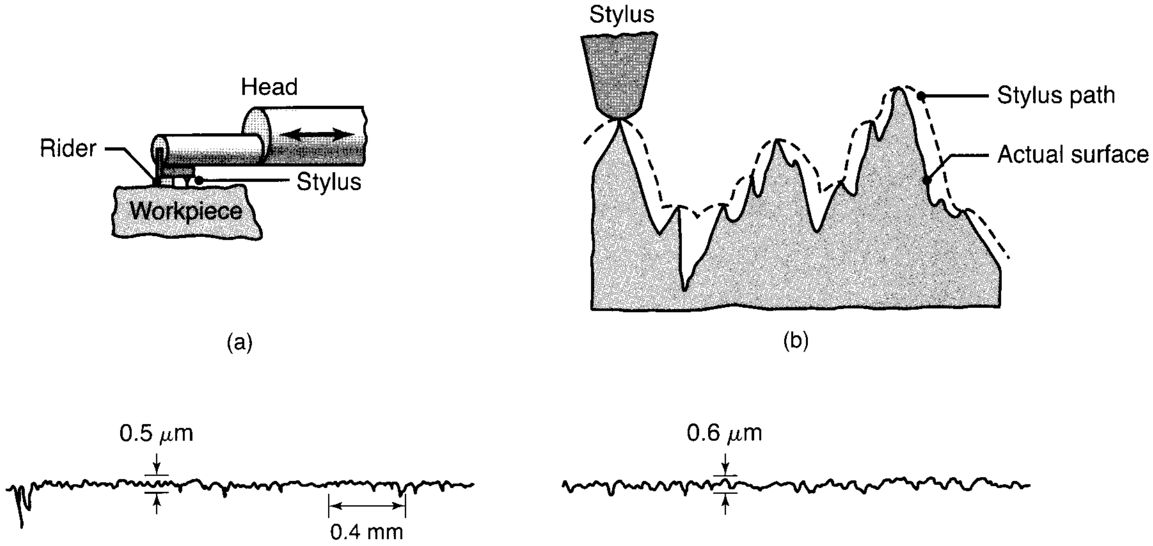
The units generally used for surface roughness are μm (microns).

Because of its simplicity, the arithmetic mean value (R_a) was adopted internationally in the mid-1950s and is used widely in engineering practice. Dividing Eq. (33.2) by Eq. (33.1) yields the ratio R_q/R_a, which, for typical surfaces produced by machining and finishing processes is 1.1 for cutting, 1.2 for grinding, and 1.4 for lapping and honing.

In general, a surface cannot be described by its R_a or R_q value alone, since 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 significantly. However, this type of surface profile can be significant in terms of friction, wear, and fatigue characteristics of a manufactured product. Consequently, it is important to analyze a surface in great detail, particularly for parts to be used in critical applications.

Symbols for Surface Roughness. Acceptable limits for surface roughness are specified on technical drawings by symbols, typically shown around the check mark in the lower portion of Fig. 33.2a, and the values of these limits are placed to the left of the check mark. The symbols and their meanings concerning the lay are given in Fig. 33.2b. Note that the symbol for the lay is placed at the lower right of the check mark. Symbols used to describe a surface specify only its roughness, waviness, and lay; they do not include flaws. Therefore, whenever necessary, a special note is included in technical drawings to describe the method that should be used to inspect for surface flaws.

Measuring Surface Roughness. Typically, instruments called **surface profilometers** are used to measure and record surface roughness. A profilometer has a *diamond*



(c) Lapping

(d) Finish grinding

(e) Rough grinding

(f) Turning

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.

stylus that travels along a straight line over the surface (Fig. 33.4a). The distance that the stylus travels is called the **cutoff**, which generally ranges from 0.08 to 25 mm. A cutoff of 0.8 mm is typical for most engineering applications. The rule of thumb is that the cutoff must be large enough to include 10 to 15 roughness irregularities, as well as all surface waviness.

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 appears to be much rougher than it actually is. The recording instrument compensates for any surface waviness; it indicates only roughness.

Because of the finite radius of the diamond stylus tip, the path of the stylus 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 commonly used stylus-tip diameter is 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.

Surface roughness can be observed directly through an *optical* or *scanning-electron microscope*. Stereoscopic photographs are particularly useful for three-dimensional views of surfaces and also can be used to measure surface roughness.

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.

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.

Atomic-force microscopes (AFMs) are used to measure extremely smooth surfaces and even 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 design in typical engineering applications vary by as much as two orders of magnitude. Some examples are as follows:

• Bearing balls	0.025 μm
• Crankshaft bearings	0.32 μm
• Brake drums	1.6 μm
• Clutch-disk faces	3.2 μm.

Because of the many material and process variables involved, the range of roughness produced even within the same manufacturing process can be significant.

33.4 Friction

Friction plays an important role in manufacturing processes because of the relative motion and the forces that always are present on tools, dies, and workpieces. Friction (a) dissipates energy, thus generating heat, which can have detrimental effects on an operation, and (b) impedes free movement at interfaces, thus significantly affecting the flow and deformation of materials in metalworking processes. However, friction is not always undesirable; for example, without friction, it would be impossible to roll metals, clamp workpieces on machines, or hold drill bits in chucks.

There are a number of explanations for the phenomenon of friction. A commonly accepted theory of friction is the **adhesion theory**, 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. 33.5). The maximum slope of the surface ranges typically from 5° to 15°. In such a situation, the normal (contact) load, N, is supported by minute **asperities** (small projections from the surface) that are in contact with each other. Therefore, the normal stresses at these asperities are high; this causes *plastic deformation* at the junctions. Their contact

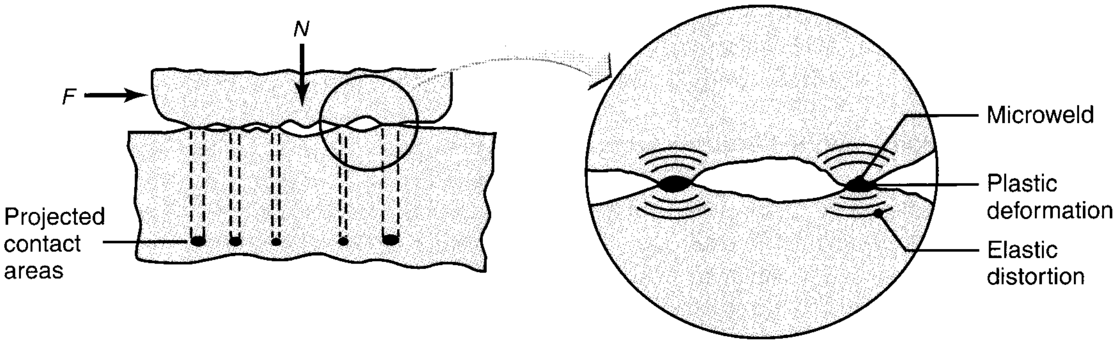


FIGURE 33.5 Schematic illustration of the interface of two bodies in contact showing real areas of contact at the asperities. In engineering surfaces, the ratio of the apparent-to-real areas of contact can be as high as 4 to 5 orders of magnitude.

creates an *adhesive bond*—the asperities form *microwelds*. Cold-pressure welding (see Section 31.2) is based on this principle.

Another theory of friction is the **abrasion theory**, which is based on the notion that an asperity from a hard surface (such as a tool or a die) penetrates and plows through a softer surface (the workpiece). Plowing may (a) cause displacement of the material and/or (b) produce small chips or slivers, as in cutting and abrasive processes. Other explanations for frictional behavior have been suggested, but for most applications in manufacturing, adhesion and abrasion mechanisms are the most relevant.

The sliding motion between two bodies with an interface as just described is possible only if a *tangential force* is applied. This force, called the **friction force**, F, is required to *shear* the junctions or *plow* through the softer material. The ratio F/N (Fig. 33.5a) is the **coefficient of friction**, μ. Depending on the materials and processes involved, coefficients of friction in manufacturing vary significantly. For example, in metal-forming processes, μ ranges from about 0.03 for cold working to 0.7 for hot working and from 0.5 to as much as 2 for machining.

Almost all of the energy dissipated in overcoming friction is converted into *heat*, which raises surface temperature. A small fraction of the energy becomes *stored energy* (Section 1.6) in the plastically deformed surfaces. The temperature increases with increasing friction and sliding speed, decreasing thermal conductivity, and decreasing specific heat of the sliding materials (see also Section 21.4). The interface temperature may be high enough to soften and even melt the surfaces and, sometimes, to cause microstructural changes in the materials involved. Note that temperature also affects the viscosity and other properties of lubricants, with a sufficiently high temperature causing their breakdown.

Friction in Plastics and Ceramics. Because their strength is low compared with that of metals (Tables 2.2 and 7.1), *plastics* generally possess low frictional characteristics. This property makes plastics better than metals for bearings, gears, seals, prosthetic joints, and general friction-reducing applications, provided that the loads are not high. Because of this characteristic, polymers sometimes are described as *self lubricating*.

The factors involved in the friction and wear of metals are generally applicable to polymers as well. In sliding, the plowing component of friction in thermoplastics and elastomers is a significant factor because of their viscoelastic behavior (i.e., they exhibit both viscous and elastic behavior) and subsequent hysteresis loss (see Fig. 7.14). This condition can easily be simulated by dragging a dull nail across the surface of a piece of rubber and observing how the rubber quickly recovers its shape.

An important factor in plastics applications 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. If the temperature rise is not controlled, 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 PTFE (*Teflon*) is attributed to its molecular structure, which has no reactivity with metals. Consequently, its adhesion is poor and thus its friction is low.

The frictional behavior of *ceramics* is similar to that of metals; hence, adhesion and plowing at interfaces contribute to the friction force in ceramics as well. Usually, adhesion is less important with ceramics because of their high hardness, whereby the real area of contact at sliding interfaces is small.

Reducing Friction. Friction can be reduced through the selection of materials that have low adhesion (such as carbides and ceramics) and through the use of surface films and coatings. *Lubricants* (such as oils) or solid films (such as graphite) interpose an adherent film between the tool, die, and workpiece. This film minimizes adhesion and interactions of one surface with the other, thus reducing friction. Friction also can be reduced significantly by subjecting the tool– or die–workpiece interface to *ultrasonic vibrations*, typically at 20 kHz. The amplitude of the vibrations periodically separates the two surfaces and allows the lubricant to flow more freely into the interface during these separations.

Friction Measurement. Although the coefficient of friction can be estimated theoretically, it usually is determined experimentally, either during actual manufacturing processes or in simulated laboratory tests using small-scale specimens of various shapes. A test that has gained wide acceptance—particularly for bulk-deformation processes—is the **ring-compression test**. A flat ring is upset plastically between two flat platens (Fig. 33.6a). As its height is reduced, the ring expands

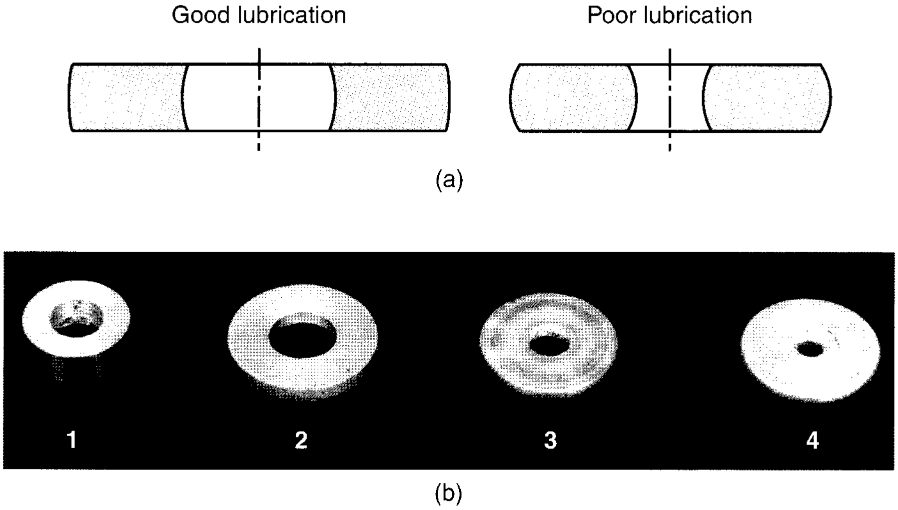


FIGURE 33.6 Ring-compression test between flat dies. (a) Effect of lubrication on type of ring-specimen barreling. (b) Test results: (1) original specimen and (2) to (4) increasing friction. *Source*: After A.T. Male and M.G. Cockcroft.

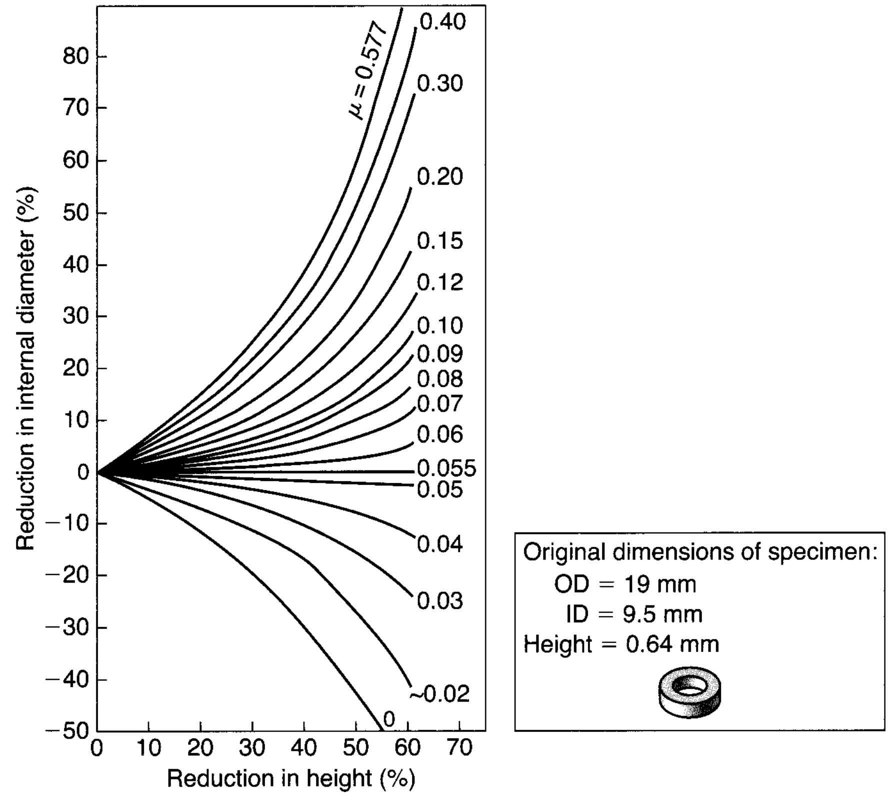


FIGURE 33.7 Chart to determine friction coefficient from a ring-compression test. Reduction in height and change in internal diameter of the ring are measured; then μ is read directly from this chart. For example, if the ring specimen is reduced in height by 40% and its internal diameter decreases by 10%, the coefficient of friction is 0.10.

radially outward. If friction at the interfaces is zero, both the inner and outer diameters of the ring expand as if it were a solid disk. With increasing friction, however, the internal diameter becomes smaller and barreling occurs. For a particular reduction in height, there is a critical friction value at which the internal diameter increases from its original value if μ is lower and decreases if μ is higher (Fig. 33.6b).

By measuring the change in the specimen's internal diameter and using the curves shown in Fig. 33.7 (which are obtained through theoretical analyses), the coefficient of friction can be determined. Note that each ring geometry and material has its own specific set of curves. The most common geometry of a specimen has an outer diameter, an inner diameter, and height proportions of 6:3:2, respectively. 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 from the appropriate chart.

EXAMPLE 33.1 Determination of Coefficient of Friction

In a ring-compression test, a specimen 10 mm in height and with an outside diameter (OD) of 30 mm and an inside diameter (ID) of 15 mm is reduced in thickness by 50%. Determine the coefficient of friction, μ, if the OD is 38 mm after deformation.

Solution First it is necessary to determine the new ID (which is obtained from volume constancy) as follows:

$$\text{Volume} = \frac{\pi}{4}(30^2 - 15^2)10 = \frac{\pi}{4}(38^2 - \text{ID}^2)5.$$

From this equation, the new ID is calculated as 9.7 mm.

Thus, the change in internal diameter is

$$\Delta\text{ID} = \frac{(9.7 - 15)}{15} = -0.35 \text{ or } 35\% \text{ (decrease)}.$$

With a 50% reduction in height and a 35% reduction in internal diameter, the friction coefficient can be obtained from Fig. 33.7 as

$$\mu = 0.21.$$

33.5 Wear

Wear has important technological and economic significance because it changes the shapes of tools and dies and, consequently, affects the tool life, tool and die dimensions, and thus the quality of the parts produced. The importance of wear is evident in the number of parts and components that continually have to be replaced or repaired. Examples of wear in manufacturing processes include dull drill bits that have to be reground, worn cutting tools that have to be indexed or changed, tools and dies that have to be repaired or replaced, and countless other situations. *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.

Although wear generally alters a part's surface topography and may result in severe surface damage, it can also have a beneficial effect. The *running-in* period for various machines and engines produces wear in order to remove the peaks from asperities (Fig. 33.8). Thus, under controlled conditions, wear may

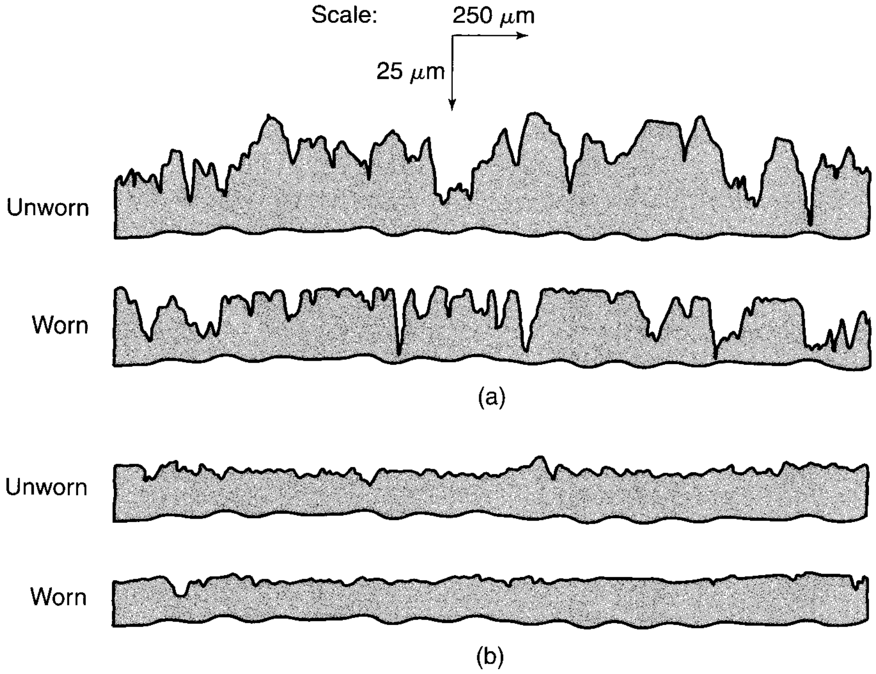


FIGURE 33.8 Changes in original (a) wire-brushed and (b) ground-surface profiles after wear. Note the difference in the vertical and horizontal scales. *Source*: After E. Wild and K.J. Mack.

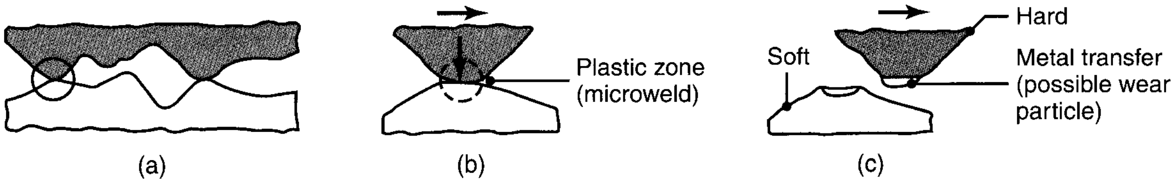


FIGURE 33.9 Schematic illustration of (a) two contacting asperities, (b) adhesion between two asperities, and (c) the formation of a wear particle.

be regarded as a type of smoothing or polishing process. Note that writing with an ordinary pencil or chalk is a wear process, and the words written actually are formed from wear particles.

Adhesive Wear. If a tangential force is applied to the model shown in Fig. 33.9, shearing can take place either (a) at the original interface or (b) along a path below or above the interface, in either case causing *adhesive wear*, also called *sliding wear*. Because of factors such as strain hardening at the asperity contact, diffusion, and mutual solid solubility, the adhesive bonds often are stronger than the base metals. Thus, during sliding, fracture usually follows a path in the weaker or softer component; that is how a wear fragment is generated. Although this fragment is typically attached to the harder component (the upper surface in Fig. 33.9c), it eventually becomes detached during further rubbing at the interface and develops into a loose **wear particle**.

In more severe cases, such as ones with high loads and strongly bonded asperities, adhesive wear is described as *scuffing*, *smearing*, *tearing*, *galling*, or *seizure* (severe wear). Oxide layers on surfaces have a large influence on adhesive wear, sometimes acting as a protective film, resulting in **mild wear**—which consists of small wear particles.

Adhesive wear can be reduced by one or more of the following methods:

a. Selecting materials that do not form strong adhesive bonds.

b. Using a harder material as one member of the pair.

c. Using materials that oxidize more easily.

d. Applying hard coatings that serve methods **a** to **c**. Coating one surface with a soft material (such as tin, silver, lead, or cadmium) also is effective in reducing wear.

e. Using an appropriate lubricant.

Abrasive Wear. This type of wear is caused by a hard, rough surface or a surface containing hard, protruding particles sliding across another surface. As a result, *microchips* or *slivers* are produced as wear particles, thereby leaving grooves or scratches on the softer surface (Fig. 33.10). In fact, processes such as filing, grinding, ultrasonic machining, and abrasive-jet and abrasive water-jet machining, described in the preceding chapters, act in this manner. The difference in these operations is that, unlike wear that is generally not intended or wanted, the process parameters are *controlled* to produce the desired shapes and surfaces through wear.

There are two basic types of abrasive wear. In **two-body wear**, abrasive action takes place between two sliding surfaces or between a hard, abrasive particle in contact with a solid body. This type is the basis of **erosive wear**. In **three-body wear**, an abrasive particle is present between two sliding solid bodies, such as a wear particle carried by a lubricant. Such a situation indicates the importance of properly filtering lubricants in metalworking operations to remove any contaminants.

The **abrasive-wear resistance** of pure metals and ceramics has been found to be directly proportional to their hardness. Thus, abrasive wear can be reduced by increasing the hardness of materials (usually by heat treating) or by reducing the normal load. Elastomers and rubbers resist abrasive wear as well, because they deform *elastically* and then recover when abrasive particles cross over their surfaces. The best examples are automobile tires, which last a long time even though they are operated on road surfaces that generally are rough and abrasive. Even hardened steels would not last long under such conditions.

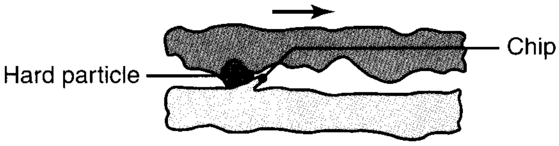


FIGURE 33.10 Schematic illustration of abrasive wear in sliding. Longitudinal scratches on a surface usually indicate abrasive wear.

Corrosive Wear. Also known as *oxidation* or *chemical wear*, this type of wear is caused by chemical and electrochemical reactions between the surface and the environment. The fine corrosive products on the surface constitute the wear particles in corrosive wear. When the corrosive layer is destroyed or removed through sliding or abrasion, another layer begins to form, and the process of removal and corrosive-layer formation is repeated. Among corrosive media are water, seawater, oxygen, acids, chemicals, and atmospheric hydrogen sulfide and sulfur dioxide.

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. Fatigue wear, also called *surface fatigue* or *surface-fracture* wear, is caused when the surface of a material is subjected to cyclic loading, such as rolling contact in bearings. The wear particles usually are formed through *spalling* or *pitting*. **Thermal fatigue** is another type of fatigue wear, whereby surface cracks are generated by thermal stresses from thermal cycling, as when a cool die repeatedly contacts hot workpieces. The cracks then join, and the surface begins to spall, producing fatigue wear in a phenomenon known as *heat checking*.

In manufacturing operations, fatigue wear usually occurs on dies in hot-working and die-casting operations. This type of 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 may act as local points for crack initiation.

Other Types of Wear. Several other types of wear can be seen in manufacturing operations.

- **Erosion** is caused by loose particles abrading a surface.
- **Fretting corrosion** occurs at interfaces that are subjected to very small reciprocal movements.
- **Impact wear** is the removal (by impacting particles) of small amounts of material from a surface.

In many cases, component wear is the result of a combination of different types of wear. Note, for example, in Fig. 33.11, that even in the same forging die, various types of wear take place in different locations; a similar situation can exist in cutting tools, as shown in Fig. 21.18.

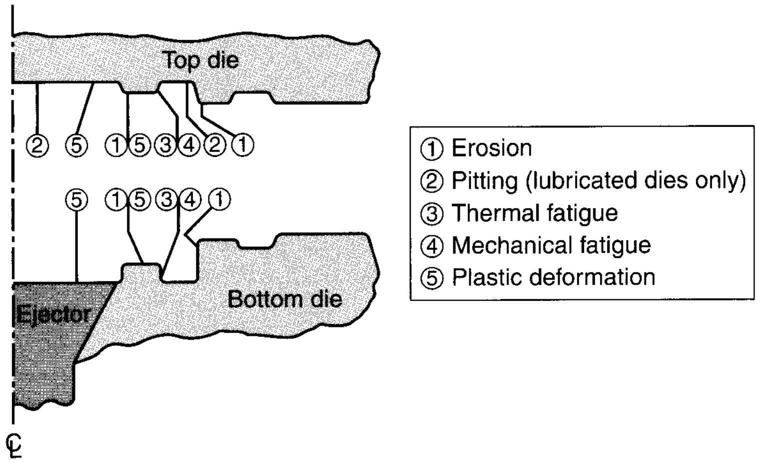


FIGURE 33.11 Types of wear observed in a single die used for hot forging. *Source*: After T.A. Dean.

Wear of Thermoplastics. The wear behavior of thermoplastics is similar to that of metals. Their abrasive-wear behavior depends partly on the ability of the polymer to deform and recover elastically, as in elastomers. Typical polymers with good wear resistance are polyimides, nylons, polycarbonate, polypropylene, acetals, and high-density polyethylene. These polymers are molded or machined to make gears, pulleys, sprockets, and similar mechanical components. Because thermoplastics can be made with a wide variety of compositions, they also can be blended with internal lubricants, such as polytetrafluoroethylene, silicon, graphite, molybdenum disulfide, and rubber particles, that are interspersed within the polymer matrix.

Wear of Reinforced Plastics. The wear resistance of reinforced plastics depends on the type, amount, and direction of reinforcement in the polymer matrix. Carbon, glass, and aramid fibers all improve wear resistance. Wear takes place when fibers are pulled out of the matrix (*fiber pullout*). Wear is highest when the sliding direction is parallel to the fibers, because they can be pulled out more easily in this case. Long fibers increase the wear resistance of composites, because they are more difficult to pull out and they prevent cracks in the matrix from propagating to the surface as easily.

Wear of Ceramics. When ceramics slide against metals, wear is caused by (a) small-scale plastic deformation and brittle surface fracture, (b) surface chemical reactions, (c) plowing, and (d) fatigue. Metals can be transferred to the oxide-type ceramic surfaces, forming metal oxides. Thus, sliding actually takes place between the metal and the metal-oxide surface.

33.6 Lubrication

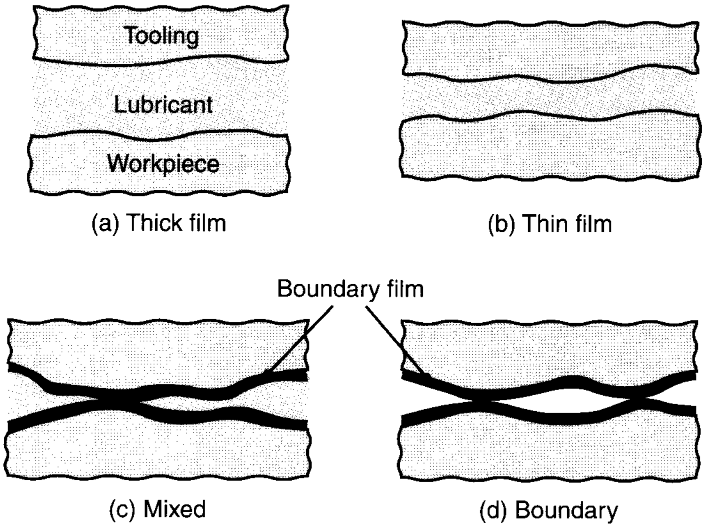
There is evidence that the lubrication of surfaces to reduce friction and wear dates back about four millenia to lubricate various linear and rotary moving components. For example, chariot wheels were lubricated with beef tallow in 1400 B.C. Various oils also were used for lubrication in metalworking processes (see Table I.2), beginning in about 600 A.D.

In manufacturing processes, as noted in various chapters, the surfaces of tools, dies, and workpieces are subjected to (a) *force* and *contact pressure*, which ranges from very low values to multiples of the yield stress of the workpiece material;

(b) *relative speed*, from very low to very high; and (c) *temperature*, which ranges from ambient to melting. In addition to selecting appropriate materials and controlling process parameters to reduce friction and wear, **lubricants**, or, more generally, **metalworking fluids**, are applied widely.

Regimes of Lubrication. There are four regimes of lubrication that are generally of interest in manufacturing operations (Fig. 33.12):

1. In **thick-film lubrication**, the surfaces are separated completely by a film of lubricant and lubricant viscosity is an important factor. Such films can develop in some regions of the workpiece in high-speed operations and also can develop from high-viscosity lubricants that become trapped at die–workpiece interfaces. A thick lubricant film results in a dull, grainy surface appearance on the workpiece after forming operations, with the degree of roughness varying with grain size. In operations such as coining and precision forging, trapped lubricants are undesirable, because they prevent accurate shape generation.
2. 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 and the process is known as **thin-film lubrication**. This condition raises the friction at the sliding interfaces and results in slight wear.
3. In **mixed 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 fluid film trapped in pockets, such as the valleys between asperities.
4. In **boundary lubrication**, the load is supported by contacting surfaces that are covered with a *boundary film* of lubricant (Fig. 33.12d)—a thin molecular lubricant layer. The film is attracted to the metal surfaces and prevents direct metal-to-metal contact of the two bodies, thus reducing wear. Boundary lubricants typically are natural oils, fats, fatty acids, esters, or soaps. However, boundary films can *break down* (a) as a result of 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 may also score severely.



(a) Thick film (b) Thin film (c) Mixed (d) Boundary

FIGURE 33.12 Regimes of lubrication generally occurring in metalworking operations. *Source*: After W.R.D. Wilson.

Other Considerations. Note that the valleys in the surface of the contacting bodies (see Figs. 33.2a, 33.4, and 33.5) can serve as local reservoirs or pockets for lubricants, thereby supporting a substantial portion of the load. The workpiece, but not the die, should have the rougher surface; otherwise, the rougher and harder die surface, 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 is another 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 die–workpiece interface. With proper selection of process parameters, a relatively thick lubricant film can be entrained and maintained.

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 now available with diverse chemistries, properties, and characteristics that fulfill these requirements. (See also Section 22.12 on *cutting fluids*.)

33.7.1 Oils

Oils maintain high film strength on the surface of a metal, as it readily 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. Moreover, it is difficult and costly to remove oils from component surfaces that are to be painted or welded, and it is difficult to dispose of them. (See also Section 34.16.)

The sources of oils may be **mineral** (*petroleum* or *hydrocarbon*), **animal**, or **vegetable**. Oils may be *compounded* with any number of additives or with other oils. Compounding is used to change such properties as viscosity–temperature behavior, surface tension, heat resistance, and boundary-layer characteristics. Undiluted mineral oils, with or without additives, are known as **neat oils**. Oils can be contaminated by the lubricants used for the slideways and various components of the machine tools and metalworking machinery. Because these oils have characteristics different from those used for the process itself, they can have adverse effects as a metalworking lubricant. When present in the metalworking fluid itself, such oils are known as **tramp oil**.

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 together—hence the term immiscible. Milky in appearance, emulsions also are known as **water-soluble oils** or **water-based coolants** and are of two types. In *indirect emulsion*, water droplets are dispersed in the oil. In *direct emulsion*, 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 has detrimental effects on tool life, the surface integrity of workpieces, and the dimensional accuracy of parts.

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 any mineral oils. 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 are typically reaction products of sodium or potassium salts with fatty acids. Alkali soaps are soluble in water, but other metal soaps generally are insoluble. Soaps are effective *boundary lubricants* and 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 and generally consist 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, as lubricants for stainless steels and high-temperature alloys.

33.7.5 Additives

Metalworking fluids usually are *blended* with various additives, including

- Oxidation inhibitors
- Rust-preventing agents
- Foam inhibitors
- Wetting agents
- Odor-controlling agents
- 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 and form adherent surface films of metallic sulfides and chlorides. These films have low shear strength and good antiweld properties and thus can reduce friction and wear effectively. However, they may preferentially attack the cobalt binder in tungsten-carbide tools and dies (through *selective leaching*), causing changes in the surface roughness and 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. Described here are four of the most commonly used *solid lubricants*.

Graphite. The general properties of graphite were described in Section 8.6. Graphite is weak in shear along its basal planes (see Fig. 1.4); thus, it has a 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. In a vacuum or an inert-gas atmosphere, friction is very high; in fact, graphite can be abrasive in these environments. 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 an alcohol.

Molybdenum Disulfide. A widely used lamellar solid lubricant, molybdenum disulfide is somewhat similar in appearance to graphite. However, unlike graphite, it has a high friction coefficient in an ambient environment. Oils commonly are used as carriers for molybdenum disulfide (MoS_2) and as a lubricant at room temperature. Molybdenum disulfide is 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; polymers such as polytetrafluoroethylene, polyethylene, and methacrylates are also used. However, these coatings have limited applications because of their lack of strength under high contact stresses, especially at elevated temperatures.

Soft metals also 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. 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.

Glasses. Although it is a solid material, glass becomes viscous at elevated temperatures and hence can serve as a liquid lubricant. Viscosity is a function of temperature (but not of pressure) and depends on the type of glass. Poor thermal conductivity also makes glass attractive, since it acts as a thermal barrier between hot workpieces and relatively cool dies. Glass lubrication is typically used in such applications as hot extrusion and hot forging.

Conversion Coatings. Lubricants may not always adhere properly to workpiece surfaces, particularly under high normal and shearing stresses. Failure to adhere has the greatest effects 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.) This reaction leaves a somewhat rough and spongy surface, which acts as a carrier for the lubricant. After treatment, any excess acid from the surface is removed with the use of borax or lime. A liquid lubricant, such as a soap, is then applied to the surface. The lubricant film adheres to the surface and cannot be scraped off easily. *Zinc-phosphate* conversion coatings often are used on carbon and low-alloy steels. *Oxalate* coatings are used for stainless steels and high-temperature alloys.

Fullerenes (Buckyballs). As described in Section 8.6.1, these are carbon molecules in the shape of soccer balls. When placed between sliding surfaces, buckyball molecules act like tiny ball bearings. They can perform well as solid lubricants and are being investigated as bearings in aerospace applications.

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 here.

In selecting an oil as a lubricant, it is important to investigate its viscosity, temperature, and pressure characteristics. Low viscosity can have significant detrimental effects and cause high friction and wear. 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, but as lubricants, they are not as effective as oils. It is estimated that water-based fluids are used in 80 to 90% of all machining operations.

Specific requirements for metalworking fluids are as follows:

- They should not leave any harmful residues that could interfere with operations.
- They should not stain or corrode the workpiece or the equipment.
- Periodic inspection is necessary to detect deterioration caused by bacterial growth, accumulation of oxides, chips, wear debris and general degradation and breakdown due to temperature and time. The presence of wear particles is particularly important, because they cause damage to the system; proper inspection and filtering are thus essential.

After the completion of manufacturing operations, workpiece surfaces usually have lubricant residues; these 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 used for this purpose are described in Section 34.16.

Biological and Environmental Considerations. Biological and environmental considerations are important factors in the selection of a metalworking fluid. Hazards may result if one contacts or inhales some of these fluids, such as inflammation of the skin (*dermatitis*) and long-term exposure to *carcinogens*. The improper disposal of metalworking fluids may 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 equipment for their proper treatment, recycling, and disposal. In the United States, 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

- 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 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).
- The geometric and material properties of surfaces can affect their friction, wear, fatigue, corrosion, and electrical and thermal conductivity properties significantly.
- The measurement and description of surface features (including their characteristics) are important aspects of manufacturing. The most common surface-roughness measurement is the arithmetic mean value. The instrument usually used to measure surface roughness is a profilometer.
- 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.
- Other important factors are the 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.
- A wide variety of metalworking fluids, including oils, emulsions, synthetic solutions, and solid lubricants, is available for specific applications.
- The selection and use of lubricants requires a careful consideration of many factors regarding the workpiece and die materials and the particular manufacturing process.
- Metalworking fluids have various lubricating and cooling characteristics. Biological and environmental considerations also are important factors in selecting a metalworking fluid.

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 average
Running-in
Selective leaching
Self lubricating
Severe wear
Soaps
Solid lubricants
Substrate
Surface defects
Surface finish
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 parts

BIBLIOGRAPHY

Astakhov, V., **Tribology of Metal Cutting**, Elsevier, 2007.

Bhushan, B., **Introduction to Tribology**, Wiley, 2002.

——— (ed.), **Handbook of Micro/Nanotribology**, 2nd ed., CRC Press, 1998.

———, **Modern Tribology Handbook**, CRC Press, 2001.

Blau, P.J., **Friction Science and Technology**, Marcel Dekker, 1995.

Booser, E.R. (ed.), **Tribology Data Handbook**, CRC Press, 1998.

Burakowski, T., and Wiershon, T., **Surface Engineering of Metals: Principles, Equipment, Technologies**, CRC Press, 1998.

Byers, J.P. (ed.), **Metalworking Fluids**, Marcel Dekker, 1994.

Chattopadhyay, R., **Surface Wear: Analysis, Treatment, and Prevention**, ASM International, 2001.

Gohar, R., **Fundamentals of Tribology**, Imperial College Press, 2008.

Ludema, K.C., **Friction, Wear, Lubrication: A Textbook in Tribology**, CRC Press, 1996.

Nachtman, E.S., and Kalpakjian, S., **Lubricants and Lubrication in Metalworking Operations**, Marcel Dekker, 1985.

Neele, M.J. (ed.), **The Tribology Handbook**, 2nd ed., Butterworth-Heinemann, 1995.

Rabinowicz, E., **Friction and Wear of Materials**, 2nd ed., Wiley, 1995.

Schey, J.A., **Tribology in Metalworking—Friction, Lubrication and Wear**, ASM International, 1983.

Stachowiak, G.W., **Wear: Materials, Mechanisms and Practice**, Wiley, 2006.

Stachowiak, G.W., and Batchelor, A.W., **Engineering Tribology**, 3rd ed., Butterworth-Heinemann, 2005.

Totten, G.E., and Liang, H., **Mechanical Tribology: Materials Characterization and Applications**, CRC Press, 2004.

Williams, J.A., **Engineering Tribology**, Oxford University Press, 2005.

REVIEW QUESTIONS

33.1. Explain what is meant by (a) surface texture and (b) surface integrity.

33.2. List and explain the types of defects typically found on surfaces.

33.3. Define the terms (a) *roughness* and (b) *waviness*.

33.4. Explain why the results from a profilometer are not a true depiction of the actual surface.

33.5. Describe the features of the ring-compression test. Does it require the measurement of forces?

33.6. List the types of wear generally observed in engineering practice.

33.7. How can adhesive wear be reduced? Abrasive wear?

33.8. Explain the functions of a lubricant in manufacturing processes?

33.9. What is the role of additives in metalworking fluids?

33.10. Describe the factors involved in lubricant selection.

QUALITATIVE PROBLEMS

33.11. Give several examples that show the importance of friction in manufacturing processes as described in Parts III and IV.

33.12. 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.

33.13. What factors would you consider in specifying the lay of a surface for a part? Explain.

33.14. Explain why identical surface-roughness values do not necessarily represent the same type of surface.

33.15. Why are the requirements for surface-roughness design in engineering applications so broad? Explain with specific examples.

33.16. What is the significance of a surface-temperature rise resulting from friction? Give some examples based on topics covered in the preceding chapters.

33.17. Explain the causes of lay on surfaces.

33.18. Give several examples of how wear on molds, tools, and dies affects a manufacturing operation.

33.19. 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.

33.20. Give two examples in which waviness on a surface would be desirable. (b) Give two examples in which it would be undesirable.

33.21. Do the same as for Problem 33.20, but for surface roughness.

33.22. Describe your observations regarding Fig. 33.7.

33.23. Give the reasons that an originally round specimen in a ring-compression test may become oval after it is upset.

33.24. Explain the reason that the abrasive-wear resistance of a material is a function of its hardness.

33.25. On the basis of your own experience, make a list of parts and components that have to be replaced because of wear.

33.26. Explain why the types of wear shown in Fig. 33.11 occur in those particular locations in the forging die.

33.27. List manufacturing operations in which high friction is desirable and those in which low friction is desirable.

QUANTITATIVE PROBLEMS

▶**33.28.** 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.

33.29. 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.

▶**33.30.** Refer to Fig. 33.6b, 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.

▶**33.31.** Using Fig. 33.7, make a plot of the coefficient of friction versus the change in internal diameter for a constant reduction in height of 35%.

▶**33.32.** 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

33.33. 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.

33.34. 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?

33.35. Describe your own thoughts regarding biological and envronmental considerations in the use of metalworking fluids.

33.36. Wear can have detrimental effects in manufacturing operations. Can you visualize situations in which wear could be beneficial? Explain, and give some examples.

33.37. 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.

33.38. 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.

33.39. 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.

Surface Treatments, Coatings, and Cleaning

CHAPTER 34

- The preceding chapters have described methods of producing desired shapes from a wide variety of materials; although material and process selection is very important, often the surface properties of a component determine its performance or commercial success.
- This chapter describes various surface-finishing operations that can be performed for technical and aesthetic reasons subsequent to manufacturing a part.
- The chapter presents the surface treatment, cleaning, and coating processes that are commonly performed and includes a discussion of mechanical surface treatments such as shot peening, laser peening, and roller burnishing, with the benefit of imparting compressive residual stresses onto metal surfaces.
- Coating operations are then examined, including cladding, thermal spray operations, physical and chemical vapor deposition, ion implantation, and electroplating.
- The benefits of diamond and diamond-like carbon coatings are also investigated.
- Finally, surface-texturing, painting, and cleaning operations are described.

34.1 Introduction

After a part is manufactured, some of its surfaces may have to be processed further in order to ensure that they receive certain properties and characteristics. It may be necessary to perform **surface treatments** in order to

- *Improve resistance to wear, erosion, and indentation* (e.g., for machine-tool slideways, as shown in Figs. 23.2 and 35.1, wear surfaces of machinery, and shafts, rolls, cams, and gears)
- *Control friction* (on sliding surfaces of tools, dies, bearings, and machine ways)
- *Reduce adhesion* (of electrical contacts)
- *Improve lubrication* (surface modification to retain lubricants)
- *Improve resistance to corrosion and oxidation* (on sheet metals for automobile bodies, 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 surface texture* (appearance, dimensional accuracy, and frictional characteristics)
- *Impart decorative features* (color and texture).

Numerous techniques are used to impart these characteristics to various types of metallic, nonmetallic, and ceramic materials on the basis of mechanisms that involve (a) plastic deformation of the workpiece surface, (b) chemical reactions, (c) thermal means, (d) deposition, (e) implantation, and (f) organic coatings and paints. We begin with surface-hardening techniques and then continue with descriptions of different types of coatings that are applied to surfaces by various means. Some of these techniques also are used in the manufacture of semiconductor devices, as described in Chapter 28.

Techniques that are used to impart texture on workpiece surfaces and types of organic coatings used for various purposes are then described. The chapter ends with a discussion of methods used for cleaning manufactured surfaces before the components are assembled into the completed product and made ready for service. Environmental considerations regarding the fluids used and the waste material from various surface-treatment processes are also included, as they are an important factor to be considered.

34.2 Mechanical Surface Treatments

Several techniques are used to mechanically improve the surface properties of manufactured components. The more common methods are the following:

Shot Peening. In shot peening, the workpiece surface is impacted repeatedly with a large number of cast steel, glass, or ceramic shot (small balls), which make overlapping indentations on the surface. This action, using shot sizes that range from 0.125 to 5 mm in diameter, causes plastic surface deformation at depths up to 1.25 mm. Because the plastic deformation is not uniform throughout the part's thickness, shot peening causes compressive residual stresses on the surface, thus improving the fatigue life of the component by delaying the initiation of fatigue cracks. Unless the process parameters are controlled properly, the plastic deformation of the surface can be so severe that it can damage the surface. The extent of deformation can be reduced by *gravity peening*, which involves larger shot sizes, but fewer impacts on the workpiece surface.

Shot peening is used extensively on shafts, gears, springs, oil-well drilling equipment, and jet-engine parts, such as turbine and compressor blades. However, note that if these parts are subjected to high temperatures, the residual stress will begin to relax (thermal relaxation) and their beneficial effects will be diminished greatly. An example is gas-turbine blades performing at their operating temperatures.

Laser Shot Peening. In this process, also called *laser shock peening* and first developed in the mid-1960s (but not commercialized until much later), the workpiece surface is subjected to planar laser shocks (pulses) from high-power lasers. This surface-treatment process produces compressive residual-stress layers that are typically 1 mm deep with less than 1% of cold working of the surface. Laser 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 necessary for the process are on the order of 100 to 300 J/cm^2 and have a pulse duration of about 30 nanoseconds. Currently, the basic limitation of laser shot peening for industrial, cost-effective applications is the high cost of the high-power lasers (up to 1 kW) that must operate at energy levels of 100 J/pulse.

Water-jet Peening. In this more recently developed process, a water jet at pressures as high as 400 MPa impinges on the surface of the workpiece, inducing compressive residual stresses and surface and subsurface hardening at the same level as in shot peening. The water-jet peening process has been used successfully on steels and aluminum alloys. The control of process variables (jet pressure, jet velocity, the design of the nozzle, and its distance from the surface) is important in order to avoid excessive surface roughness and surface damage.

Ultrasonic Peening. This process uses a hand tool based on a piezoelectric transducer. Operating at a frequency of 22 kHz, it can have a variety of heads for different applications.

Roller Burnishing. In this process, also called *surface rolling*, the surface of the component is cold worked by a hard and highly polished roller or set of rollers. The process is used on various flat, cylindrical, or conical surfaces (Fig. 34.1). Roller burnishing improves surface finish by removing scratches, tool marks, and pits and induces beneficial 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 residual stresses and minimal plastic deformation.

Internal cylindrical surfaces also are burnished by a similar 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.

Roller burnishing is used to improve the mechanical properties of surfaces as well as their surface finish. It can be used either by itself or in combination with other finishing processes, such as grinding, honing, and lapping. The equipment can be mounted on various CNC machine tools for improved productivity and consistency of performance. All types of metals (soft or hard) can be roller burnished. Roller burnishing is typically used on hydraulic-system components, seals, valves, spindles, and fillets on shafts.

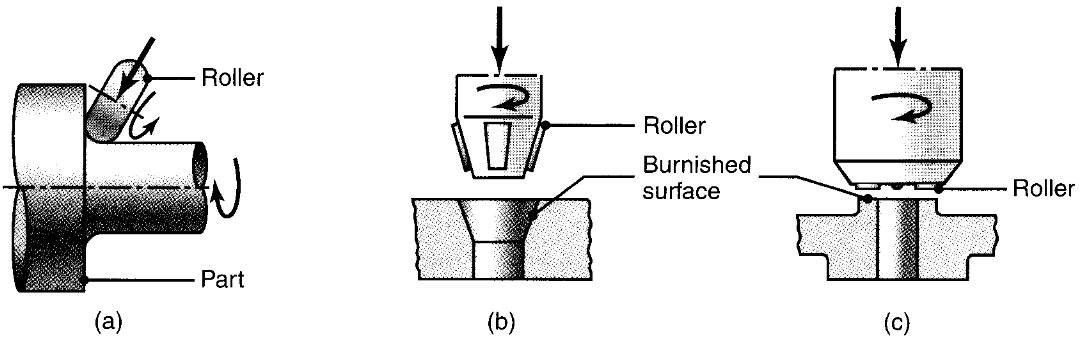


FIGURE 34.1 Burnishing tools and roller burnishing of (a) the fillet of a stepped shaft to induce compressive surface residual stresses for improved fatigue life; (b) a conical surface; and (c) a flat surface.

Explosive Hardening. In explosive hardening, the surfaces are subjected to high transient pressures through the placement and detonation of a layer of an explosive sheet directly on the workpiece surface. The contact pressures that develop as a result can be as high as 35 GPa and can last about 2 to 3 μs. Major increases in surface hardness can be achieved with this method, with very little change (less than 5%) in the shape of the component. Railroad rail surfaces, for example, are explosively hardened.

34.3 Mechanical Plating and Cladding

Mechanical Plating. In mechanical plating (also called *mechanical coating, impact plating*, or *peen plating*), fine metal particles are compacted over the workpiece surfaces by glass, ceramic, or porcelain beads that are propelled by rotary means (such as tumbling). This process, which is basically one of cold-welding particles onto a surface, typically is used for hardened-steel parts for automobiles, with plating thickness usually less than 25 μm.

Cladding. In this process, also called *clad bonding*, metals are bonded with a thin layer of corrosion-resistant metal through the application of pressure by rolls or other means. A typical application is the cladding of aluminum (*Alclad*), in which a corrosion-resistant layer of aluminum alloy (usually in sheet or tubular form) is clad over an aluminum-alloy body (core). The cladding layer is anodic to the core and usually has a thickness that is less than 10% of the total thickness.

Examples of cladding are 2024 aluminum clad with 1230 aluminum, and 3003, 6061, and 7178 aluminum clad with 7072 aluminum. Other applications are steels clad with stainless-steel or nickel alloys. The cladding material also may be applied with dies (as in cladding steel wire with copper) or explosives. Multiple-layer cladding is also utilized in special applications.

Laser cladding consists of the fusion of a different material over the substrate. It has been applied successfully to metals and ceramics, especially for enhanced friction and good wear behavior of the components.

34.4 Case Hardening and Hard Facing

Surfaces may be hardened by thermal means in order to improve their frictional and wear properties, as well as their resistance to indentation, erosion, abrasion, and corrosion. The most common methods are described next.

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 (gas and electricity), an electron beam or laser beam can be used as a heat source in surface hardening of both metals and ceramics. Case hardening, as well as some of the other surface-treatment processes described in this chapter, induces residual stresses on surfaces. The formation of martensite during case hardening causes compressive residual stresses on surfaces. Such stresses are desirable, because they improve the fatigue life of components by delaying the initiation of fatigue cracks.

Hard Facing. In this process, a relatively thick layer, edge, or point of wear-resistant hard metal is deposited on the workpiece surface by the fusion-welding techniques described in Chapter 30. Numerous layers (known as *weld overlay*) can be deposited to repair worn parts. Hard facing enhances the wear resistance of the materials; hence, such materials are used in the manufacture of tools, dies, and various industrial components. Worn parts also can be hard faced for extended use.

Spark Hardening. Hard coatings of tungsten, chromium, or molybdenum carbides can be deposited by an electric arc in a process 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 in spark hardening. Typical applications for these alloys are as 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 applied to 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** has also been used), and these processes are under continuous refinement. The surfaces to be sprayed are first cleaned of oil and dirt, and then roughened by, for example, grit blasting, to improve their bond strength (see Section 26.8). The coating material can be in the form of wire, rod, or powder, and when the droplets or particles impact the workpiece, they solidify and bond to the surface.

Depending on the process, particle velocities typically range from a low of about 150 to 1000 m/s, but can be higher for special applications. Temperatures are in the range from 3000° to 8000°C. The coating is hard and wear resistant, with a layered structure of deposited material. However, the coating can have a porosity as high as 20% due to entrapped air and oxide particles because of the high temperatures involved. Bond strength depends on the particular process and techniques used and is mostly mechanical in nature (hence the importance of roughening the surface prior to spraying), but can be metallurgical in some cases. 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 those used in rebuilding worn parts), structures, storage tanks, tank cars, rocket motor nozzles, and components that require resistance to wear and corrosion. In an automobile, thermal spraying typically can be applied to crankshafts, valves, fuel-injection nozzles piston rings, and engine blocks. The process is also used in the gas and petrochemical industries, for the repair of worn parts and to restore dimensional accuracy to parts that have not been machined or formed properly.

The source of energy in thermal-spraying processes is of two types: combustion and electrical.

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 flame wire spraying, but uses a metal powder instead of the wire.

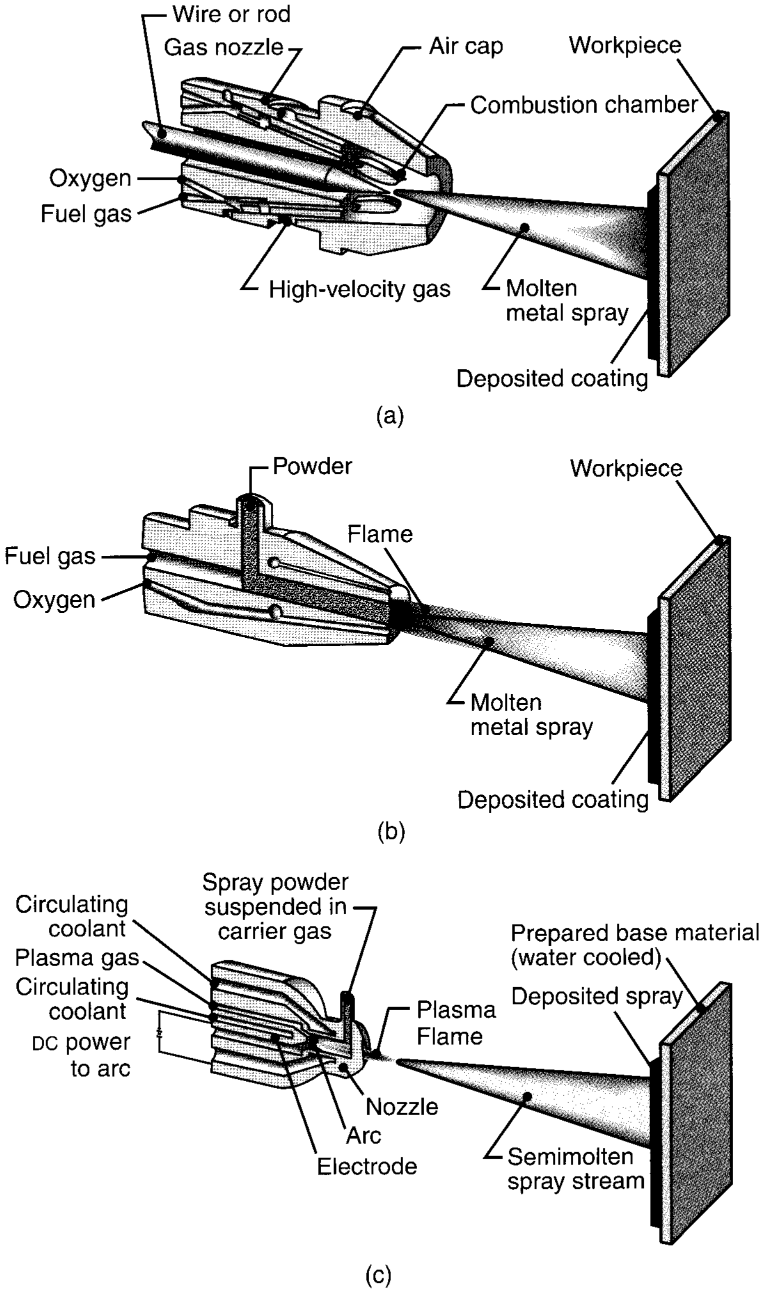


FIGURE 34.2 Schematic illustrations of thermal-spray operations: (a) thermal wire spray, (b) thermal metal-powder spray, and (c) plasma spray.

- **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 produces a high performance similar to that of the detonation gun, but is less expensive.

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. In this more recent development, 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.

EXAMPLE 34.1 Repair of a Worn Turbine-engine Shaft by Thermal Spraying

The shaft of the helical gear for a GE T-38 gas-turbine engine had two worn regions on its nitrided surfaces. The case-hardened depth was 0.3 mm. Even though the helical gears were in good condition, the part was considered scrap because there was no approved method of repair. The worn regions first were machined undersize, grit blasted, and coated with tungsten carbide (12% cobalt content; see Section 22.5) using the high-velocity oxyfuel-gas thermal-spraying (HVOF) technique. Then the part was finish machined to the dimensions of the original shaft. The total cost of repair was a fraction of the projected cost of replacing the part.

Source: Courtesy of Plasma Technology, Inc.

34.6 Vapor Deposition

Vapor deposition is a process in which the substrate (workpiece surface) is subjected to chemical reactions by gases that contain chemical compounds of the material to be deposited. The coating thickness is usually a few microns, which is much less than the thicknesses that result from the techniques described in Sections 34.2 and 34.3. The deposited materials can consist of metals, alloys, carbides, nitrides, borides, ceramics, or oxides. Control of the coating composition, thickness, and porosity are important. The substrate may be metal, plastic, glass, or paper. Typical applications for vapor deposition are the coating of cutting tools, drills, reamers, milling cutters, punches, dies, and wear surfaces.

There are two major vapor-deposition processes: physical vapor deposition and chemical vapor deposition.

34.6.1 Physical Vapor Deposition

The three basic types of *physical vapor deposition* (PVD) processes are (1) vacuum deposition, or arc evaporation; (2) sputtering; and (3) ion plating. These processes are carried out 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 being carried 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 is usually at room temperature or slightly higher for improved bonding). Coatings of uniform thickness can be deposited, even on complex shapes. In **arc deposition** (PV/ARC), the coating material (cathode) is evaporated by several arc evaporators

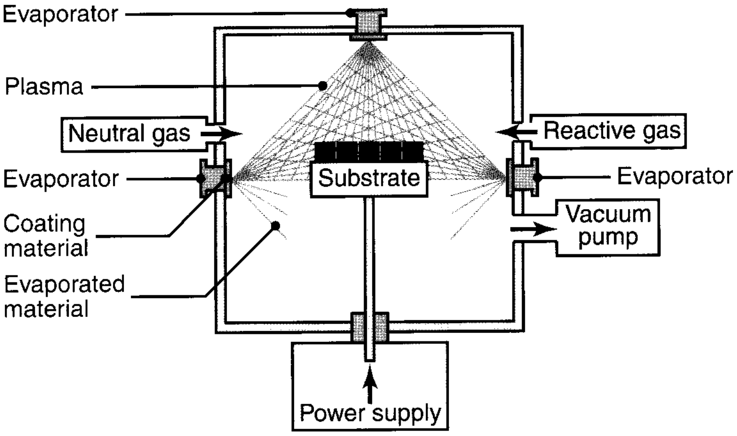


FIGURE 34.3 (a) Schematic illustration of the physical-vapor-deposition process. Note that there are three arc evaporators and the parts to be coated are placed on a tray inside the chamber.

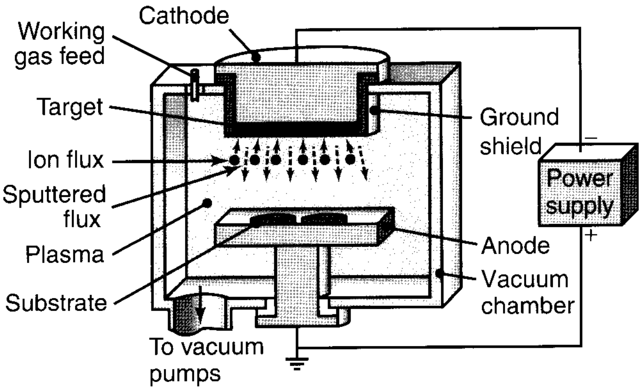


FIGURE 34.4 Schematic illustration of the sputtering process.

(Fig. 34.3) using highly localized electric arcs. The arcs produce a highly reactive plasma, which consists of ionized vapor of the coating material. The vapor condenses on the substrate (anode) and coats it. Applications of this process are both functional (oxidation-resistant coatings for high-temperature applications, electronics, and optics) and decorative (hardware, appliances, and jewelry). **Pulsed-laser deposition** is a more recent, related process in which the source of energy is a pulsed laser.

Sputtering. In *sputtering*, an electric field ionizes an inert gas (usually argon). The positive ions bombard the coating material (cathode) and cause sputtering (ejection) of its atoms. The atoms then 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 the oxides are deposited. Carbides and nitrides also are deposited by reactive sputtering. 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 that describes various combined processes of sputtering and vacuum evaporation. An electric field causes a glow, generating a plasma (Fig. 34.5). The vaporized atoms in this process are ionized only partially. **Ion-beam-enhanced (assisted) deposition** is capable of producing thin films as coatings for semiconductor, tribological, and optical applications. Bulky parts can be coated in large chambers using high-current power supplies of 15 kW and voltages of 100,000 DC. **Dual ion-beam deposition** is a hybrid coating technique that combines PVD with simultaneous ion-beam bombardment. This technique results in good adhesion on metals, ceramics, and polymers. Ceramic bearings and dental instruments are examples of its applications.

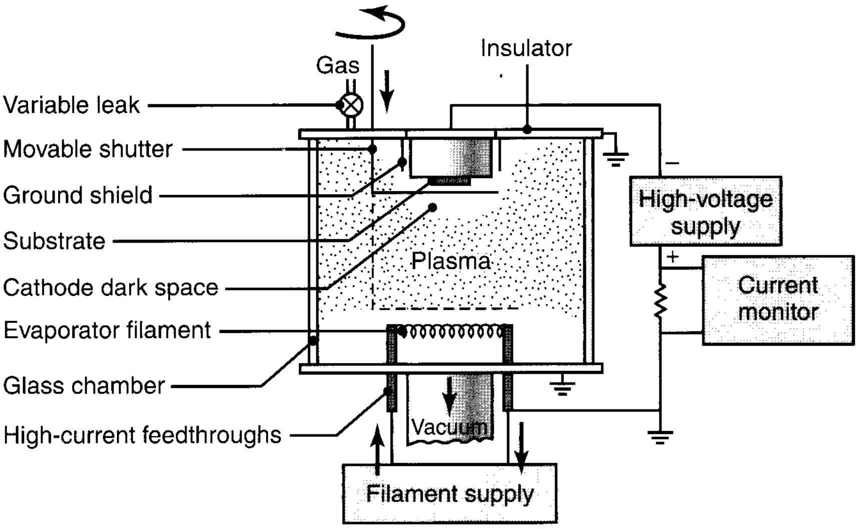


FIGURE 34.5 Schematic illustration of an ion-plating apparatus.

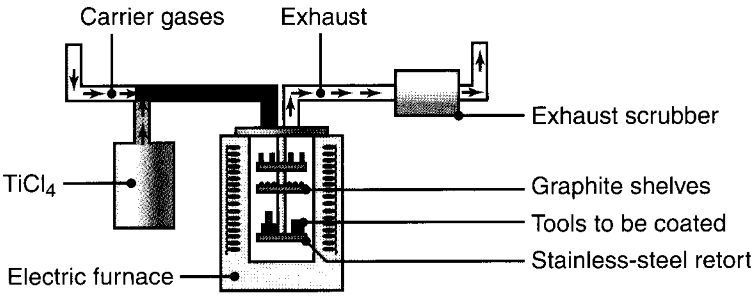


FIGURE 34.6 Schematic illustration of the chemical-vapor-deposition process. Note that parts and tools to be coated are placed on trays inside the chamber.

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 (TiN), the tools are placed on a graphite tray and heated at 950° to 1050°C at atmospheric pressure in an inert atmosphere. Titanium tetrachloride (a vapor), hydrogen, and nitrogen are then introduced into the chamber. The chemical reactions form titanium nitride on the tool surfaces. For a coating of titanium carbide, methane is substituted for the other gases.

Deposited coatings usually are thicker than those obtained with PVD. A typical cycle for CVD 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 the flow rates of the gases used, the time, and the temperature.

The types of coatings and the workpiece materials allowable are fairly unrestricted in CVD. Almost any material can be coated and any material can serve as a substrate, although bond strength may vary. The CVD process is also used to produce diamond coatings without binders, unlike polycrystalline diamond films, which use

1 to 10% binder materials. A more recent development in CVD is **medium-temperature CVD** (MTCVD). This technique results in a higher resistance to crack propagation than CVD affords.

34.7 Ion Implantation and Diffusion Coating

In *ion implantation*, ions (charged atoms) are introduced into the surface of the workpiece material. 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) 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.

Ion implantation is particularly effective on materials such as aluminum, titanium, stainless steels, tool and die steels, carbides, and chromium coatings. The process is typically used on cutting and forming tools, dies and molds, and metal prostheses, such as artificial hips and knees. When used in some specific applications, such as semiconductors (Section 28.3), ion implantation is called **doping**—meaning "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. The process has acquired different names (depending on the diffused element), as shown in Table 4.1, which lists diffusion processes such as *carburizing, nitriding*, and *boronizing*.

34.8 Laser Treatments

As described in various chapters of this book, lasers are having increasingly broader applications in manufacturing processes (laser machining, forming, joining, rapid prototyping, and metrology) and surface engineering (laser peening, alloying, surface treatments, and texturing). Powerful, efficient, reliable, and less expensive lasers are now available for a variety of cost-effective surface treatments, as outlined in Fig. 34.7.

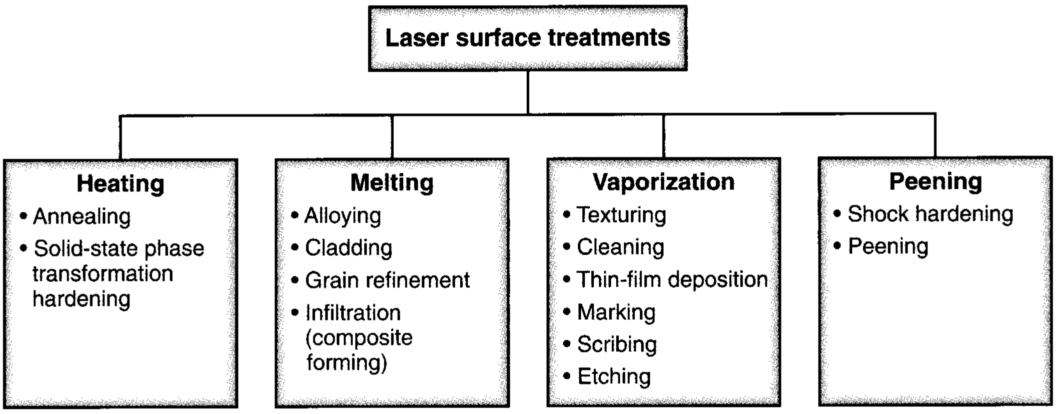


FIGURE 34.7 An outline of laser surface-engineering processes. *Source*: After N.B. Dahotre.

EXAMPLE 34.2 Applications of Laser Surface Engineering

In this example, several applications of lasers in engineering practice are given. The most commonly used lasers are Nd:YAG and CO_2; excimer lasers are generally used for surface texturing (see also Table 27.2).

Localized surface hardening—Cast irons: diesel-engine cylinder liners, automobile steering assemblies, and camshafts. Carbon steels: gears and electromechanical parts.

Surface alloying—Alloy steels: bearing components. Stainless steels: diesel-engine valves and seat inserts. Tool and die steels: dies for forming and die casting.

Cladding—Alloy steels: automotive valves and valve seats. Superalloys: turbine blades.

Ceramic coating—Aluminum–silicon alloys: automotive-engine bore.

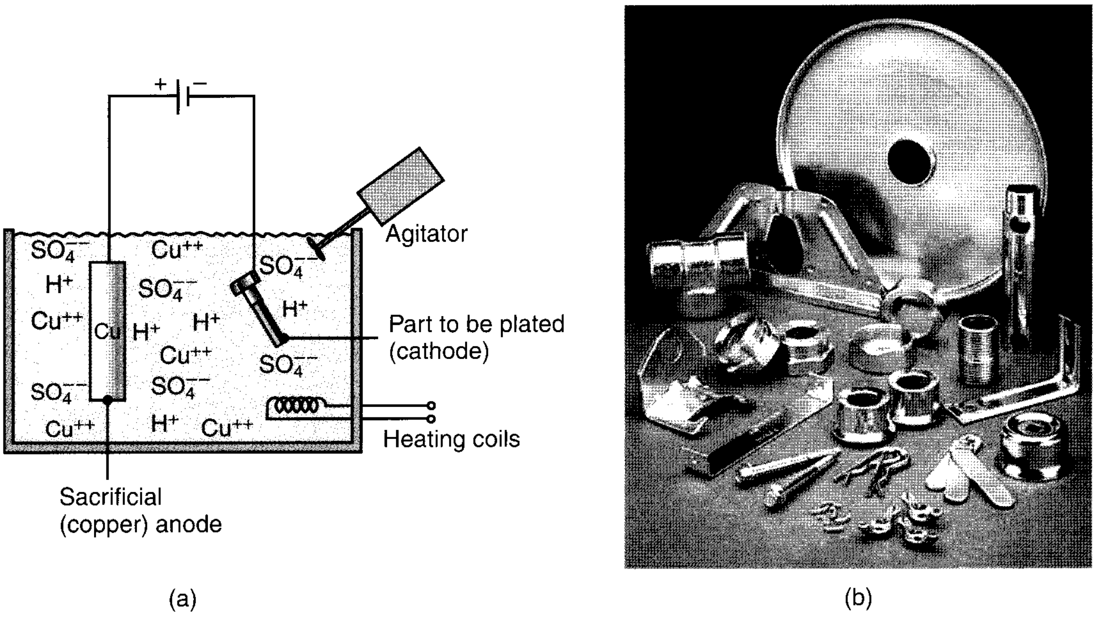
Surface texturing—Metals, plastics, ceramics, and wood: all types of products.

34.9 Electroplating, Electroless Plating, and Electroforming

Plating, like other coating processes, imparts the properties of resistance to wear, resistance to corrosion, high electrical conductivity, and better appearance and reflectivity, as well as similar desirable properties.

Electroplating. In electroplating, the workpiece (cathode) is plated with a different metal (anode), which is transferred through a water-based electrolytic solution (Fig. 34.8). Although the plating process involves a number of reactions, the process consists basically of the following sequence:

1. The metal ions from the anode are discharged by means of the potential energy from the external source of electricity, or are delivered in the form of metal salts.



(a) (b)

FIGURE 34.8 (a) Schematic illustration of the electroplating process. (b) Examples of electroplated parts. *Source*: Courtesy of BFG Electroplating.

2. The metal ions are dissolved into the solution.
3. The metal ions are deposited on 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 plate metal, the electrolyte, and the efficiency of the system; typically, c is in the range from 0.03 to 0.1 mm^3/amp-s. Note that, for the same volume of material deposited, the larger the workpiece surface plated, the thinner is the layer. The time required for elecroplating is usually long, because the deposition rate is typically on the order of 75 μm/hour. 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 plated from the solution, it needs to be periodically replenished, and this is accomplished through two principal methods: salts of metals are occasionally added to the solution, or 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 main forms of electroplating:

1. In **rack plating**, the parts to be plated are placed in a rack, which is then conveyed through a series of process tanks.
2. In **barrel plating**, small parts are placed inside a permeable barrel, which is placed inside the process tank(s). This form of electroplating is commonly performed with small parts, such as bolts, nuts, gears, and fittings. Electrolytic fluid can penetrate through the barrel and provide the metal for plating, and electrical contact is provided through the barrel and through contact with other parts.
3. In **brush processing**, the electrolytic fluid is pumped through a handheld brush with metal bristles. The workpiece can be very large in this circumstance, and the process is suitable for field repair or plating and can be used to apply coatings on large equipment without disassembly.

Simple electroplating can be achieved in a single-process bath or tank, but more commonly, a sequence of operations is used in a plating line. For example, the following tanks and processes may be part of an electroplating operation:

- Chemical cleaning and degreasing tanks will be used to remove surface contaminants and enhance surface adhesion of the plated coating.
- The workpieces may be exposed to a strong acid bath (pickling solution) to reduce or eliminate the thickness of the oxide coating on the workpiece.
- A base coating may be applied. This may involve the same or a different metal than that of the ultimate surface. For example, if the desired metal coating will not adhere well to the substrate, an intermediate coating can be applied. Also, if thick films are desired, a plating tank can be used to quickly develop a film, and a subsequent tank with brightener additives in the electrolytic solution is used to develop the ultimate surface finish.
- A separate tank performs final electroplating.
- Rinse tanks will be used throughout the sequence.

Rinse tanks are necessary for a number of reasons. Some plating is performed with cyanide salts delivering the required metal ions. If any residue acid (such as that from a pickling tank) is conveyed to the cyanide-solution tank, poisonous hydrogen-cyanide gas is exhausted. (This is a significant safety concern, and environmental

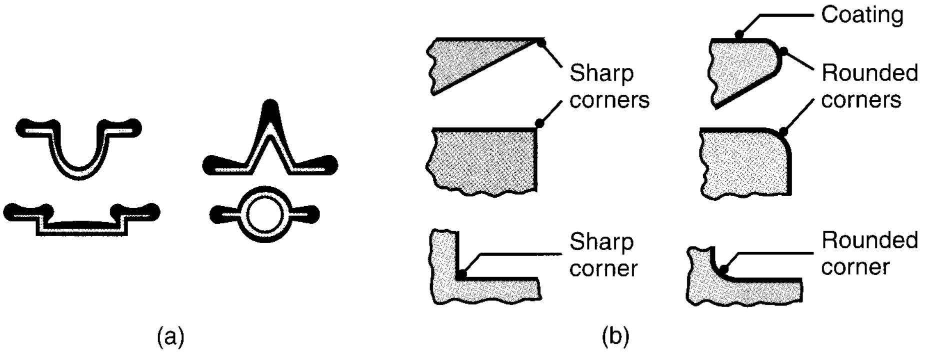


FIGURE 34.9 (a) Schematic illustration of nonuniform coatings (exaggerated) in electroplated parts. (b) Design guidelines for electroplating. Note that sharp external and internal corners should be avoided for uniform plating thickness.

controls are essential in plating facilities.) Also, residue plating solution will contain some metal ions, and it is often desirable to recover those ions by capturing them in a rinse tank.

The rate of film deposition depends on the local current density and is not necessarily uniform on a part. Workpieces with complex shapes may require an altered geometry because of varying plating thicknesses, as shown in Fig. 34.9.

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** is done by 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. 2.14) and a thickness of about 0.05 mm or more. 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. It is also used to rebuild worn parts.

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 gold, silver, and platinum are important electroplating materials in the electronics and jewelry industries for electrical contact and decorative purposes, respectively.

Plastics (such as ABS, polypropylene, polysulfone, polycarbonate, polyester, and nylon) also can be electroplating substrates. Because they are not electrically conductive, plastics must 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 and without the use of 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 workpiece. 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

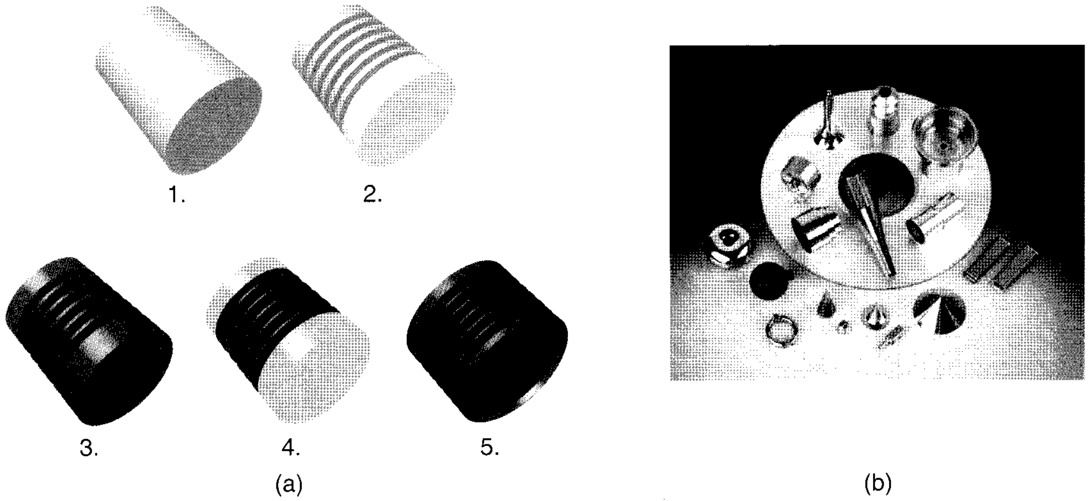


FIGURE 34.10 (a) Typical sequence in electroforming. (1) A mandrel is selected with the correct nominal size. (2) The desired geometry (in this case, that of a bellows) is machined into the mandrel. (3) The desired metal is electroplated onto the mandrel. (4) The plated material is trimmed if necessary. (5) The mandrel is dissolved through chemical machining. (b) A collection of electroformed parts. *Source*: Courtesy of Servometer®, Cedar Grove, NJ.

and ceramics. The process is more expensive than electroplating. However, unlike that of electroplating, the coating thickness of electroless plating is always uniform.

Electroforming. A variation of electroplating, electroforming actually is a metal-fabricating process. Metal is electrodeposited on a *mandrel* (also called a *mold* or a *matrix*), which is then removed; thus, the coating itself becomes the product (Fig. 34.10). Both simple and complex shapes can be produced by electroforming, with wall thicknesses as small as 0.025 mm. Parts may weigh from a few grams to as much as 270 kg. Production rates can be increased through the use of multiple mandrels.

Mandrels are made from a variety of materials: metallic (zinc or aluminum) or nonmetallic (which can be made electrically conductive with the proper coatings). Mandrels should be able to be removed physically without damaging the electroformed part. They also may be made of low-melting alloys, wax, or plastics, all of which can be melted away or dissolved with suitable chemicals.

The electroforming 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.

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. Various metals (particularly steel, aluminum, and zinc) can be conversion coated. Oxides that naturally form on their surfaces represent a form of conversion coating.

Phosphates, chromates, and *oxalates* are used to produce these coatings, for purposes such 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). The two common methods of coating are *immersion* and *spraying*. The equipment required depends on the method of application, the type of product, and quality considerations.

Anodizing. This is an oxidation process (*anodic oxidation*) in which the workpiece surfaces are converted to a hard and porous oxide layer that provides corrosion resistance and a decorative finish. The workpiece is the anode in an electrolytic cell immersed in an acid bath, which results in chemical adsorption of oxygen from the bath. Organic dyes of various colors (usually black, red, bronze, gold, or gray) can be used to produce stable, durable surface films. Typical applications for anodizing are aluminum furniture and utensils, architectural shapes, automobile trim, picture frames, keys, and sporting goods. 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 chemical compounds such as oxides, chromates, and phosphates. An example is the *blackening* of iron and steels, a process that utilizes solutions of hot, caustic soda and results in chemical reactions that produce a lustrous, black oxide film on surfaces.

34.11 Hot Dipping

In *hot dipping*, the workpiece (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 many other products.

A typical continuous *hot-dipped galvanizing* line for sheet steel is shown in Fig. 34.11. The rolled sheet is first cleaned electrolytically and 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* (similar to air-drying in car washes). Proper draining for the removal of excess coating materials is important.

The coating thickness is usually given in terms of coating weight per unit surface area of the sheet, typically 150 to 900 g/m^2. The service life depends on the thickness of the zinc coating and the environment to which it is exposed. Various **precoated sheet steels** are used extensively in automobile bodies. Proper draining to remove excess coating materials is an important consideration.

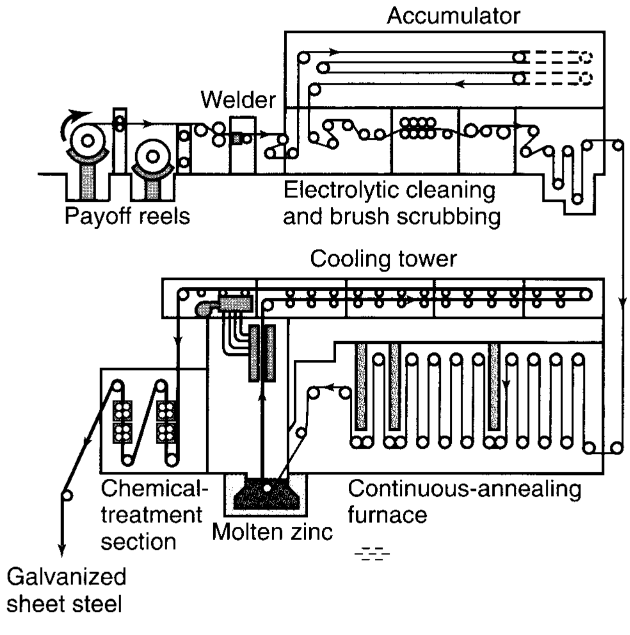


FIGURE 34.11 Flow line for the continuous hot-dipped galvanizing of sheet steel. The welder (upper left) is used to weld the ends of coils to maintain continuous material flow. *Source*: Courtesy of the American Iron and Steel Institute.

34.12 Porcelain Enameling; Ceramic and Organic Coatings

Metals can be coated with a variety of glassy (*vitreous*) coatings to provide corrosion and electrical resistance, and 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 by heating both of them at 425° to 1000°C to liquefy the oxides. The coating may be applied by dipping, spraying, or electrodeposition, and thicknesses are usually from 0.05 to 0.6 mm. The viscosity of the material can be controlled using binders so that the coating adheres 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. Porcelain enamels also are used as protective coatings on jet-engine components. Metals coated are typically steels, cast iron, and aluminum. Glasses are used as a lining (for chemical resistance) where the thickness of the glass is much greater than that of the 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. Materials such as powders of hard metals, aluminum oxide, and zirconium oxide are applied to a substrate at room temperature by means of binders. These coatings act as thermal barriers and have been applied (usually by thermal spraying techniques) to hot-extrusion dies, turbine blades, diesel-engine components, and nozzles for rocket motors to extend the life of these components. They also are used for electrical-resistance applications to withstand repeated arcing. Plasma arcs are used where temperatures may reach 15,000°C, which is much higher than those obtained by flames.

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 the coil stock on continuous lines (see Fig. 13.10), with thicknesses generally from 0.0025 to 0.2 mm. Such coatings have a wide range of properties: flexibility, durability, hardness, resistance to abrasion and chemicals, color, texture, and gloss. Coated sheet metal is subsequently formed into various products, such as TV cabinets, appliance housings, paneling, shelving, residential-building siding, gutters, and metal furniture.

More critical applications involve, for example, the protection of naval aircraft, which are subjected to high humidity, rain, seawater, pollutants (such as those from ship exhaust stacks), aviation fuel, deicing fluids, and battery acid, as well as being impacted by particles such as dust, gravel, stones, and deicing salts. For aluminum structures, organic coatings consist typically of an epoxy primer and a polyurethane topcoat with a lifetime of four to six years. Primer performance is an important factor in the durability of the coating.

EXAMPLE 34.3 Ceramic Coatings for High-temperature Applications

Certain product characteristics (such as wear resistance and thermal and electrical insulation—particularly at elevated temperatures) can be imparted through ceramic coatings, rather than to the base metals or materials themselves. Selecting materials with such bulk properties can be expensive and may not meet the structural strength requirements of a particular application.

For example, a wear-resistance component does not have to be made completely from a wear-resistant material, since the properties of only a thin layer of its surface are relevant to wear. Consequently, coatings have important 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—as is done in multiple-layer coated cutting tools (see Fig. 22.8).

TABLE 34.1

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), zironium 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

34.13 Diamond Coating and Diamond-like Carbon

The properties of *diamond* that are relevant to manufacturing engineering were described in Section 8.7. Important advances have been made in the diamond coating of metals, glass, ceramics, and plastics using various techniques, such as chemical vapor deposition (CVD), 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); sunglasses; 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; diamond-coated speakers for stereo systems; turbine blades; and fuel-injection nozzles.

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, unlike the hazy gray diamond film formerly produced. This film is then laser cut to desired shapes and brazed onto cutting tools (for example).

The development of these techniques, combined with the important properties of diamond (hardness, wear resistance, high thermal conductivity, and transparency to ultraviolet light and microwave frequencies), has enabled the production of various aerospace and electronic parts and components.

Studies also are continuing regarding the growth of diamond films on crystalline-copper substrate by the implantation of carbon ions. An important application is in making computer chips (see Chapter 28). Diamond can be doped to form *p*- and *n*-type ends on semiconductors to make transistors, and its high thermal conductivity allows the closer packing of chips than would be possible with silicon or gallium-arsenide chips, significantly increasing the speed of computers. Diamond is also an attractive material for future MEMS devices (see Chapter 29), because of its favorable friction and wear characteristics.

Diamond-like Carbon. *Diamond-like 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. Less expensive than diamond films, but with similar properties (such as low friction, high hardness, and chemical inertness, as well as having a smooth surface), DLC has applications in such areas as tools and dies, gears, engine components, bearings, MEMS devices, and microscale probes. As a coating on cutting tools, DLC has a hardness of about 5000 HV (compared with about double that for diamond).

34.14 Surface Texturing

As stated throughout the preceding chapters, each manufacturing process (such as casting, forging, powder metallurgy, injection molding, machining, grinding, polishing, electrical-discharge machining, grit blasting, and wire brushing) produces a certain surface texture and appearance. Obviously, some of these processes can be used to modify the surface produced by a previous process—for example, grinding the surface of a cast part. However, manufactured surfaces can be modified further by secondary operations for technical, functional, optical, or aesthetic reasons.

Called *surface texturing*, these additional processes generally consist of the following techniques:

- **Etching:** Using chemicals or sputtering techniques.
- **Electric arcs.**
- **Lasers:** Using excimer lasers with pulsed beams; applications include molds for permanent-mold casting, rolls for temper mills, golf-club heads, and computer hard disks.
- **Atomic oxygen:** Reacting with surfaces to produce a fine, cone-like surface texture.

The possible adverse effects of these processes on material properties and the performance of the textured parts are important considerations.

34.15 Painting

Because of its decorative and functional properties (such as environmental protection, low cost, relative ease of application, and the range of available colors), *paint* has been widely used as a surface coating. The engineering applications of painting range from appliances and machine tools to automobile bodies and aircraft fuselages. Paints generally are classified as

- **Enamels:** Produce a smooth coat with a glossy or semiglossy appearance.
- **Lacquers:** Form a 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 than the first two types.

Paints are available with good resistance to abrasion, temperature extremes, and fading; are easy to apply; and dry quickly. The selection of a particular paint

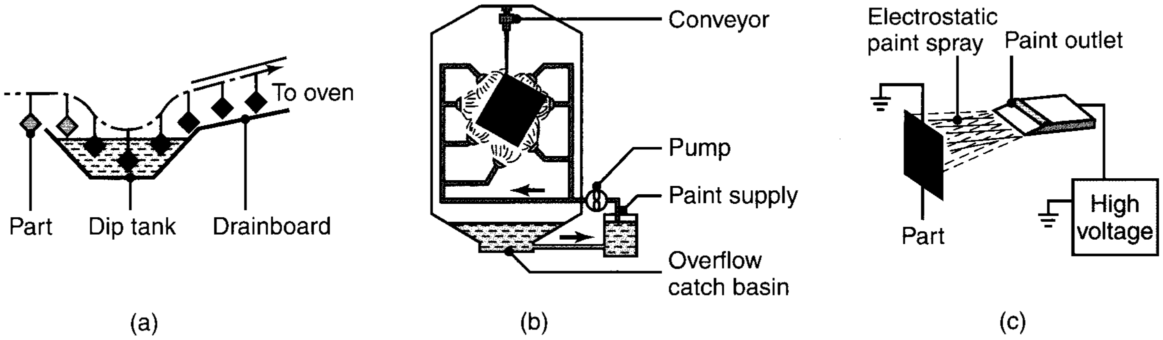


FIGURE 34.12 Methods of paint application: (a) dip coating, (b) flow coating, and (c) electrostatic spraying (used particularly for automotive bodies).

depends on specific requirements. Among these are resistance to mechanical actions (abrasion, marring, impact, and flexing) and to chemical reactions (acids, solvents, detergents, alkalis, 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 to be painted, producing a uniformly adherent coating. Unlike paint losses in conventional spraying, which may be as much as 70% of the paint, the loss can be as little as 10% in electrostatic spraying. However, deep recesses and corners can be difficult to coat with this method. The use of *robotic controls* for guiding the spray nozzles is common (see Section 37.6.3).

34.16 Cleaning of Surfaces

The importance of surfaces in manufacturing and the effects of deposited or adsorbed layers of various elements and contaminants on surface characteristics have been stressed throughout this text. A clean surface can have both beneficial and detrimental effects. Although a surface that is not clean may reduce the tendency for adhesion and galling, cleanliness generally is essential for a more effective application of coatings, painting, adhesive bonding, welding, brazing, and soldering, as well as for the reliable functioning of manufactured parts in machinery, assembly operations, and food and beverage containers.

Cleaning involves the removal of solid, semisolid, or liquid contaminants from a surface and is an important part of manufacturing operations and the economics of production. The word *clean* or the degree of cleanliness of a surface is somewhat difficult to define. Two simple and common tests are as follows:

1. Wiping the surface of, say, a dinner plate with a clean cloth and observing any residues on the cloth.
2. Observing whether water continuously coats the surface of a plate (the *waterbreak test*). If water collects as individual droplets, the surface is not clean. (You can test this phenomenon by wetting dinner plates that have been cleaned to different degrees.)

The type of cleaning process required depends on the type of *metalworking-fluid residues* and *contaminants* to be removed. For example, water-based fluids are easier and less expensive to remove than oil-based fluids. Contaminants (also called *soils*) may consist of rust, scale, chips (and other 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 disturbing the contaminants, often with wire or fiber brushing, abrasive blasting (jets), tumbling, or steam jets. Many of these processes are particularly effective in removing rust, scale, and other solid contaminants. **Ultrasonic cleaning** is also placed into this category.

Electrolytic Cleaning. In this process, a charge is applied to the part to be cleaned in an aqueous (often alkaline) cleaning solution. The charge results in bubbles of hydrogen or oxygen (depending on polarity) being released at the part's surface. The bubbles are abrasive and aid in the removal of contaminants.

Chemical Cleaning. Chemical cleaning usually involves the removal of oil and grease from surfaces. The operation consists of one or more of the following processes:

- **Solution:** The soil dissolves in the cleaning solution.
- **Saponification:** A chemical reaction converts animal or vegetable oils into a soap that is soluble in water.
- **Emulsification:** The cleaning solution reacts with the soil or lubricant deposits 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 surface-active elements in the cleaning solution.
- **Aggregation:** Lubricants are removed from the surface by various agents in the cleanser and are collected as large dirt particles.

Cleaning Fluids. Common cleaning fluids used in conjunction with electrochemical processes for more effective cleaning include the following:

- **Alkaline solutions:** A complex combination of water-soluble chemicals, alkaline solutions are the least expensive and most widely used cleaning fluids in manufacturing operations. 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 generally consist of kerosene and oil-in-water and various types of emulsifiers.
- **Solvents:** Typically petroleum solvents, chlorinated hydrocarbons, and mineral spirits, solvents generally are used for short runs. Fire and toxicity are major hazards.
- **Hot vapors:** Chlorinated solvents can be used to remove oil, grease, and wax. The solvent is boiled in a container and then condensed. This hot-vapor process is simple, and the cleaned parts are dry.
- **Acids, salts,** and **mixtures of organic compounds:** These are 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 design guidelines are as follows:

- Avoid deep, blind holes.
- Make several smaller components instead of one large component, which may be difficult to clean.
- Provide 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 processes. They are used to impart specific chemical, physical, and mechanical properties, such as appearance, and corrosion, friction, wear, and fatigue resistance. Several techniques are available for modifying surfaces.
- The processes used include mechanical working and surface treatments, such as heat treatment, deposition, and plating. Surface coatings include enamels, nonmetallic materials, and paints.
- Clean surfaces can be important in the further processing (e.g., coating, painting, or welding) and use of the product. Cleaning can have a significant economic impact on manufacturing operations. Various mechanical and chemical cleaning methods may be utilized.

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
Vacuum evaporation
Vapor deposition
Waterbreak test
Water-jet peening

BIBLIOGRAPHY

ASM Handbook, Vol. 5, **Surface Engineering**, ASM International, 1994.

Bhushan, B., and Gupta, B.K., **Handbook of Tribology: Materials, Coatings, and Surface Treatments**, McGraw-Hill, 1991.

Bunshah, R.F. (ed.), **Handbook of Hard Coatings: Deposition Technologies, Properties and Applications**, Noyes, 2001.

Burakowski, T., and Wierschon, T., **Surface Engineering of Metals: Principles, Equipment, Technologies**, CRC Press, 1998.

Davis, J.R. (ed.), **Surface Engineering for Corrosion and Wear Resistance**, IOM Communications and ASM International, 2001.

——— (ed.), **Handbook of Thermal Spray Technology**, Thermal Spray Society and ASM International, 2004.

——— (ed.), **Surface Hardening of Steels**, ASM International, 2002.

Edwards, J., **Coating and Surface Treatment Systems for Metals: A Comprehensive Guide to Selection**, ASM International, 1997.

Handbook of Surface Treatments and Coatings, ASME Press, 2003.

Holmberg, K., and Matthews, A., **Coating Tribology: Properties, Techniques, and Applications**, Elsevier, 1994.

Inagaki, N., **Plasma Surface Modification and Plasma Polymerization**, Technomic, 1996.

Kawai, S. (ed.), **Anodizing and Coloring of Aluminum Alloys**, Metal Finishing Information Services and ASM International, 2002.

Lambourne, R., and Strivens, T.A., **Paint and Surface Coatings: Theory and Practice**, 2nd ed., Woodhead, 1999.

Liberto, N., **User's Guide to Powder Coating**, 4th ed., Society of Manufacturing Engineers, 2003.

Lindsay, J.H. (ed.), **Coatings and Coating Processes for Metals**, ASM International, 1998.

Mattox, D.M., **Handbook of Physical Vapor Deposition (PVD) Processing**, Noyes, 1999.

Park, J., **Chemical Vapor Deposition**, ASM International, 2001.

Prelas, M.A., Popovichi, G., and Bigelow, L.K. (eds.), **Handbook of Industrial Diamonds and Diamond Films**, Marcel Dekker, 1998.

Schlesinger, M., and Paunovic, M. (eds.) **Modern Electroplating**, 4th ed., Wiley, 2001.

Stern, K.H. (ed.), **Metallurgical and Ceramic Protective Coatings**, Chapman & Hall, 1997.

Sudarshan, T.S. (ed.), **Surface Modification Technologies**, ASM International, 1998.

Tamarin, Y., **Protective Coatings for Turbine Blades**, ASM International, 2002.

Tracton, A.A. (ed.), **Coatings Technology Handbook**, CRC, 2006.

Utech, B., **A Guide to High Performance Coating**, Society of Manufacturing Engineers, 2002.

van Ooij, W.J., Bierwagen, G.P., Skerry, B.S., and Mills, D., **Corrosion Control of Metals by Organic Coatings**, CRC Press, 1999.

REVIEW QUESTIONS

34.1. Explain why surface treatments may be necessary for various parts made by one or more processes.

34.2. What are the advantages of roller burnishing?

34.3. Explain the difference between case hardening and hard facing.

34.4. Describe the principles of physical and chemical vapor deposition. What applications do these processes have?

34.5. What is the principle of electroforming? What are the advantages of electroforming?

34.6. Explain the difference between electroplating and electroless plating.

34.7. How is hot dipping performed?

34.8. What is an air knife? How does it function?

34.9. Describe the common painting systems presently in use in industry.

34.10. What is a conversion coating? Why is it so called?

34.11. Describe the difference between thermal spraying and plasma spraying.

34.12. What is cladding, and why is it performed?

QUALITATIVE PROBLEMS

34.13. Describe how roller-burnishing processes induce compressive residual stresses on the surfaces of parts.

34.14. Explain why some parts may be coated with ceramics. Give some examples.

34.15. Give examples of part designs that are suitable for hot-dip galvanizing.

34.16. Comment on your observations regarding Fig. 34.9.

34.17. 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.

34.18. Make a list of the coating processes described in this chapter and classify them in relative terms as "thick" or "thin."

34.19. Why is galvanizing important for automotive-body sheet metals?

34.20. Explain the principles involved in various techniques for applying paints.

QUANTITATIVE PROBLEMS

34.21. Taking a simple example, such as the parts shown in Fig. 34.1, estimate the force required for roller burnishing. (*Hint:* See Sections 2.6 and 14.4.)

▶**34.22.** Estimate the plating thickness in electroplating a 20-mm solid-metal ball using a current of 10 A and a plating time of 1.5 hours. Assume that $c = 0.08$ in Eq. (34.1).

SYNTHESIS, DESIGN, AND PROJECTS

34.23. Which surface treatments are functional, and which are decorative? Are there any treatments that serve both functions? Explain.

34.24. 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.

34.25. 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.

34.26. It is known that a mirror-like surface finish can be obtained by plating workpieces that are ground; that is, the surface finish improves after coating. Explain how this occurs.

34.27. 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.

34.28. Outline the reasons that the topics described in this chapter are important in manufacturing processes and operations.

34.29. 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?

34.30. 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.

34.31. 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.

34.32. 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.)

34.33. 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.

34.34. 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.

Engineering Metrology, Instrumentation, and Quality Assurance

PART VIII

The preceding chapters have described the techniques used to modify the surfaces of components and products to obtain certain desirable properties, discussing the advantages and limitations of each technique along the way. Although dimensional accuracies obtained in individual manufacturing processes were described, we have not yet described how parts are *measured* and *inspected* before they are assembled into products.

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. The vast majority of manufactured parts are components or a subassembly of a product, and they must fit and be assembled properly so that the product performs its intended function during its service life. For example, (a) a piston should fit into a cylinder within specified tolerances, (b) a turbine blade should fit properly into its slot on a turbine disk, and (c) the slideways of a machine tool must be produced with a certain accuracy so that the parts produced on that machine are accurate within their desired specifications.

Measurement of the relevant 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 a machine is worn and has to be replaced, all one has to do is purchase a similar one with the same specification or part number. The same is now done with all products, ranging from bolts and nuts, to gears, to electric motors.

The first of the next two chapters describes the principles involved in, and the various instruments and modern machines used for, measuring dimensional features such as length, angle, flatness, and roundness. Testing and inspecting parts are important aspects of manufacturing operations; thus, the methods used for the nondestructive and destructive testing of parts are also described.

One of the most important aspects of manufacturing is **product quality**. Chapter 36 discusses the technological and economic importance of *building quality into a product* rather than inspecting the product after it is made, as has been done traditionally. This concept is even more significant in view of competitive manufacturing in a global economy.

CHAPTER 35 Engineering Metrology and Instrumentation

EXAMPLES:

- This chapter describes the importance of the measurement of manufactured parts, noting that measurement of parts and their certification to a certain standard is essential to ensuring part fit and thus proper operation.
- A wide variety of measurement strategies, gages, equipment, and machines has been developed, as described in this chapter.
- The topics discussed include traditional measurement with simple rulers; gages and instruments, such as micrometers and calipers; and digital equipment and computer-controlled equipment, such as coordinate-measurement machines.
- The chapter describes features of measuring instruments and the importance of automated measurements, ending with an introduction to the principles of dimensioning and tolerancing.

35.1 Introduction

This chapter presents the principal methods of measurement and the characteristics of the instruments used in manufacturing. **Engineering metrology** is defined as the measurement of dimensions such as length, thickness, diameter, taper, angle, flatness, and profile. Consider, for example, the slideways for machine tools (Fig. 35.1); these components must have specific dimensions, angles, and flatness in order for the machine to function properly and with the desired dimensional accuracy.

Traditionally, measurements have been made *after* the part has been produced—an approach known as **postprocess inspection.** Here, the term *inspection* means "checking the dimensions of what has been produced or is being produced and determining whether those dimensions comply with the specified dimensional tolerances and other specifications." Today, however, measurements are being made while the part is being produced on the machine—an approach known as **in-process, online,** or **real-time inspection.**

An important aspect of metrology in manufacturing processes is the **dimensional tolerance** (i.e., the permissible variation in the dimensions of a part). Tolerances are important because of their impact on the proper functioning of a product, part interchangeability, and manufacturing costs. Generally, the smaller the tolerance, the higher are the production costs. The chapter ends with a discussion of dimensional limits and fits used in engineering practice.

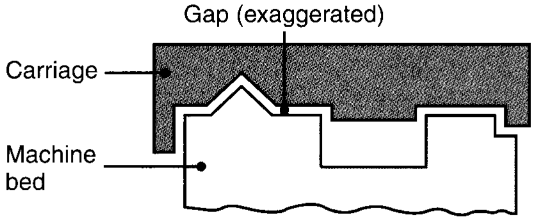


FIGURE 35.1 Cross section of a machine-tool slideway; see also Fig. 23.2. The width, depth, angles, and other dimensions all must be produced and measured accurately for the machine tool to function as expected.

35.2 Measurement Standards

Our earliest experience with measurement is usually with a simple *ruler* to measure lengths (linear dimensions). Rulers are used as a *standard* against which dimensions are measured. Traditionally, in English-speaking countries, the units *inch* and *foot* have been used, which originally were based on parts of the human body. Consequently, it was common to find significant variations in the length of 1 foot.

In most of the world, however, the *meter* has been used as a length standard. Originally, 1 meter was defined as one ten-millionth of the distance between the North Pole and the equator. The original meter length subsequently 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 officially was 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 a unit of length in the Système International d'Unités (SI) and is the international standard.

Numerous measuring instruments and devices are used in engineering metrology, each of which has its own application, resolution, precision, and other features. Two terms commonly used to describe the type and quality of an instrument are as follows:

1. **Resolution** is 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, is the degree to which the 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 by being held by the hand.

In engineering metrology, the words **instrument** and **gage** often are used interchangeably. 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. In the interest of accuracy, measurements should be taken in controlled environments maintaining the standard temperature, usually within ±0.3°C.

EXAMPLE 35.1 Length Measurements throughout History

Many standards for length measurement have been developed during the past 6000 years. A common standard in Egypt around 4000 B.C. was the King's elbow, which was equivalent to 0.4633 m. One *elbow* was equal to 1.5 feet (or 2 hand spans, 6 hand widths, or 24 finger thicknesses). In 1101 A.D., King Henry I declared a new standard called the *yard* (0.9144 m), which was the distance from his nose to the tip of his thumb.

During the Middle Ages, almost every kingdom and city established its own length standard—some with identical names. In 1528, the French physician

Jean Fernel proposed the distance between Paris and Amiens (a city 120 km north of Paris) as a general length reference. During the 17th century, some scientists suggested that the length of a certain pendulum be used as a standard. In 1661, British architect Sir Christopher Wren suggested that a pendulum with a period of one-half second be used. The Dutch mathematician Christian Huygens proposed a pendulum that had a length one-third of Wren's and a period of 1 second.

To put an end to the confusion of length measurement, a definitive length standard began to be developed in 1790 in France with the concept of a *mètre* (from the Greek word *metron*, meaning "measure"). A gage block 1 meter long was made of pure platinum with a rectangular cross section and was placed in the National Archives in Paris in 1799. Copies of this gage were made for other countries over the years.

During the three years from 1870 to 1872, international committees met and decided on an international meter standard. The new bar was made of 90% platinum and 10% iridium, with an x-shaped cross section and overall dimensions of 20 × 20 mm. Three marks were engraved at each end of the bar. The standard meter is the distance between the central marks at each end, measured at 0°C. Today, extremely accurate measurement is based on the speed of light in a vacuum, which is calculated by multiplying the wavelength of the standardized infrared beam of a laser by its frequency.

35.3 Geometric Features of Parts; Analog and Digital Measurements

In this section, we list the most common quantities and geometric features that typically are measured in engineering practice and in products made by the manufacturing processes described throughout this book:

- *Length*—including all linear dimensions of parts.
- *Diameter*—outside and inside, including parts with different outside and inside diameters (steps) along their length.
- *Roundness*—including out-of-roundness, concentricity, and eccentricity.
- *Depth*—such as that of drilled or bored holes and cavities in dies and molds
- *Straightness*—such as that of shafts, bars, and tubing.
- *Flatness*—such as that of machined and ground surfaces.
- *Parallelism*—such as that of two shafts or slideways in machines.
- *Perpendicularity*—such as that of a threaded bar inserted into a flat plate.
- *Angles*—including internal and external angles.
- *Profile*—such as curvatures in castings, in forgings, and on car bodies.

A wide variety of instruments and machines is available to accurately and rapidly measure the preceding quantities on stationary parts or on parts that are in continuous production. Because of major and continuing trends in automation and the computer control of manufacturing operations, modern measuring equipment is now an *integral* part of production machines. The implementation of digital instrumentation and developments in computer-integrated manufacturing (described in Part IX of the book) have together led to the total integration of measurement technologies within manufacturing systems.

It is important to recognize the advantages of *digital* over *analog* instruments. As will be obvious from our description of traditional measuring equipment in Section 35.4, accurate measurement on 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, a digital caliper does not require any particular skills, because measurements are indicated directly (Fig. 35.2b).

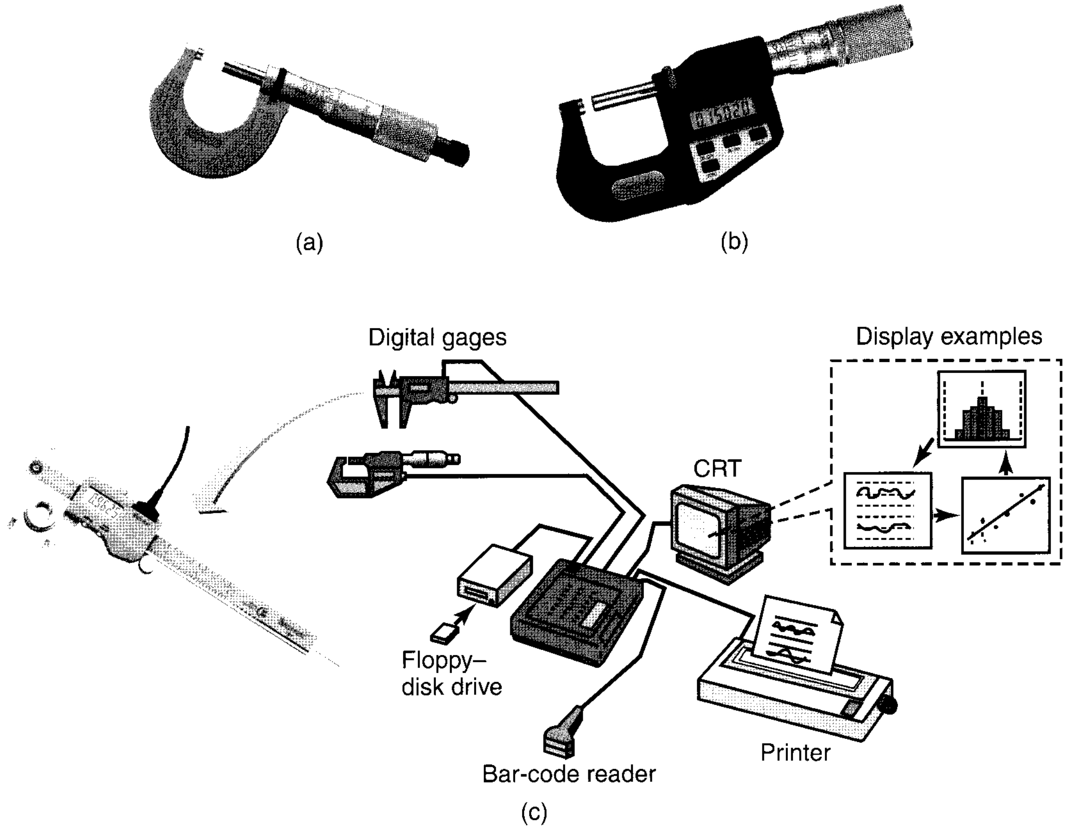


FIGURE 35.2 (a) A vernier (analog) micrometer. (b) A digital micrometer with a range of 0 to 1 in. (0 to 25 mm) and a resolution of 50 μin. (1.25 μm). Generally, it is much easier to read dimensions on this instrument than on analog micrometers. (c) Schematic illustration showing the integration of digital gages with miroprocessors for real-time data acquisition for statistical process control. *Source*: (a) Courtesy of L.C. Starrett Co. (b) Courtesy of Mitutoyo Corp.

More importantly, digital equipment can be integrated easily into other equipment (Fig. 35.2c), including production machinery and systems for statistical process control (SPC), as described in detail in Chapter 36.

35.4 Traditional Measuring Methods and Instruments

This section describes the characteristics of traditional measuring methods and instruments that have been used over many years and are still used extensively in many parts of the world. However, these instruments are rapidly being replaced with more efficient and advanced instruments and measuring machines, as described in Section 35.5.

35.4.1 Line-graduated Instruments

These instruments are used for measuring lengths or angles. *Graduated* means "marked to indicate a certain quantity."

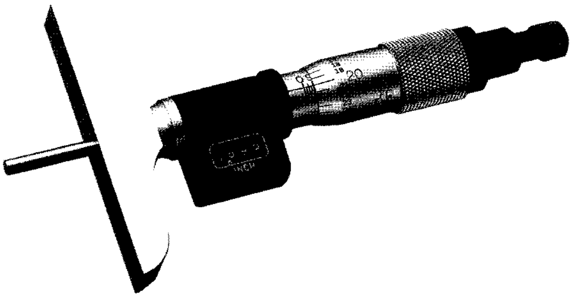


FIGURE 35.3 A digital micrometer depth gage. *Source*: Courtesy of Starrett Co.

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.
- **Calipers:** These instruments can be used to measure inside or outside lengths. Also called *caliper gages* and *vernier calipers* (named for P. Vernier, who lived in the 1600s), they have a graduated beam and a sliding jaw. *Digital calipers* are in increasingly wider use.
- **Micrometers:** These instruments are commonly used for measuring the thickness and inside or outside dimensions of parts. *Digital micrometers* are equipped with digital readouts (Fig. 35.2b) in metric or English units. Micrometers also are 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, threaded-rod diameters, and wall thicknesses of tubes and curved sheets.

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 rule. Because of the experience required to use them and their dependence on graduated scales, the accuracy of indirect-measurement tools is limited. *Telescoping gages* can be used for the 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, and the angle is read directly on the vernier scale. Another common type of bevel protractor is the *combination square*, which is a steel rule equipped with devices for measuring 45° and 90° angles.
- **Sine bar:** Measuring with this method involves placing the part on an inclined bar (sine bar) or plate and adjusting the angle by placing gage blocks on a surface plate. After the part is placed on the sine bar, a dial indicator is used to scan the top surface of the part. Gage blocks (see Section 35.4.4) are added or removed as necessary until the top surface is parallel to the surface plate. The angle on the part is then calculated from trigonometric relationships.
- **Surface plates:** These plates are used to place both parts to be measured and the measuring instruments. They typically are made of cast iron or natural stones (such as granite) and are used extensively in engineering metrology. Granite surface plates have the desirable properties of being resistant to corrosion, being nonmagnetic, and having 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 or deviations in the distance between two or more surfaces. These instruments, of which the most common example is a **dial indicator** (Fig. 35.4), compare dimensions

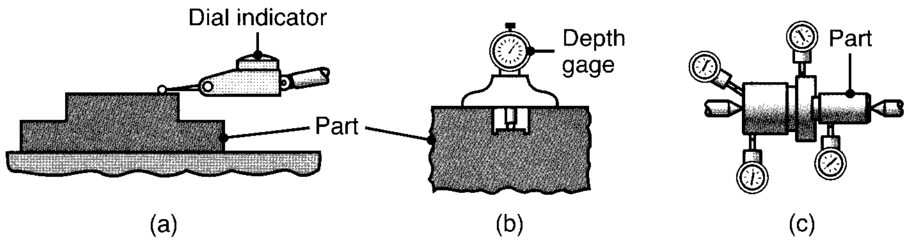


FIGURE 35.4 Three uses of dial indicators: to measure (a) roundness and (b) depth, and (c) for multiple-dimension gaging of a part.

(hence the word *comparative*). They are all simple mechanical devices that convert linear displacements of a pointer to the amount of rotation of an indicator on a circular dial. The indicator is set to zero at a certain reference surface, and 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 (as either plus or minus some number) to accuracies as high as 1 μm. Dial indicators with electrical and fluidic amplification mechanisms and with a digital readout also are available.

35.4.2 Measuring Geometric Features

Straightness. *Straightness* commonly can be checked with a straightedge or a dial indicator (Fig. 35.5). 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 now commonly used to align individual machine elements in the assembly of machine components.

Flatness. *Flatness* can be measured by mechanical means with a *surface plate* and a *dial indicator*. This method can be used to measure perpendicularity, which also can be measured by precision-steel squares.

Another method for measuring flatness is **interferometry**, which uses an *optical flat*. This device is a glass disk or fused-quartz disk with parallel flat surfaces that is placed on the surface of the workpiece (Fig. 35.6a). When a *monochromatic* light beam (a light beam with one wavelength) is aimed at the surface at an angle, the optical flat splits the light beam into two beams, appearing as light and dark bands to the naked eye (Fig. 35.6b). 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. 35.6c). Consequently, a truly flat workpiece surface (i.e., one in which the angle between the two surfaces is zero) will not split the light beam, and fringes will not appear. When surfaces are not flat, the fringes are curved (Fig. 35.6d). The interferometry method is also used for observing surface textures and scratches (Fig. 35.6e).

Diffraction gratings consist of two optical flat glasses of different lengths with closely spaced parallel lines scribed on their surfaces. The grating on the shorter glass is inclined slightly. As a result, *interference fringes* develop when it is 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 with the use of electronic counters and photoelectric sensors, a resolution of 2.5 μm can be obtained with gratings having 40 lines/mm.

Roundness. This feature usually is described as a deviation from true roundness (which, mathematically, is manifested in a circle). The term *out-of-roundness* (ovality) is actually more descriptive of the shape of the part (Fig. 35.7a) than the word

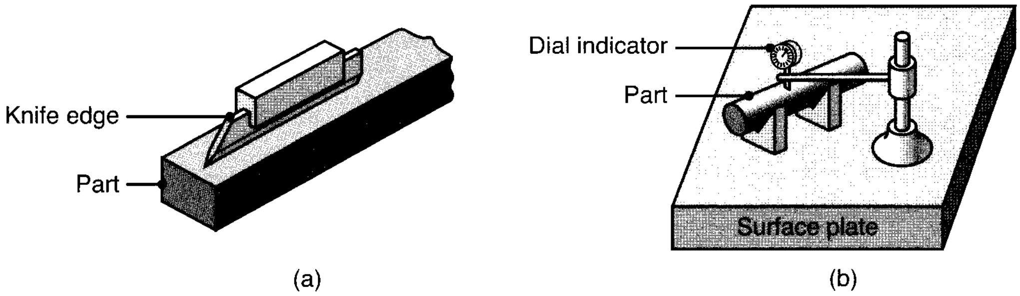


FIGURE 35.5 Measuring straightness manually with (a) a knife-edge rule and (b) a dial indicator. *Source*: After F.T. Farago.

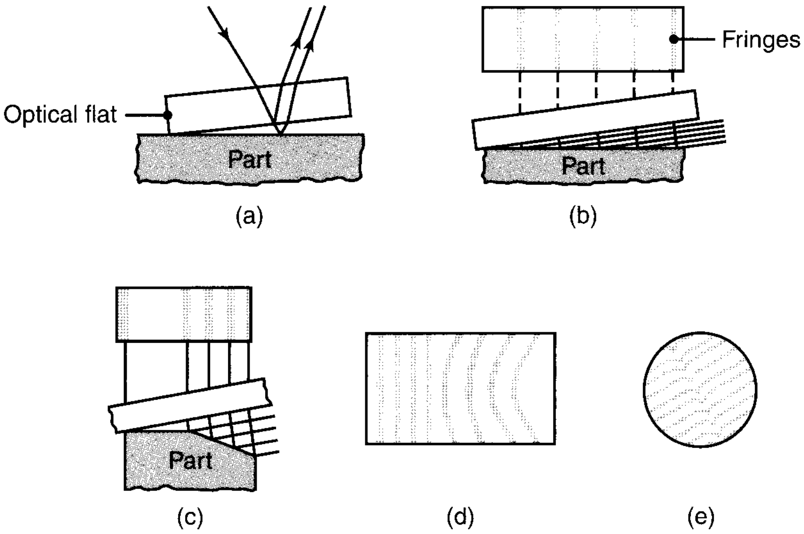


FIGURE 35.6 (a) Interferometry method for measuring flatness with an optical flat. (b) Fringes on a flat, inclined surface. An optical flat resting on a perfectly flat workpiece surface will not split the light beam, and no fringes will be present. (c) Fringes on a surface with two inclinations. *Note*: The greater the incline, the closer together are the fringes. (d) Curved fringe patterns indicate curvatures on the workpiece surface. (e) Fringe pattern indicating a scratch on the surface.

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 placed on a *V-block* or between centers (Figs. 35.7b and c, respectively) and is rotated while the point of a dial indicator is in contact with the part surface. After a full rotation of the workpiece, the difference between the maximum and minimum readings on the dial is noted. This difference is called the **total indicator reading** (TIR) or the **full indicator movement**. This method can also be used to measure the straightness (squareness) of end faces of shafts that are machined, such as the facing operation shown in Fig. 23.1e.
2. In *circular tracing*, the part is placed on a platform, and its roundness is measured by rotating the platform (Fig. 35.7d). Alternatively, the probe can be rotated around a stationary part to take the measurement.

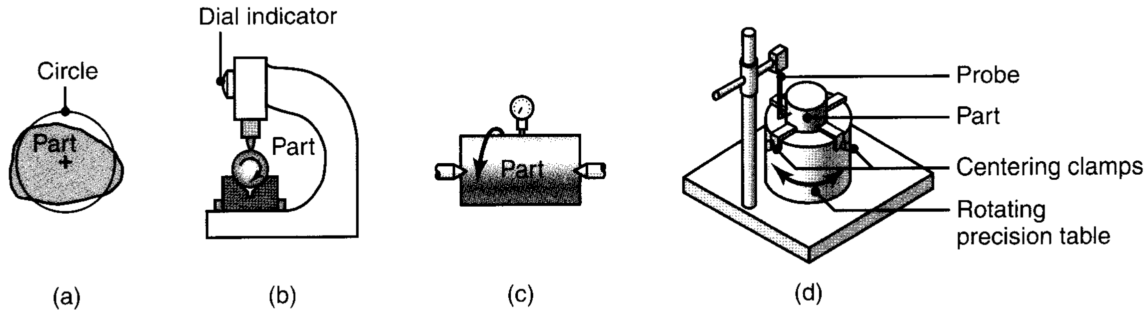


FIGURE 35.7 (a) Schematic illustration of out-of-roundness (exaggerated). Measuring roundness with (b) a V-block and dial indicator, (c) a round part supported on centers and rotated, and (d) circular tracing. *Source*: After F.T. Farago.

Profile. *Profile* may be measured by means such as (a) comparing the surface with a template or profile gage (as in the measurement of radii and fillets) for conformity and (b) using a number of dial indicators or similar instruments. The best method, however, is using the advanced measuring machines described in Section 35.5.

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 gages used 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. 35.8a), and (c) micrometers using pins or balls of various diameters (Fig. 35.8b). Advanced methods include the use of optical projectors and coordinate-measuring machines.

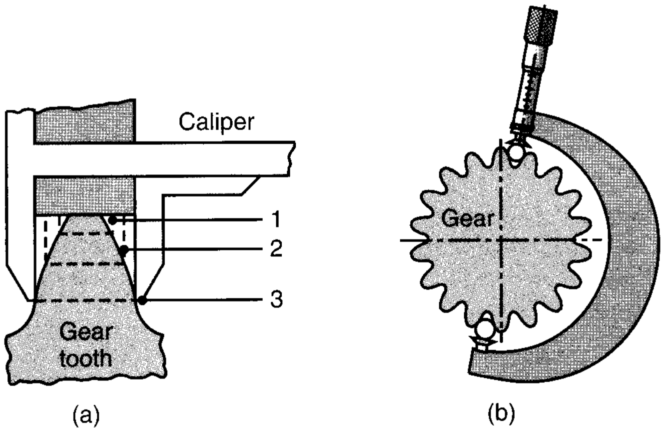


FIGURE 35.8 Measuring gear-tooth thickness and profile with (a) a gear-tooth caliper and (b) pins or balls and a micrometer. *Source*: Courtesy of American Gear Manufacturers Association.

35.4.3 Optical Contour Projectors

These instruments, also called **optical comparators**, were first developed in the 1940s to check the geometry of cutting tools for machining screw threads, but are now used for checking all profiles (Fig. 35.9). 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

This section describes several common gages that have simple solid shapes and cannot be classified as instruments although they are very valuable in metrology.

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 carefully. *Angle blocks* are made the same way and are used for angular gaging. Gage blocks have a flatness within 1.25 μm. Environmental temperature control is important when gages are used for high-precision measurements.

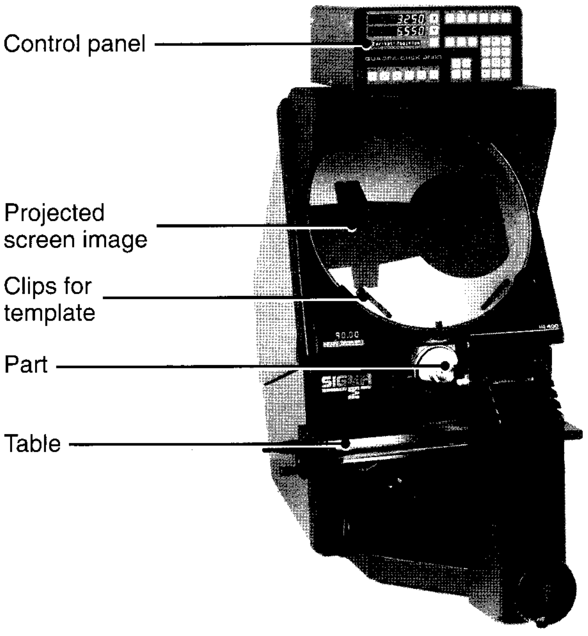


FIGURE 35.9 A bench-model horizontal-beam contour projector with a 16-in.-diameter (400-mm) screen with 150-W tungsten halogen illumination. *Source*: Courtesy of L.S. Starrett Company, Precision Optical Division.

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 compared with an established standard.

- **Plug gages** are commonly used for holes (Figs. 35.10a and b). The *GO gage* is smaller than the *NOT GO* (or *NO GO*) *gage* and slides into any hole that has a dimension smaller than the diameter of the gage. The *NOT GO* gage must not go into the 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 also are available, for measuring internal tapers (in which deviations between the gage and the part are indicated by the looseness of the gage), splines, and threads (in which the *GO* gage must screw into the threaded hole).
- **Ring gages** (Fig. 35.10c) are used to measure shafts and similar round parts. Ring thread gages are used to measure external threads. The *GO* and *NOT*

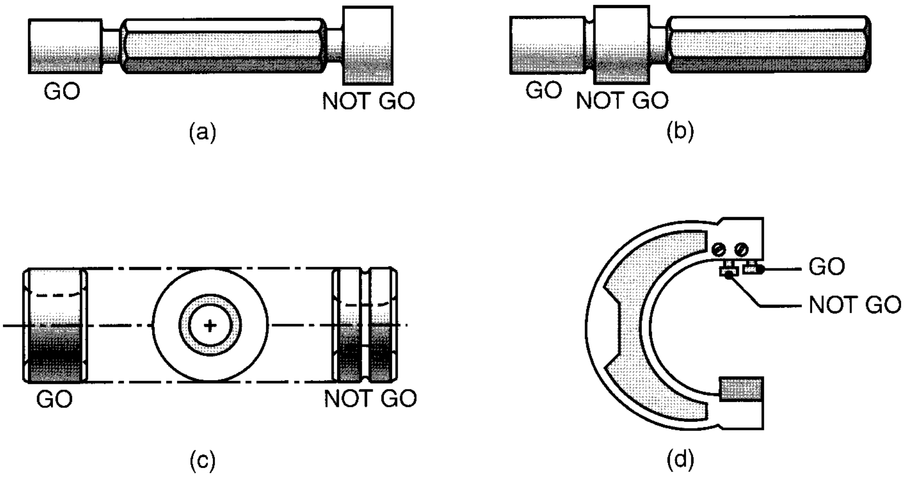


FIGURE 35.10 (a) Plug gage for holes, with *GO* and *NOT GO* on opposite ends of the gage. (b) Plug gage with *GO* and *NOT GO* on one end. (c) Plain ring gages for gaging round rods. Note the difference in knurled surfaces to identify the two gages. (d) Snap gage with adjustable anvils.

GO features on these gages are identified by the type of knurling on the outside diameters of the rings.

- **Snap gages** (Fig. 35.10d) commonly are 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. 35.11a. 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 sensed

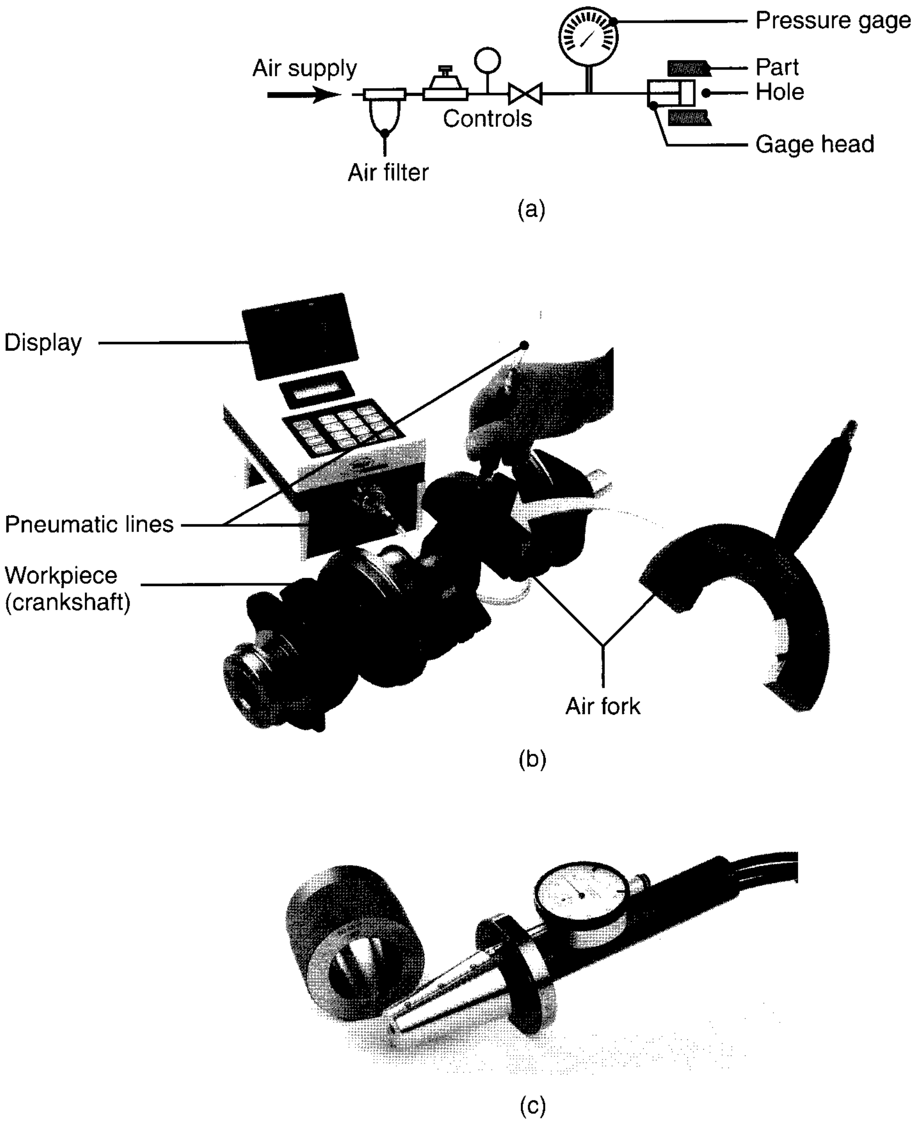


FIGURE 35.11 (a) Schematic illustration of the principle of an air gage. (b) Illustration of an air-gage system used to measure the main bearing dimension on a crankshaft. (c) A conical head for air gaging; note the three small airholes on the conical surface. *Source*: (b) Courtesy of Mahr Federal, Inc. (c) Courtesy of Stotz Gaging Co.

and indicated by a pressure gage, is calibrated to measure the dimensional variations of holes.

The air gage can be rotated during use to indicate and measure any out-of-roundess of the hole. The outside diameters of parts (such 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. 35.11b). Various shapes of air heads, such as the conical head shown in Fig. 35.11c, can be prepared for use in specialized applications on parts with different geometric features.

Air gages are easy to use, and the resolution can be as fine as 0.125 μm. If the surface roughness of the part is too high, the readings may be unreliable. The compressed-air supply must be clean and dry for proper operation. The part being measured does not have to be free of dust, metal particles, or similar contaminants, because the air will blow them away. The noncontacting nature and the low pressure of an air gage has the benefit of not distorting or damaging the measured part, as could be the case with mechanical gages—thus giving erroneous readings.

35.5 Modern Measuring Instruments and Machines

A wide variety of measuring instruments and gages has been developed. They range from simple, hand-operated devices to computer-controlled machines with very large workspaces.

Electronic Gages. Unlike mechanical systems, *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 displayed as linear dimensions with a digital readout. A handheld electronic gage for measuring bore diameters is shown in Fig. 35.12. When its handle is squeezed slightly, the tool can be inserted into the bore, and the bore diameter is read directly. A microprocessor-assisted electronic gage for measuring vertical length is shown in Fig. 35.13.

A commonly used electronic gage is the **linear-variable differential transformer** (LVDT), for measuring small displacements. Electronic caliper gages with diamond-coated edges are available. The chemical vapor deposition (CVD) coating on these gages has a wear resistance superior to that of steel or tungsten-carbide edges; it also resists corrosion.

Although they are more expensive than other types of gages, electronic gages have advantages in ease of operation, rapid response, a digital readout, less possibility of human error, versatility, flexibility, and the capability to be integrated into automated systems through microprocessors and computers.

FIGURE 35.12 An electronic gage for measuring bore diameters. The measuring head is equipped with three carbide-tipped steel pins for wear resistance. The LED display reads 29.158 mm. *Source*: Courtesy of TESA SA.

Laser Micrometers. In this instrument, a laser beam scans the workpiece (Fig. 35.14), typically at a rate of 350 times per second. 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 continuous, high-speed production. In addition, 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. Laser micrometers are of the shadow type or are charge-coupled device (CCD) based for in-line measurement while a part is in production.

Laser micrometers are available with various capacities and features. They can be handheld for manual operation, or they can be mounted on and integrated with computer-controlled machines and statistical-process control units.

Laser Interferometry. This technique is used to check and calibrate machine tools for various geometric features during 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.

Photoelectric Digital Length Measurement. This type of measurement is done by an instrument that can measure the overall dimensions, thickness, and depth of a variety of parts. Resolution settings can range from 5 to 0.01 μm.

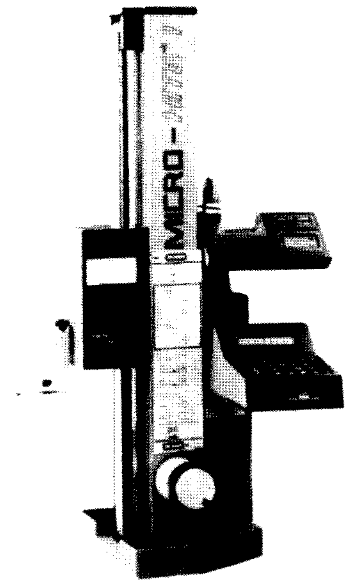


FIGURE 35.13 An electronic vertical-length measuring instrument with a resolution of 1 μm (40 μin). *Source*: Courtesy of TESA SA.

35.5.1 Coordinate-measuring Machines

As schematically shown in Fig. 35.15a, a *coordinate-measuring machine* (CMM) consists basically of a platform on which the workpiece being measured is placed and is then moved linearly or rotated. A probe (Fig. 35.15b; see also Fig. 25.6) is attached to a head (capable of various movements) and records all measurements. In addition to the tactile probe shown, other types of probes are scanning, laser (Fig. 35.15c), and vision probes, all of which are nontactile. A CMM for inspection of a typical part is shown in Fig. 35.15d.

Coordinate-measuring machines are very versatile and capable of recording measurements of complex profiles with high resolution (0.25 μm) and high speed. They are built rigidly and ruggedly to resist environmental effects in manufacturing plants, such as temperature variations and vibration. They can be placed close to machine tools for efficient inspection and rapid feedback; that way, processing parameters are corrected before the next part is made. Although large CMMs can be expensive, most machines with a touch probe and computer-controlled three-dimensional movement are suitable for use in small shops and generally cost under $20,000.

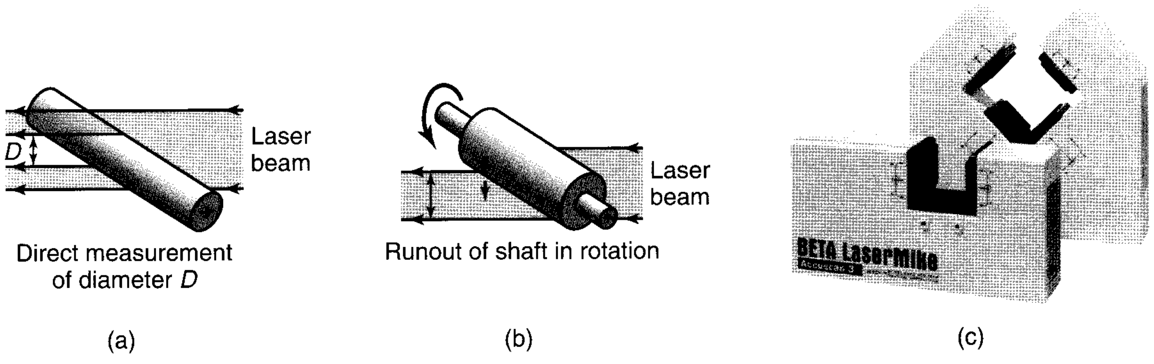


FIGURE 35.14 (a) and (b) Two types of measurements made with a laser scan micrometer. (c) Two types of laser micrometers. Note that the instrument in the front scans the part (placed in the opening) in one dimension; the larger instrument scans the part in two dimensions. *Source*: Courtesy of BETA LaserMike.

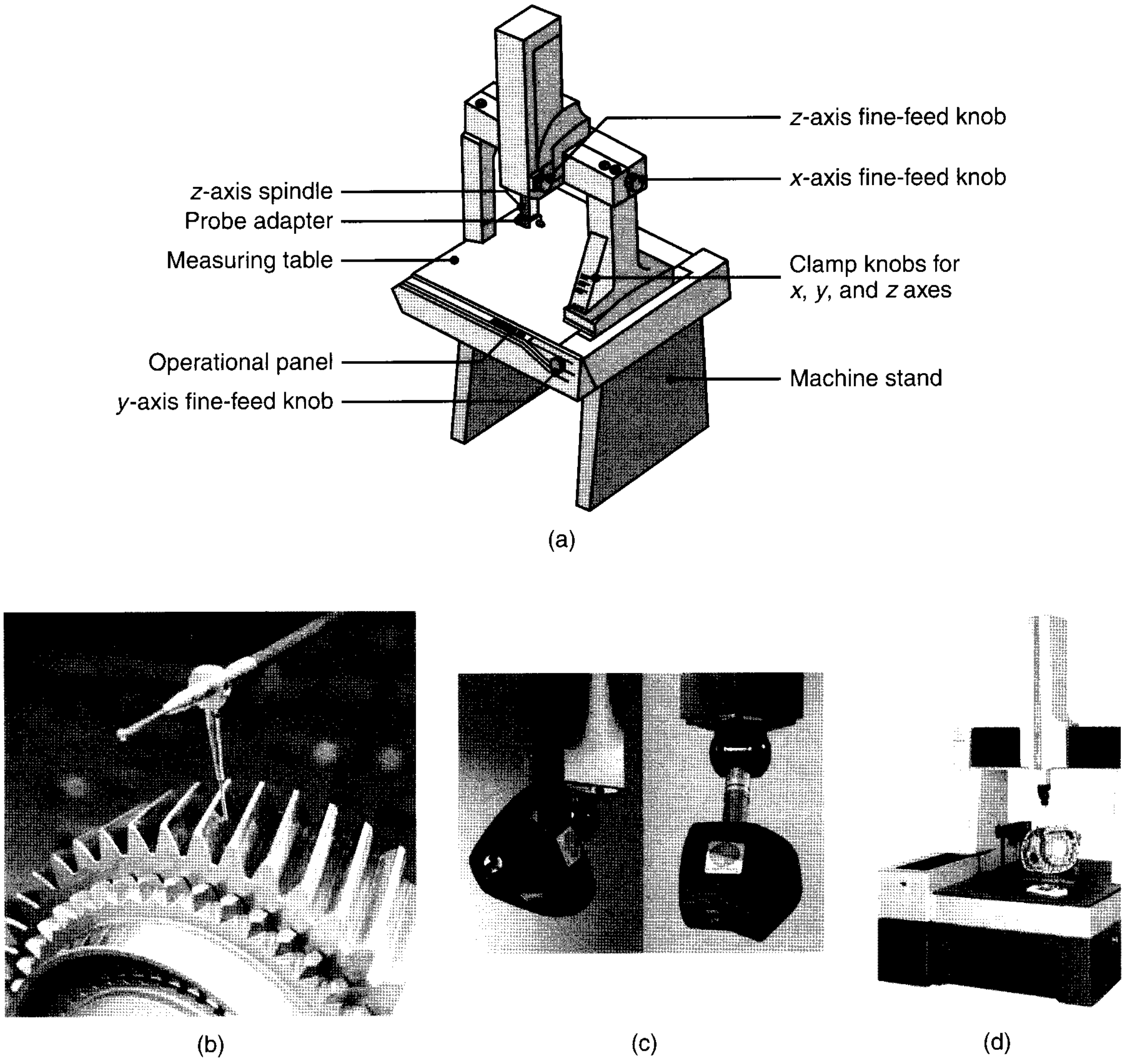


FIGURE 35.15 (a) Schematic illustration of a coordinate-measuring machine. (b) A touch signal probe. (c) Examples of laser probes. (d) A coordinate-measuring machine with a complex part being measured. *Source*: (b) through (d) Courtesy of Mitutoyo America Corp.

EXAMPLE 35.2 Coordinate-measuring Machine for Car Bodies

A large horizontal CNC coordinate-measuring machine used to measure all dimensions of a car body is shown in Fig. 35.16. This machine has a measuring range of 6 × 1.6 × 2.4 m high and a resolution of 0.1 μm. The system has temperature compensation within a range from 16° to 26°C to maintain 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 measuring speed is 5 mm/s. The probes are software controlled, and the machine is equipped with safety devices to prevent the

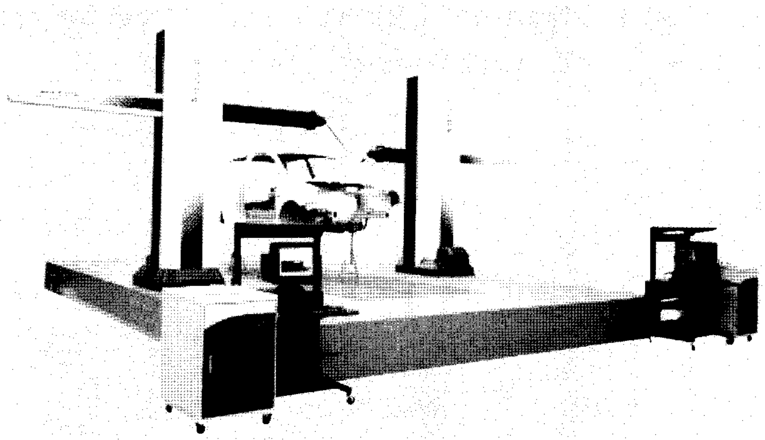


FIGURE 35.16 A large coordinate-measuring machine with two heads measuring various dimensions on a car body.

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 that controls all movements and records all measurements.

Source: Courtesy of Mitutoyo America Corporation.

35.6 Automated Measurement and Inspection

Automated measurement and inspection is based on various **online sensor systems** that monitor the dimensions of parts while they are being made and, if necessary, use these measurements as input to make corrections (Section 37.7). Manufacturing cells and flexible manufacturing systems (Chapter 39) have led to the adoption of advanced measuring techniques and systems.

To appreciate the importance of online monitoring of dimensions, consider the following question: If a machine has been producing a certain part 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, as well as human, factors involved:

- Static and dynamic *deflections* of the machine because of vibrations and fluctuating forces are caused by machine characteristics and variations in the properties and dimensions of the incoming material.
- *Distortion* of the machine because of thermal effects are caused by such factors as changes in the temperature of the environment, changes of metalworking fluids, and changes of machine bearings and various components.
- *Wear* of tools, dies, and molds can affect the dimensional accuracy of the parts produced.
- Human *errors* and miscalculations cause problems.

As a result of these factors, the dimensions of parts will vary, 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:** The ratio of instrument output to the input dimension; also called *magnification*.
- **Calibration:** The adjustment or setting of an instrument to give readings that are accurate within a reference standard.
- **Drift:** An instrument's capability to maintain its calibration over time; also called *stability*.
- **Linearity:** The accuracy of the readings of an instrument over its full working range.
- **Magnification:** The ratio of instrument output to the input dimension; also called *amplification*.
- **Precision:** Degree to which an instrument gives repeated measurement of the same standard.
- **Repeat accuracy:** The same as accuracy, but repeated many 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.
- **Stability:** An instrument's capability to maintain its calibration over time; also called *drift*.

The 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 (temperature, humidity, dust, and so on), (c) the operator skills required, and (d) the cost of equipment.

35.8 Geometric Dimensioning and Tolerancing

Individually manufactured parts and components eventually are assembled into products. We take it for granted that when a thousand lawn mowers are manufactured and assembled, each part of the mower will mate properly with its intended components. For example, the wheels of the lawn mower will slip easily into their axles, or the pistons will fit properly into the cylinders, being neither too tight nor too loose.

Likewise, when we have to replace a broken or worn bolt on an old machine, we purchase an identical bolt. We are confident from similar experiences in the past that the new bolt will fit properly in the machine. The reason we feel confident is that the bolt is manufactured according to certain standards and the dimensions of all similar bolts will vary by only a small, specified amount that do not affect their function.

In other words, all bolts are manufactured within a certain range of dimensional tolerance; thus, all similar bolts are *interchangeable*. We also expect that the new bolt will function satisfactorily for a certain length of time, unless it is abused or misused. Bolts are periodically subjected to various tests during their production to make sure that their quality is within certain specifications.

Dimensional Tolerance. *Dimensional tolerance* is defined as the permissible or acceptable variation in the dimensions (height, width, depth, diameter, and angles) of a part. The root of the word "tolerance" is the Latin *tolerare*, meaning "to endure" or "to put up with." Tolerances are unavoidable, because it is virtually impossible and unnecessary to manufacture two parts that have precisely the same dimensions.

Furthermore, because close dimensional tolerances can increase the product cost significantly, a narrow tolerance range is economically undesirable. However, for some parts, close tolerances are necessary for their proper functioning and are worth the added expense associated with narrow tolerance ranges. Examples are precision measuring instruments and gages, hydraulic pistons, and bearings for aircraft engines.

Measuring dimensional tolerances and features of parts rapidly and reliably can be a challenging task. For example, each of the 6 million parts on a Boeing 747-400 aircraft requires the measurement of about 25 features, representing a total of 150 million measurements. Surveys have shown that the dimensional tolerances on state-of-the-art manufactured parts are shrinking by a factor of 3 every 10 years and that this trend will continue. It is estimated that accuracies of (a) conventional turning and milling machines will rise from the present 7.5 to 1 μm, (b) diamond-wheel wafer-slicing machines for semiconductor fabrication to 0.25 μm, (c) precision diamond turning machines to 0.01 μm, and (d) ultraprecision ion-beam machines to less than 0.001 μm. (See also Fig. 25.16.)

Importance of Tolerance Control. Dimensional tolerances become important only when a part is to be assembled or mated with another part. Surfaces that are free and not functional do not need close tolerance control. For example, the accuracy of the holes and the distance between the holes for a connecting rod are far more critical than the rod's width and thickness at various locations along its length (see Fig. 14.7).

To appreciate the importance of dimensional tolerances, let's assemble a simple round shaft (axle) and a wheel with a round hole. Assume that we want the axle's diameter to be 25 mm (Fig. 35.17). We go to a hardware store and purchase a 25-mm round rod and a wheel with a 25-mm hole. Will the rod fit into the hole without our having to force it, or will it be loose in the hole? The 25-mm dimension is the **nominal size** of the shaft. If we purchase such a rod from a different store, at a different time, and select one randomly from a large lot, the chances are that each rod will

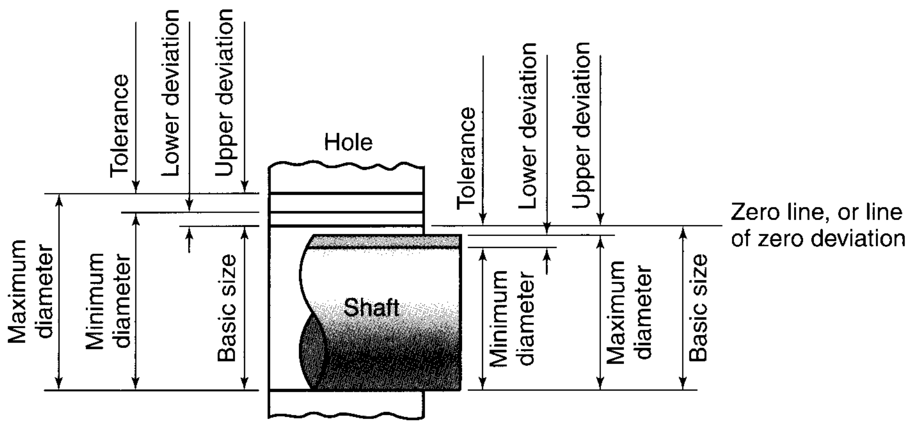


FIGURE 35.17 Basic size, deviation, and tolerance on a shaft, according to the ISO system.

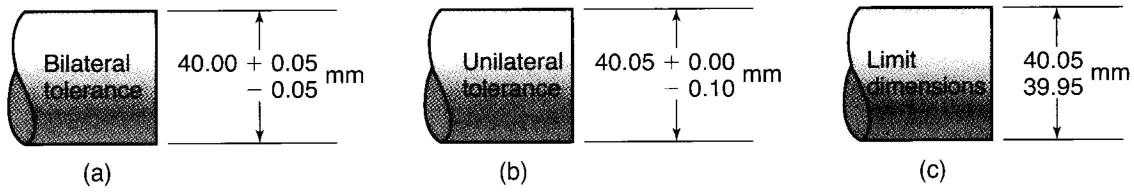


FIGURE 35.18 Various methods of assigning tolerances on a shaft: (a) bilateral tolerance, (b) unilateral tolerance, and (c) limit dimensions.

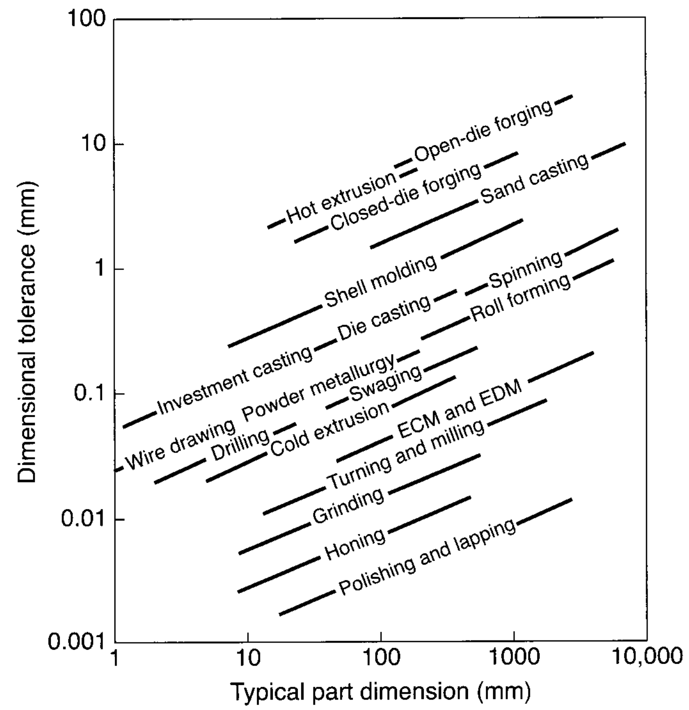


FIGURE 35.19 Dimensional tolerances as a function of part size for various manufacturing processes. Note that because many factors are involved, there is a broad range for tolerances.

have a slightly different diameter. Machines with the same setup may produce rods of slightly different diameters, depending on a number of factors, such as speed of operation, temperature, lubrication, and variations in the properties of the incoming material. If we now specify a *range* of diameters for both the rod and the hole of the wheel, we can predict the type of fit correctly.

Certain terminology has been established to clearly define these geometric quantities. One such system is the International Organization for Standardization (ISO) system shown in Fig. 35.17. Note that both the shaft and the hole have minimum and maximum diameters, the difference being the tolerance for each member. A proper engineering drawing would specify these parameters with numerical values, as shown in Fig. 35.18.

The range of dimensional tolerances possible in manufacturing processes is given in various figures and tables throughout this book. There is a general relationship between tolerances and part size (Fig. 35.19) and between tolerances and the surface finish of parts manufactured by various processes (Fig. 35.20). Note the wide range of tolerances and surface finishes obtained. Also, the larger the part, the greater is its obtainable 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. Because of the complex geometric relationships involved among all of the parts to be assembled, the definitions of these terms can be somewhat confusing.

The commonly used terms for geometric characteristics are defined briefly as follows, in alphabetical order:

- **Allowance:** The specified difference in dimensions between mating parts; also called *functional dimension* or *sum dimension.*
- **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.

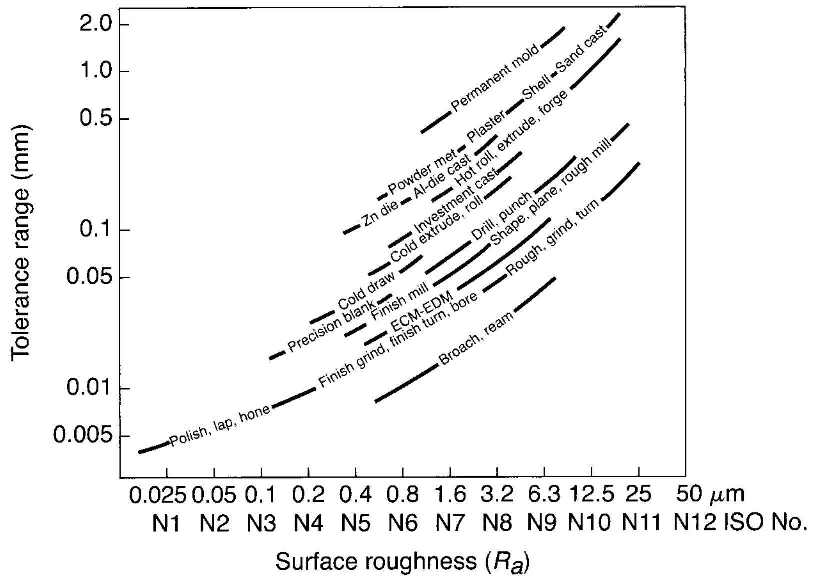


FIGURE 35.20 Dimensional tolerance range and surface roughness obtained in various manufacturing processes. These tolerances apply to a 25-mm workpiece dimension.
Source: After J.A. Schey.

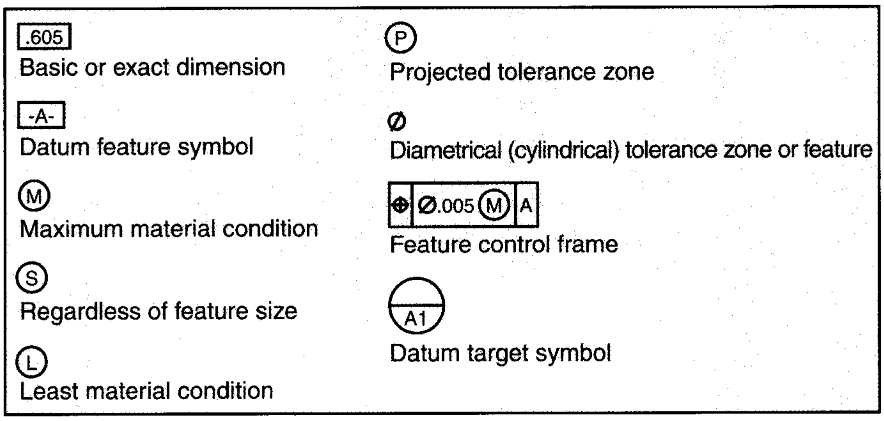
- **Clearance:** The space between mating parts.
- **Clearance fit:** Fit that allows for rotation or sliding between mating parts.
- **Datum:** A theoretically exact axis, point, line, or plane.
- **Feature:** A physically identifiable portion of a part, such as hole, slot, pin, or chamfer.
- **Fit:** 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:** Tolerances based on a zero line on the hole; also called *standard hole practice* or *basic hole system*.
- **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 dimensions:** The maximum and minimum dimensions of a part; also called *limits*.
- **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:** Tolerances based on a zero line on the shaft; also called *standard shaft practice* or *basic shaft system*.
- **Standard size:** Nominal size in integers and common subdivisions of length.
- **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 Figs. 35.21a and b.

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 35.21 Geometric characteristic symbols to be indicated on engineering drawings of parts to be manufactured. *Source*: Courtesy of The American Society of Mechanical Engineers. .

Limits and Fits. *Limits and fits* are essential in specifying dimensions for holes and shafts. There are two standards on 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 has been developed in greater detail. In these standards, capital letters always refer to the hole and lowercase letters to the shaft.

SUMMARY

- In modern manufacturing technology, many parts are processed to a high degree of precision and thus require measuring instrumentation with several features and characteristics.
- Many devices are available for inspection—from simple gage blocks to electronic gages with high resolution. The selection of a particular measuring instrument depends on factors such as the type of measurement for which it will be used, the environment in which it will be used, and the accuracy of measurement required.
- 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. Tolerances not only affect the accuracy and operation of all types of machinery and equipment, but also can influence product cost significantly.
- The smaller (tighter) the range of tolerances specified, the higher is the cost of production. Tolerances should be as broad as possible, but should also maintain the functional 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

Bentley, J.P., **Principles of Measurement Systems**, 4th ed., Prentice Hall, 2005.

Bosch, J.A. (ed.), **Coordinate Measuring Machines and Systems**, Marcel Dekker, 1995.

Cogorno, G., **Geometric Dimensioning and Tolerancing for Mechanical Design**, McGraw-Hill, 2006.

Creveling, C.M., **Tolerance Control: A Handbook for Developing Optimal Specifications**, Addison-Wesley, 1996.

Curtis, M.A., **Handbook of Dimensional Measurement**, 4th ed., Industrial Press, 2007.

Drake, P.J., **Dimensioning and Tolerancing Handbook**, McGraw-Hill, 1999.

Gooldy, G., **Geometric Dimensioning and Tolerancing**, rev. ed., Prentice Hall, 1995.

Kimura, F., **Computer-Aided Tolerancing**, Springer, 1997.

Krulikowski, A., **Fundamentals of Geometric Dimensioning and Tolerancing**, Delmar, 1997.

Liggett, J.V., **Dimensional Variation Management Handbook: A Guide for Quality, Design, and Manufacturing Engineers**, Prentice Hall, 1993.

Madsen, D.A., **Geometric Dimensioning and Tolerancing**, Goodheart-Wilcox, 2003.

Meadows, J.D., **Measurement of Geometric Tolerances in Manufacturing**, Marcel Dekker, 1998.

Morris, A.S., **Measurement and Calibration for Quality Assurance**, Wiley, 1998.

———, **Measurement and Instrumentation Principles**, Butterworth-Heinemann, 2001.

Puncochar, D.E., **Interpretation of Geometric Dimensioning and Tolerancing**, 2nd ed., Industrial Press, 1997.

Whitehouse, D.J., **Handbook of Surface Metrology**, IOP Publishing, 1994.

Wilson, B.A., **Design Dimensioning and Tolerancing**, Goodheart-Wilcox, 1996.

———, **Dimensioning and Tolerancing Handbook**, Genium, 1998.

Winchell, W., **Inspection and Measurement in Manufacturing**, Society of Manufacturing Engineers, 1996.

REVIEW QUESTIONS

35.1. Explain what is meant by standards for measurement.

35.2. What is the basic difference between direct-reading and indirect-reading linear measurements? Name the instruments used in each category.

35.3. What is meant by comparative length measurement?

35.4. Explain how flatness is measured. What is an optical flat?

35.5. Describe the principle of an optical comparator.

35.6. Why have coordinate measuring machines become important instruments?

35.7. What is the difference between a plug gage and a ring gage?

35.8. What are dimensional tolerances? Why is their control important?

35.9. Explain the difference between tolerance and allowance.

35.10. What is the difference between bilateral and unilateral tolerance?

35.11. How is straightness measured?

QUALITATIVE PROBLEMS

35.12. Why are the words "accuracy" and "precision" often incorrectly interchanged?

35.13. Why do manufacturing processes produce parts with a wide range of tolerances? Explain, giving several examples.

35.14. Explain the need for automated inspection.

35.15. Dimensional tolerances for nonmetallic parts usually are wider than for metallic parts. Explain why. Would this also be true for ceramics parts?

35.16. Comment on your observations regarding Fig. 35.20. Why does dimensional tolerance increase with increasing surface roughness?

35.17. Review Fig. 35.19, and comment on the range of tolerances and part dimensions produced by various manufacturing processes.

35.18. In the game of darts, is it better to be accurate or to be precise? Explain.

35.19. What are the advantages and limitations of *GO* and *NOT GO* gages?

35.20. Comment on your observations regarding Fig. 35.18.

35.21. Why is it important to control temperature during the measurement of dimensions? Explain, with examples.

35.22. Describe the characteristics of electronic gages.

35.23. What method would you use to measure the thickness of a foam-rubber part? Explain.

QUANTITATIVE PROBLEMS

▶**35.24.** 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 30.00 mm at room temperature?

▶**35.25.** If the same steel rule as in Problem 35.24 is used to measure aluminum extrusions, what will be the indicated diameter at room temperature? What if the part were made of a thermoplastic?

▶**35.26.** A shaft must meet a design requirement of being at least 28.0 mm in diameter, but it can be 0.380 mm oversized. Express the shaft's tolerance as it would appear on an engineering drawing.

SYNTHESIS, DESIGN, AND PROJECTS

35.27. Describe your thoughts on the merits and limitations of digital measuring equipment over analog instruments. Give specific examples.

35.28. 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.

35.29. Obtain a digital micrometer and a steel ball of, say, 6-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.

35.30. Repeat Problem 35.29, 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.

35.31. What is the significance of the tests described in Problems 35.29 and 35.30?

35.32. Explain the relative advantages and limitations of a tactile probe versus a laser probe.

35.33. 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.

35.34. 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.

35.35. 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.

35.36. Explain how you would jusify the considerable cost of a coordinate-measuring machine such as that shown in Fig. 35.16.

CHAPTER

36

Quality Assurance, Testing, and Inspection

- This chapter outlines the procedures used to ensure the manufacture of high-quality products.
- The chapter describes the mathematical tools and inspection techniques that have been developed to produce higher quality products; among such tools and techniques are total quality management and Taguchi methods, as well as 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 ends with a discussion of the methods used in destructive and non-destructive testing of materials and products.

36.1 Introduction

Throughout this book, it has been repeatedly emphasized that manufactured products develop certain external and internal characteristics which result, in part, from the type of production processes used. External characteristics most commonly involve dimensions, size, and surface finish and integrity considerations, such as surface damage from cutting tools or friction during the processing of the workpiece. Internal characteristics include defects such as porosity, impurities, inclusions, phase transformations, embrittlement, cracks, debonding of laminations, and residual stresses.

Some of these defects may exist in the original stock, and some are introduced or induced during the particular manufacturing operation. Before they are marketed, manufactured parts and products are inspected for several characteristics. This inspection routine is important in order to

- Ensure dimensional accuracy so that parts fit properly into other components during assembly.
- Identify products whose failure or malfunction may have serious implications, including bodily injury or fatality. Typical examples are elevator cables, switches, brakes, grinding wheels, railroad wheels, turbine blades, pressure vessels, and welded joints.

This chapter identifies and describes the various methods that are commonly used to inspect manufactured products.

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, particularly for large corporations in industrialized countries. In Japan, the single term **kaizen** is used to signify *never-ending improvement.*

The prevention of defects in products and online inspection of parts are major goals in all manufacturing activities. Again, *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 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 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

We all have used terms such as "poor quality" or "high quality" to describe a particular product or the products of a particular company. What is quality? Although we may recognize it when we see or use a product, quality, unlike most technical terms, is difficult to define precisely. Simply and generally, quality may be defined as *a product's fitness for use.*

Several aspects of quality that generally are identified are performance, durability, reliability, robustness, availability, cost, and serviceability, as well as aesthetics and perceived quality. 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.

For example, consider the following: (a) The handle on a kitchen utensil is installed crookedly, or the handle discolors or cracks during its use, (b) a weighing scale functions erratically, (c) a plastic toy breaks easily, (d) a vacuum cleaner requires frequent repairs, (e) the stem of a screwdriver bends, and (f) a machine tool cannot maintain the dimensional tolerances of the workpiece because of lack of stiffness or poor construction. These examples all lead us to believe that the product is of low quality. 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. Some examples of this type of product are good-quality kitchen utensils, refrigerators, washing machines, tools, bicycles, and automobiles.

Note that, in describing good- or poor-quality products, this book has stated neither the intended lifetimes of those products nor any of their technical specifications. Design and manufacturing engineers have the joint responsibility of selecting and specifying materials for the components of the products to be made. For example, in selecting the type of metal for a screwdriver stem, we can specify materials that have high strength and high resistance to wear and corrosion and then process them with the use of appropriate manufacturing techniques, including heat treatments and coatings.

It is important to note, however, that materials possessing better properties generally are more expensive and may be more difficult to process than those with

TABLE 36.1

Average Life Expectancy of Some Products	
U.S. dollar bill	18 months
Car battery	4 years
Hair dryer	5 years
Water heater (gas)	12 years
Vacuum cleaner	10 years
Air-conditioning unit	15 years
Nuclear reactor	40 years
Automobile disk brake	65,000 km
Muffler	50,000 km
Tire	65,000–100,000 km

poorer properties. The level of quality that a manufacturer chooses for its products may depend on the market for which the products are intended. For example, low-cost, low-quality tools have their own world market niche.

As described in Section 40.9, the total product cost depends on several variables, including the level of automation in the manufacturing plant. Thus, there are many ways for the engineer to review and modify overall product design and manufacturing processes to minimize a product's cost without affecting its quality. 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.
- Result in the need for in-field repairs.
- Have the significant built-in cost of customer dissatisfaction.

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. Typical life expectancies of some products are given in Table 36.1 (see also Table I.4 in the General Introduction).

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.

These 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.

Quality assurance is the responsibility of everyone involved with design and manufacture. The often-repeated statement that quality must be built into a product reflects this important concept. Although a finished product can be inspected for quality, and rejected if conditions are not met, quality cannot be inspected into a finished product.

An important aspect of quality assurance is the capability to (a) analyze defects as they occur on the production line and (b) promptly eliminate them or reduce them 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 as **total quality control** and, in an even larger sense, **total quality management**.

It is clear that, 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 to determine whether it meets the specifications for dimensional tolerances, surface finish, defects, and other characteristics.

36.4 Total Quality Management

Total quality management (TQM) is a system which emphasizes the concept that quality must be designed and built into a product. It 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 here.

Leadership and *teamwork* in the organization are essential. They ensure that the goal of **continuous improvement** in manufacturing operations is foremost, because they reduce product *variability* and they improve customer satisfaction. The TQM concept also requires control of the *processes*, and not the *parts produced*, so that process variability can be reduced and no defective parts are allowed to continue through the production line.

Quality Circle. The basic concept of a *quality circle* consists of regular meetings by groups of employees (workers, supervisors, and managers) who 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 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, as described in Section 39.6.

Quality Engineering as a Philosophy. Experts in quality control have put 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 a major impact on modern manufacturing.

36.4.1 Deming Methods

During World War II, W.E. Deming (1900–1993) 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* in (a) the performance of machines and of people and (b) the quality and dimensions of raw

TABLE 36.2

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 everybody in the company to work to accomplish the transformation.

materials. The efforts of these pioneers involved not only statistical methods of analysis, but also a new way of looking at manufacturing operations—that is, from the perspective of *improving quality while lowering costs.*

Deming recognized that manufacturing organizations are systems of management, workers, machines, and products. His basic ideas are summarized in the now-well-known *14 points*, given in Table 36.2. These points are not to be seen as a checklist or menu of tasks; they are the characteristics that Deming recognized in companies that produce high-quality goods. He placed great emphasis on communication, direct worker involvement, and education in statistics and modern manufacturing technology.

36.4.2 Juran Methods

A contemporary of Deming, J.M. Juran (1904–2008) emphasized the importance of the following ideas:

- 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 (the end users who purchase the product or service), or they may be internal (the 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 are then responsible for implementing both quality control and continued improvement in quality.

36.5 Taguchi Methods

In G. Taguchi's (1924–) 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 by Taguchi to manufacture high-quality products. One fundamental viewpoint put forward is the quality challenge facing manufacturers: Provide products that delight your customers. To delight customers, manufacturers should offer products with the following 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 product characteristics clearly are the goals of manufacturers striving to provide high-quality products. Although it is very challenging to actually 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 the following concepts:

- **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. They involve the use of *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.

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.

An important concept introduced by Taguchi is that any deviation from a design objective constitutes a loss in quality. Consider, for example, the tolerancing standards for a shaft, given in Fig. 35.18. On the one hand, there is a range of dimensions over which a part is acceptable; on the other hand, the Taguchi philosophy calls for a minimization of deviation from the design objective. Thus, with Fig. 35.18a as an example, a shaft with a diameter of 40.03 mm normally would be considered acceptable and thus would pass inspections. In the Taguchi approach, however, a shaft with this diameter represents a deviation from the design objective. Such deviations generally reduce the robustness and performance of products, especially in complex systems.

36.5.1 Robustness

Another aspect of quality is a concept originally suggested by Taguchi that continuously has grown in importance and is referred to as *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 function and performance) have minimal sensitivity to its input variations (such as variations in environment, load, and power source). Moreover, a robust product or machine is insensitive to changes in tolerance over its intended life.

For example, in a robust design, a part will function sufficiently well even if the loads applied, or their directions, go beyond anticipated values. Likewise, a robust machine or system will undergo minimal deterioration in performance even if it experiences variations in environmental conditions, such as temperature, humidity, air quality, and vibrations. Also, a robust machine will 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 then 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 first design. Even though the precision of hole location remains the same and the manufacturing cost is also the same, the variability in the top edge of the bracket (from the horizontal) has now been reduced by one-half. However, if the bracket is subjected to vibration, the bolts may loosen over time.

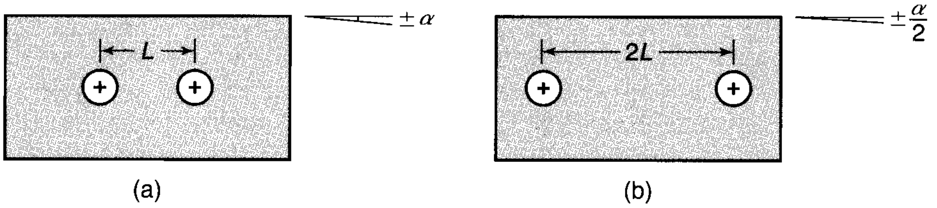


FIGURE 36.1 A simple example of robust design. (a) Location of two mounting holes on a sheet-metal bracket where the deviation keeping the top surface of the bracket from being perfectly horizontal is $\pm\alpha$. (b) New locations of holes; now the deviation (keeping the top surface of the bracket from being perfectly horizontal) is reduced to $\pm\alpha/2$.

An even more robust design approach would be to use an adhesive to hold the threads in place or to use some type of fastener that would not loosen over time.

36.5.2 Taguchi Loss Function

The *Taguchi loss function* was introduced in the early 1980s because traditional accounting practices had no real way of calculating losses on parts that met design specifications. In the traditional accounting approach, a part is defective and incurs a loss to the company when it exceeds its design tolerances; otherwise, there is no loss to the company.

The Taguchi loss function 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

$$\text{Loss cost} = k[(Y - T)^2 + \sigma^2], \tag{36.1}$$

where Y is the mean value from manufacturing, T is the target value from design, σ is the standard deviation of parts from manufacturing, 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

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 ± 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. Determine the Taguchi loss function and the payback period for the investment.

Solution Let's first identify the quantities involved: 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

$$k = \frac{(\$10.00)}{(3.2 - 2.6)^2} = \$27.28.$$

The loss cost before the improvement is then

$$\begin{aligned}\text{Loss cost} &= (27.78)[(2.6 - 2.6)^2 + 0.2^2] \\ &= \$1.11 \text{ per unit.}\end{aligned}$$

After the improvement, the standard deviation is 0.1 mm; thus, the loss cost is

$$\begin{aligned}\text{Loss cost} &= (27.78)[(2.6 - 2.6)^2 + 0.1^2] \\ &= \$0.28 \text{ per unit.}\end{aligned}$$

The savings are then (\$1.11 − \$0.28)(10,000) = \$8300 per month. Hence, the payback period for the investment is \$50,000/(\$8300/month) = 6.02 months.

CASE STUDY 36.1 Manufacture of Television Sets by Sony Corporation

Sony Corporation executives found a confusing situation confronting them in the mid-1980s. Televisions manufactured in Japanese production facilities sold faster than those produced in a San Diego facility, even though they were produced from identical designs. There were no identifications to distinguish the televisions made in Japan from those made in the United States, so there was no apparent reason for this discrepancy. However, investigations revealed that the televisions produced in Japan were superior to the U.S. versions; color sharpness was better and hues were more brilliant. Since the televisions were on display in stores, consumers could easily detect and purchase the model that had the best picture.

The difference in picture quality was obvious, but the reasons for that 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. The Japanese facility did not have a total quality program, but there was an emphasis on reducing variation 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) makes 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 the uniform distribution without defects to be superior to the distribution in which a few defects are produced but the majority of parts are closer to the design target values. However, consumers can readily detect which product is superior, and the marketplace proves that minimizing deviations is a worthwhile quality goal.

Source: After D.M. Byrne and G. Taguchi.

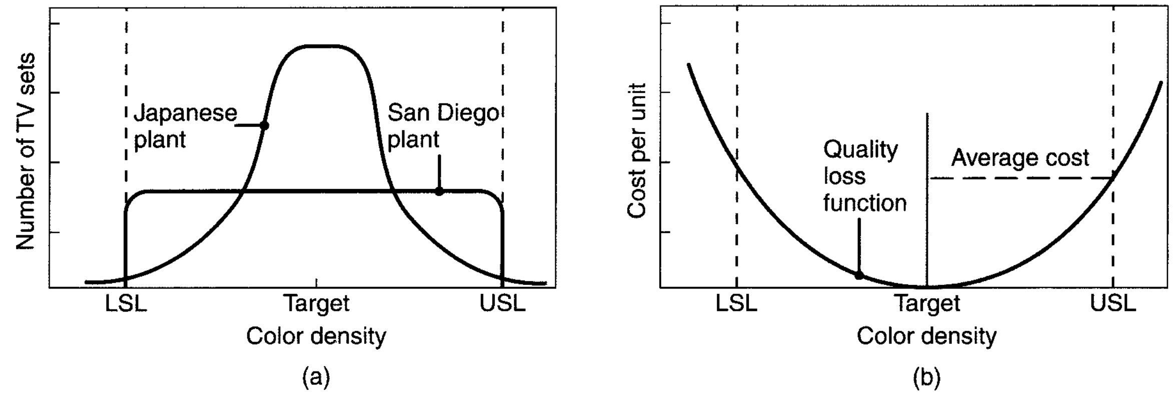


FIGURE 36.2 (a) Objective-function value distribution of color density for television sets. (b) Taguchi loss function, showing the average replacement cost per unit to correct quality problems. *Source*: After G. Taguchi.

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

With increasing international trade and global competition, customers worldwide increasingly are 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 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, equally important concerns regarding the environment and quality of life also are being addressed. This section describes the standards relevant to product quality and environmental issues.

36.6.1 The ISO 9000 Standard

First published in 1987 and then revised in 1994, the ISO 9000 standard (**Quality Management and Quality Assurance Standards**) is a deliberately generic series of quality system-management standards. The ISO 9000 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. The cost depends 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 August of 1994. Prior to the development of QS 9000, each of these automotive companies had its own standard for quality system requirements. Tier I suppliers have been required to obtain third-party registration to QS 9000 before the dates established by each of the Big Three companies. Very often, QS 9000 has been described as an "ISO 9000 chassis with a lot of extras." This is a good description, given that all of the ISO 9000 clauses serve as the foundation of QS 9000. However, the "little" extras are substantial.

The February 1995 edition of QS 9000 has three sections. Section I contains all 20 of the ISO 9001 clauses, but almost every clause has additional requirements for QS 9000. Section II has three sections: "Production Part Approval Process," "Continuous Improvement," and "Manufacturing Capabilities." Section III is entitled "Customer-Specific Requirements" and contains separate sections for Chrysler, General Motors, Ford, and truck manufacturers, respectively. Existing QS 9000 registrations are being upgraded continuously to comply with new editions of QS 9000.

36.6.3 The ISO 14000 Standard

ISO 14000 is a family of standards first published in September of 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.

Companies in most countries have rapidly been obtaining certification for this standard. 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. The following list describes some of the commonly observed variables in manufacturing:

- Cutting tools, dies, and molds are subject to wear; thus, part dimensions and surface characteristics vary over time.

- Machinery performs differently depending on its quality, age, condition, and maintenance; thus, older machines tend to chatter and vibrate, are difficult to adjust, and do not maintain tolerances as well as new machines.
- 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 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, and surface characteristics.
- Operator attention may vary during the day, from machine to machine or from operator to operator.

Those events which occur *randomly*—that is, without any particular trend or pattern—are called **chance variations** or **special causes.** Those events which can be traced to *specific causes* are called **assignable variations** or **common causes.**

Although the existence of variability in production operations has been recognized for centuries, it was Eli Whitney (1765–1825) who first understood its full significance when he found 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).

36.7.1 Statistical Quality Control

To understand *statistical quality control* (SQC), 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 should 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 be considered as representative of the population.

The sample is inspected for several characteristics and features, such as tolerances, surface finish, and defects, with the instruments and techniques described in Chapter 35 and in Sections 36.10 and 36.11. These characteristics fall into two categories: those which are measured quantitatively (method of variables) and those which 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 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, or dents in sheet-metal products) in each of the units in the group under consideration. The sample size for attributes-type data generally is larger than for variables-type data.

Assume now that you are using a micrometer to measure the diameters of turned shafts produced on a lathe (Fig. 35.2). You will soon note that the diameters vary, even though you ideally want all of the shafts to be exactly the same size. Let's now turn to the consideration of statistical quality-control techniques, which allow us to evaluate these variations and set limits on the acceptance of parts. If you list the measured diameters of the turned shafts in a given population, you will see that one or more parts have the smallest diameter and one or more have the largest diameter. The rest of the turned shafts have diameters that lie between these two extremes.

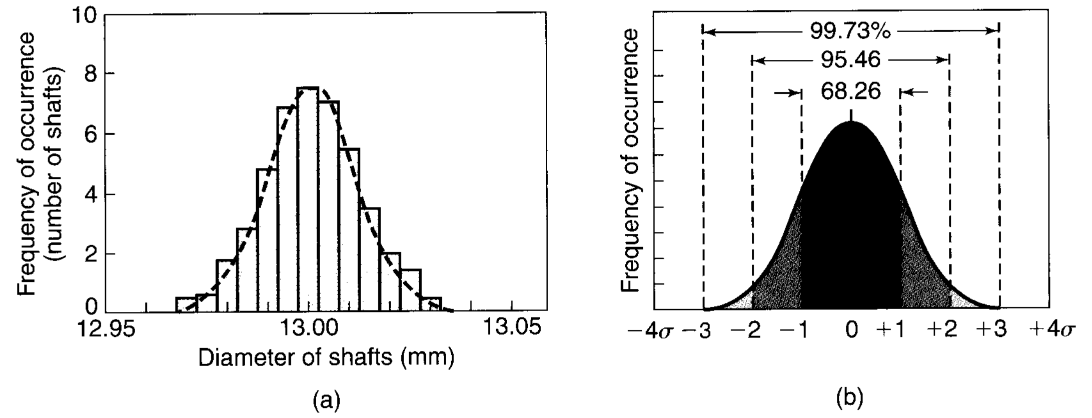
All the diameter measurements may be grouped and plotted in a *histogram* (bar graph) representing the number of parts in each diameter group (Fig. 36.3a). The bars show a **distribution** (also called a **spread** or **dispersion**) of the shaft-diameter measurements. The *bell-shaped curve* in Fig. 36.3a is called a **frequency distribution** and shows the frequencies with which parts of each diameter size are being produced.

Data from manufacturing processes often fit curves represented by a mathematically derived **normal-distribution curve** (Fig. 36.3b). This type of curve is also called *Gaussian*, after K.F. Gauss (1777–1855), who developed it on the basis of *probability*.

The bell-shaped normal distribution curve fitted to the data in Fig. 36.3a has two important features. First, it shows that most part diameters tend to cluster around an *average value* (**arithmetic mean**). This average is designated as $\bar{x}$ and is 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).



(a)

(b)

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.

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:

$$R = x_{\text{max}} - x_{\text{min}}. \tag{36.4}$$

The dispersion is estimated by the **standard deviation,** σ, and is 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. Note also that σ has the same units as x_i. Since we know the number of turned parts that fall within each group, we can calculate the percentage of the total population represented by each group. 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 $\pm 1\sigma$.

It can be seen that only 0.27% fall outside the $\pm 3\sigma$ range. This means that there are 2700 defective parts per million produced. In modern manufacturing, that is an unacceptable rate, in view of the observation that, at this level of defects, no modern computer would function reliably. However, note that these quantities are only valid for distributions that are normal, as shown in Fig. 36.3, and that are not skewed.

36.7.2 Six Sigma

Six sigma is a set of statistical tools based on the well-known 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 (Section 36.8.2), delivering defect-free products, and thus ensuring customer satisfaction. This approach consists of a clear focus on defining the quality problems, measuring relevant quantities, and analyzing, controlling, and improving processes and operations.

As mentioned in Section 36.7.1, three sigma in manufacturing would result in 0.27% (or 2700 parts per million) defective parts, an unacceptable rate in modern manufacturing. Also, in the service industries, 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 (i.e., the number of defective parts) increases during a production run, we must be able to determine the cause (incoming materials, machine controls, degradation of metalworking fluids, operator boredom, or 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 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, in order to avoid producing further defective parts. Known as *statistical process control* (SPC), this technique consists of various elements:

- Control charts and control limits
- Capabilities of the particular manufacturing process
- Characteristics of the machinery involved.

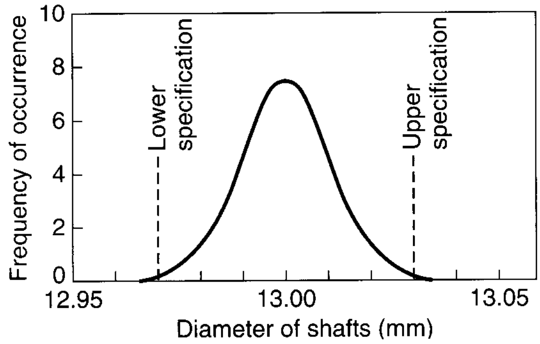


FIGURE 36.4 Frequency distribution curve showing lower and upper specification limits.

36.8.1 $\bar{x}$ and R Charts (Shewhart Control Charts)

The frequency distribution curve in Fig. 36.3b shows 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, which 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 plotted during production. Typically, there are two plots. The quantity $\bar{x}$ (Fig. 36.5a) is the average for each subset of samples taken and inspected—say, each subset consists of 5 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.

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 in these control charts 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 make sure that all parts are within three standard deviations of the mean $(\pm 3\sigma)$.

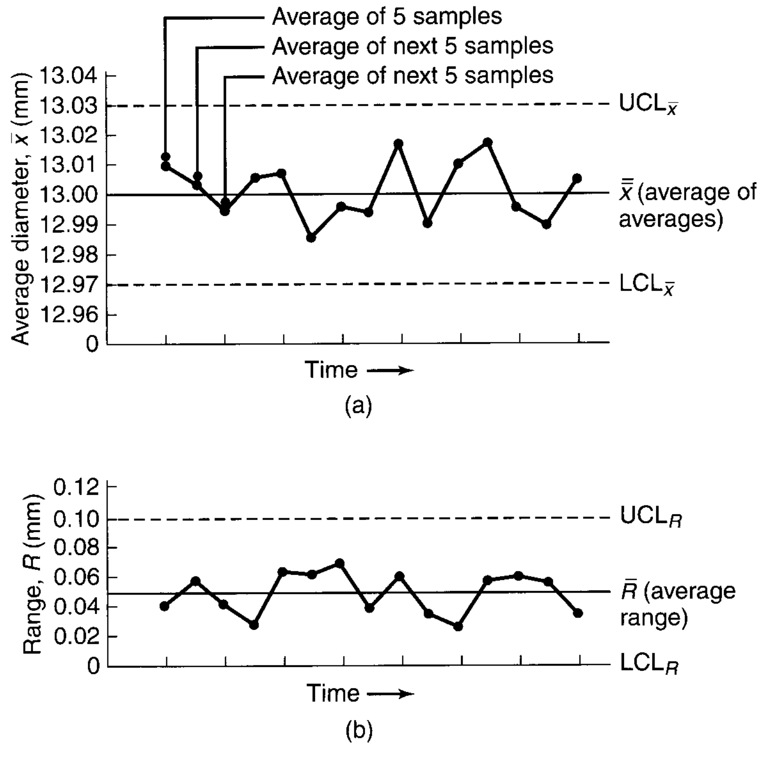


FIGURE 36.5 Control charts used in statistical quality control. The process shown is in good statistical control because all points fall within the lower and upper control limits. In this illustration, the sample size is 5 and the number of samples is 15.

TABLE 36.3

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

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 36.3 and $\bar{R}$ is the average of R values. The quantities $\bar{\bar{x}}$ and $\bar{R}$ are estimated from the measurements taken.

These control limits are calculated on the basis of the past production capability of the equipment itself; they 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 (Fig. 36.5b) shows 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 36.3. 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, we say that the process is *in good statistical control.* In other words,

- 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.

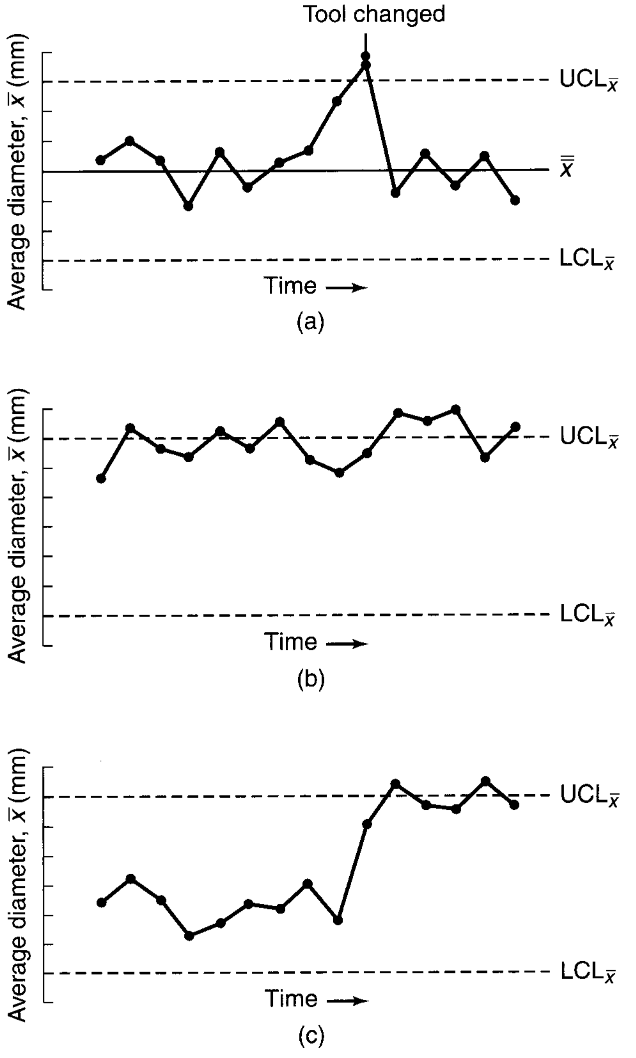


FIGURE 36.6 Control charts. (a) Process begins to become out of control because of such factors as tool wear (*drift*). The tool is changed and the process is then in statistical control. (b) Process parameters are not set properly; thus, all parts are around the upper control limit (*shift in mean*). (c) Process becomes out of control because of factors such as a change in the properties of the incoming material (*shift in mean*).

However, it can be seen that, in curves such as those in Figs. 36.6a, b, and c, there are certain *trends*. For example, note that in the middle of the curve in Fig. 36.6a the diameter of the shafts is increasing with time. The reason for this increase 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 caused by factors such 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 $\pm 2\sigma$.

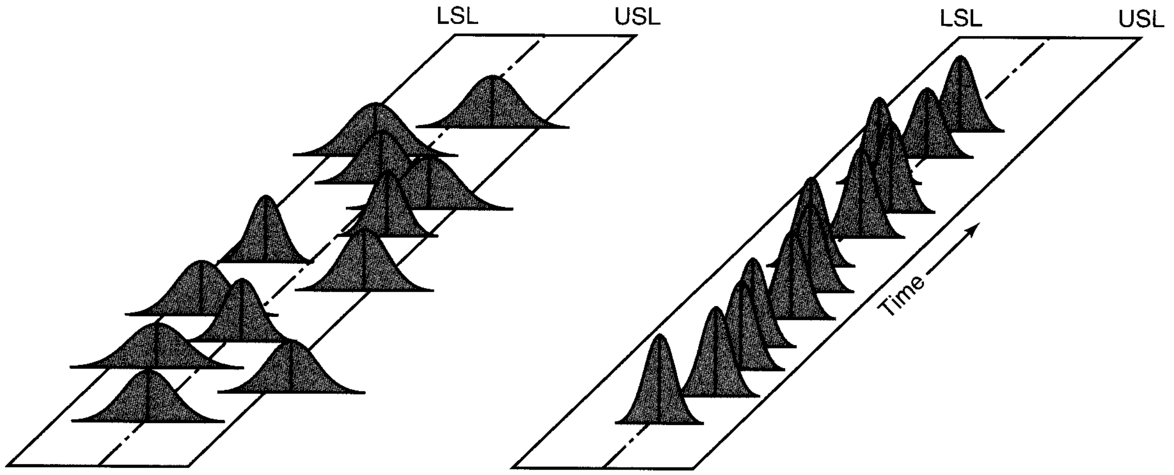
Analyzing patterns and trends in control charts requires considerable experience so that one may identify the specific cause(s) of an out-of-control situation. Among such causes may be one or more of those variables listed at the beginning of Section 36.7. Overcontrol of the manufacturing process (i.e., setting upper and lower control limits too close to each other, resulting in a smaller standard-deviation range) is a further cause of out-of-control situations and is the reason that 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 in the interest of efficiency of operation, operators should also have some responsibility for making adjustments in production processes that are beginning to be out of control. The capabilities of individual operators should be taken into account as well, so that they are not overloaded with data input and hence 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 produce defect-free products in controlled production. It tells us that the manufacturing process can produce parts consistently and repeatedly within specific limits of precision (Fig. 36.7).



(a) Unstable (b) Stable

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 have lower standard deviations than those of the distributions in part (a) and have means closer to the desired value. *Source*: After K. Crow.

Various indicators (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. Among the factors to be considered are variations in machine performance, operator skills, and incoming raw materials.

EXAMPLE 36.3 Calculation of Control Limits and Standard Deviation

The data given in Table 36.4 show length measurements (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.

Determine the upper and lower control limits and the standard deviation for the population of machined parts.

Solution We first calculate the average of averages, $\bar{\bar{x}}$:

$$\bar{\bar{x}} = \frac{1125.14}{10} = 112.51 \text{ mm}$$

TABLE 36.4

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

Then we calculate the average of the R values:

$$\bar{R} = \frac{26.0}{10} = 2.60 \text{ mm}$$

Since the sample size is 5, the following constants can be determined from Table 36.3: $A_2 = 0.577$, $D_4 = 2.115$, and $D_3 = 0$. The control limits now can be calculated from Eqs. (36.4) through (36.7). Thus, for averages,

$$\text{UCL}_{\bar{x}} = 112.51 + (0.577)(2.60) = 114.01 \text{ mm}$$

and

$$\text{LCL}_{\bar{x}} = 112.51 - (0.577)(2.60) = 111.01 \text{ mm}$$

For ranges,

$$\text{UCL}_R = (2.115)(2.60) = 5.50 \text{ mm}$$

and

$$\text{LCL}_R = (0)(2.60) = 0 \text{ mm}$$

From Eq. (36.10), we can now estimate the standard deviation σ for the population for a value of $d_2 = 2.326$:

$$\alpha = \frac{2.60}{2.326} = 1.18 \text{ mm}$$

CASE STUDY 36.2 Dimensional Control of Plastic Parts in Saturn Automobiles

A typical Saturn automobile has some 38 different injection-molded interior plastic parts (polycarbonate, polypropylene, and ABS), such as door panels, air-inlet ducts, consoles, and trim. All of these parts must conform to tight dimensional tolerances so that they fit and snap properly during assembly without unsightly gaps or buckles. However, the dimensions of these plastic parts change with temperature and humidity, and because of their flexibility, the parts also tend to bend and curl.

For this reason, measurement and inspection of plastic parts, (including the use of coordinate-measuring machines (CMMs; see Section 35.5.1) can be difficult. Although traditional gages also are used for monitoring process parameters in making these parts, a superior inspection system has been developed whereby feedback from statistical process control is received from a direct computer-controlled CMM so that the parts are molded properly.

The system compensates for the flexibility of the parts, allows automatic measurement of various part features, and makes measurements of the mold at periodic intervals. The data are analyzed on a regular basis, and when necessary, corrective actions are taken and changes are made in materials, processing, or mold design, so that the parts being molded will maintain good dimensional stability.

Source: Courtesy of Saturn Corp. and *Manufacturing Engineering*.

36.8.3 Acceptance Sampling and Control

Acceptance sampling consists of taking only 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 (MIL STD 105), this statistical technique is used widely and has become valuable. Acceptance sampling is particularly useful for inspecting high-production-rate parts when 100% inspection would be too costly. However, there are certain critical devices, such as pacemakers, prosthetic devices, and components of the space shuttle, 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. 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, and the lower is the probability of the lot's acceptance. **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 which 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 incentive for the manufacturer to control the production process better.

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 have been developed so that 100% inspection of all parts is possible and inspection can also be economical.

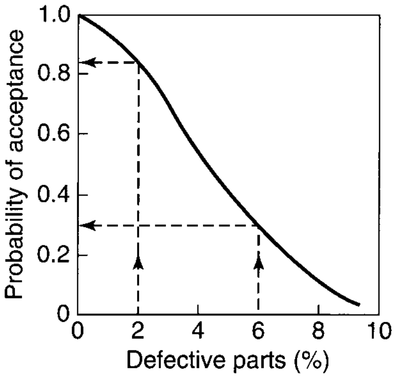


FIGURE 36.8 A typical operating-characteristics curve used in acceptance sampling. The higher the percentage of defective parts, the lower is the probability of acceptance by the consumer.

36.9 Reliability of Products and Processes

All products eventually fail in some manner or other: Automobile tires become worn and the treads become smooth, electric motors burn out, water heaters begin to leak, dies and cutting tools wear out, and machinery stops functioning properly. **Product reliability** may be defined as the probability that a product will perform its intended function in a given environment for a specified period while in normal use by the customer and without failure.

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 must be much higher than that of a kitchen faucet or a mechanical pencil. From the topics described in this chapter, it can be seen that, as the quality of each component of a product increases, so, too, does the reliability of the whole product.

For an ordinary steel chain, the reliability of each link in the chain is critical. Similarly, the reliability of each gear in a gear train for a machine or an automobile is critical. This condition is known as **series reliability**. By contrast, for a steel cable consisting of many individual wires, the reliability of each individual wire is not as critical because the cable consists of many wires. This condition is known as **parallel reliability**. The parallel reliability concept is important in the design of backup systems, which permit a product to continue functioning in the event that one of its components fails. Electrical or hydraulic systems in an aircraft, for example, typically are backed up by mechanical systems. Such systems are called *redundant systems*.

Predicting reliability is an important science and 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 may be defined as the capability of a particular manufacturing process to operate predictably and smoothly over time. Thus, it is implicit that there will be no deterioration in performance, which otherwise would require downtime on machines, interrupt production, and result in economic loss.

36.10 Nondestructive Testing

Nondestructive testing (NDT) is carried out in such a manner that product integrity and surface texture remain unchanged. Nondestructive-testing techniques generally require considerable operator skill, and interpreting test results accurately may be difficult because the observations can be subjective. However, the use of computer graphics and other enhancement techniques have significantly reduced the likelihood of human error. Current systems have various capabilities for data acquisition and for qualitative and quantitative inspection and analysis.

Listed here are the basic principles of major nondestructive-testing techniques.

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.

This 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 that 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 defect then becomes a magnet due to flux leakages where 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 by this method,

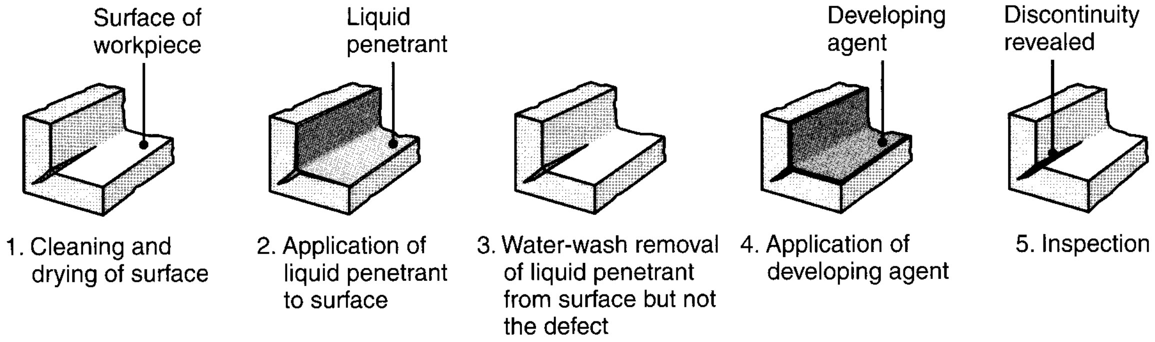


FIGURE 36.9 Sequence of operations for liquid-penetrant inspection to detect the presence of cracks and other flaws in a workpiece. *Source*: ASM International.

provided that they are not deep. The ferromagnetic particles may be colored with pigments for better visibility on metal surfaces.

The magnetic fields can be generated with either direct current a alternating current, and yokes, bars, and coils. Subsurface defects can be detected best with direct current. 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 it may be stationary.

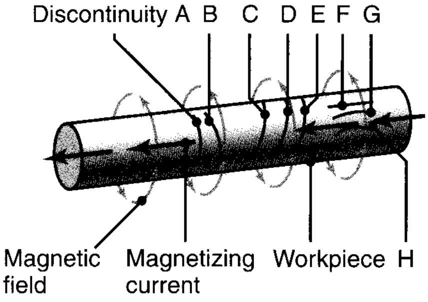


FIGURE 36.10 Schematic illustration of magnetic-particle inspection of a part with a defect in it. Cracks that are in a direction parallel to the magnetic field (such as discontinuity A) would not be detected, whereas the others shown would. Discontinuities F, G, and H are the easiest to detect. *Source*: ASM International.

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* (see Section 3.7) using materials such as quartz, lithium sulfate, or various ceramics. Most inspections are carried out at a frequency range from 1 to 25 MHz. Couplants are used to transmit the ultrasonic waves from the transducer to the test piece; typical couplants are water, oil, glycerin, and grease.

The ultrasonic-inspection method has high penetrating power and sensitivity. It can 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** 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 (see also Section 21.5.4).

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 one taps walls, desktops, or countertops in various locations with a finger or a hammer and listens to the sound emitted. Vitrified grinding wheels (Section 26.2) are tested in a similar manner (*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. However, the results depend on the geometry and mass of the part, so a reference standard is necessary for identifying flaws.

Radiography. *Radiography* uses X-ray inspection to detect such internal flaws as cracks and porosity. The technique detects differences in density within a part. For example, on an X-ray film, the metal surrounding a defect is typically denser and, hence, shows up as lighter than, the flaws. This effect is similar to the way bones and teeth show up lighter than the rest of the body on X-ray films. The source of 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

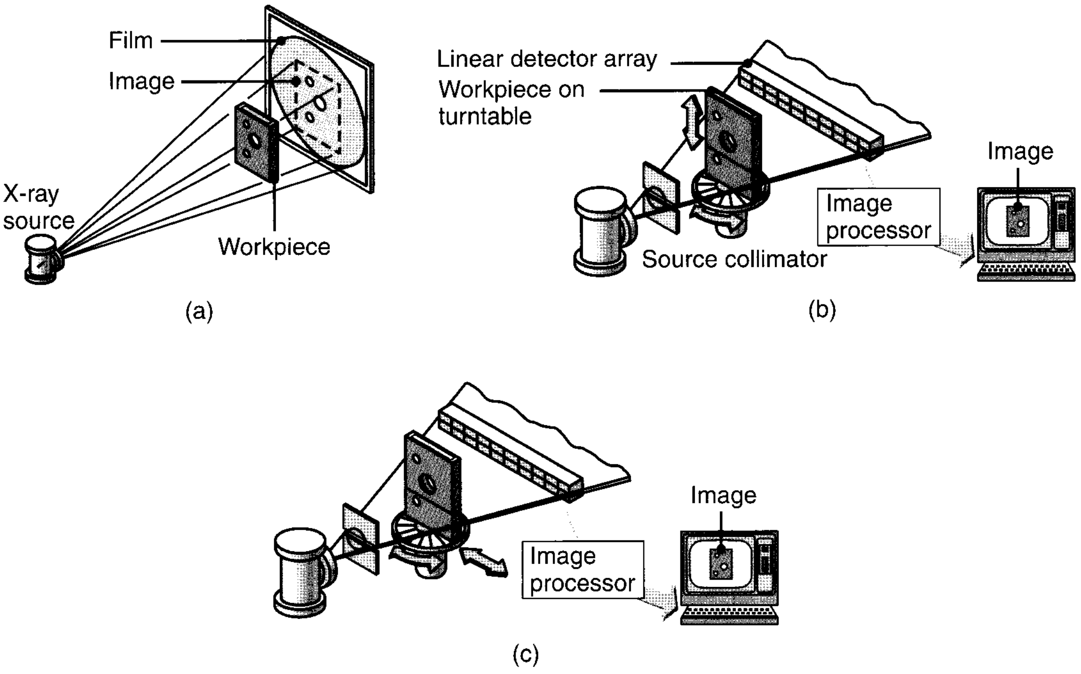


FIGURE 36.11 Three methods of radiographic inspection: (a) conventional radiography, (b) digital radiography, and (c) computed tomography. *Source*: ASM International.

images very quickly, and fluoroscopy 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 techniques are as follows:

- In **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 Figs. 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** is based on the same system as described for digital radiography, except 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 view the object precisely.
- **Computer-assisted tomography (CAT scan)** is based on the same principle and is used widely in medical practice and diagnosis.

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 ranging from 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 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 geometry of the part being inspected. Parts must be conductive electrically, and flaw depths

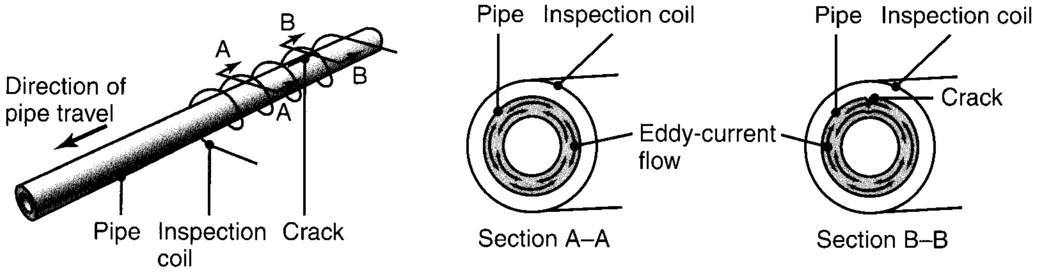


FIGURE 36.12 Changes in eddy-current flow caused by a defect in a workpiece. *Source*: ASM International.

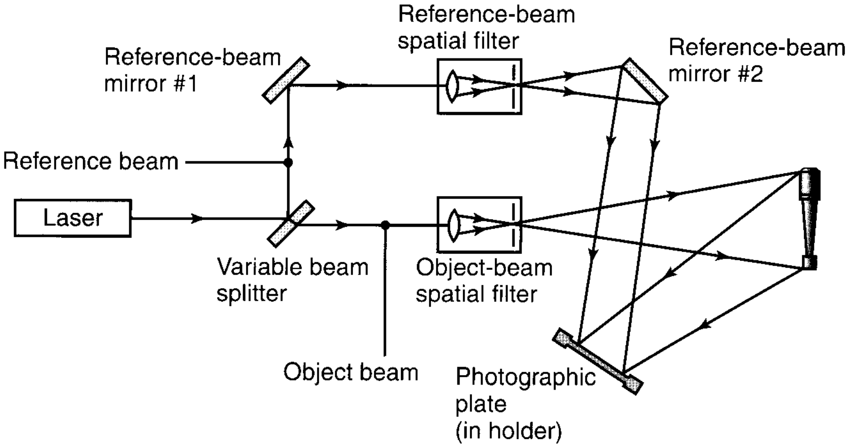


FIGURE 36.13 Schematic illustration of the basic optical system used in holography elements in radiography for detecting flaws in workpieces. *Source*: ASM International.

detected usually are limited to 13 mm. The technique requires the use of a standard reference sample to set the sensitivity of the tester.

Thermal Inspection. *Thermal inspection* involves using contact- or noncontact-type heat-sensing devices that detect temperature changes. Defects in the workpiece (such as cracks, debonded regions in laminated structures, and poor joints) 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 1°C. *Thermometric inspection* utilizes devices such as thermocouples, radiometers, and pyrometers, and sometimes meltable materials, such as wax-like 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. In response to double- and multiple-exposure techniques while the part is being subjected to external forces or time-dependent variations, changes in the images reveal defects in the part.

- 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 workpiece and two ultrasonic transducers (one for the object beam and the other for the reference beam) are immersed in a water-filled tank. The 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 usually is 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 via *destructive-testing* methods no longer maintains its integrity, original shape, or surface characteristics. The 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 are the speed testing of grinding wheels to determine their bursting speed 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. However, microhardness tests 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). Generally, most glasses, highly heat treated metals, and ceramics are notch sensitive. Consequently, a small indentation produced by the indenter can reduce their strength and toughness significantly.

36.12 Automated Inspection

Traditionally, individual parts and assemblies of parts 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 kept and used on the basis of having a certain acceptable deviation from the standard.

In contrast, *automated inspection* uses a variety of sensor systems that monitor the relevant parameters *during* the manufacturing operation (**online 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 is unnecessary. Parts also may be inspected immediately after they are produced (**in-process inspection**).

The development of accurate sensors and advanced computer-control systems has enabled automated inspection to be integrated into manufacturing operations (Chapters 37 and 38). Such a system ensures that no part is moved from one manufacturing process to another (e.g., 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, reliability, and dimensional accuracy.

Sensors for Automated Inspection. Continued advances in *sensor technology*, described in Section 37.7, have made online or real-time monitoring of manufacturing processes feasible. Directly or indirectly, and with the use of various *probes*, sensors can detect dimensions, surface roughness, temperature, force, power, vibration, tool wear, and the presence of external or internal defects.

Sensors may operate on the principles of strain gages, inductance, capacitance, ultrasonics, acoustics, pneumatics, infrared radiation, optics, lasers, or various electronic gages. Sensors 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 the rapid online adjustment of any processing parameter, thus resulting in the production of parts that consistently are within specified standards of dimensional tolerance and quality. For example, such systems already are standard equipment on many metal-cutting machine tools and grinding machines 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.
- Statistical quality control and process control are indispensable in modern manufacturing. They are particularly important in the production of interchangeable parts and in the 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 control in manufacturing.
- Several nondestructive and destructive testing techniques (each of which has its own applications, advantages, and limitations) are available for inspection of completed parts and products.
- The traditional approach of inspecting the part or product after it is manufactured has been replaced largely by online and 100% inspection of all parts and products being manufactured.

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

Aft, L.S., **Fundamentals of Industrial Quality Control**, 3rd ed., Addison-Wesley, 1998.

Bendell, A., **Taguchi Methods**, Springer, 2007.

Bentley, J.P., **Introduction to Reliability and Quality Engineering**, 2nd ed., Prentice-Hall, 1999.

Besterfield, D.H., **Quality Control**, 8th ed., Prentice Hall, 2008.

Block, M.R., and Marash, I.R., **Integrating ISO 14001 into a Quality Management System**, 2nd ed., ASQ Quality Press, 2001.

Bothe, D.R., **Measuring Process Capability: Techniques and Calculations for Quality and Manufacturing Engineers**, McGraw-Hill, 1997.

Bray, D.E., and Stanley, R.K., **Nondestructive Evaluation: A Tool in Design, Manufacturing, and Service**, rev. ed., CRC Press, 1997.

Cartz, L., **Nondestructive Testing**, ASM International, 1995.

Deming, W.E., **Out of the Crisis**, MIT Press, 1986.

DeVor, R.E., Chang T., and Sutherland, J.W., **Statistical Quality Design and Control**, 2nd ed., Prentice Hall, 2006.

Fowlkes, W.Y., and Creveling, C.M., **Engineering Methods for Robust Product Design: Using Taguchi Methods in Technology and Product Development**, Addison-Wesley, 1995.

Imai, M., **Gemba Kaizen: A Commonsense, Low-Cost Approach to Management**, McGraw-Hill, 1997.

Joglekar, A., **Statistical Methods for Six Sigma: In R&D and Manufacturing**, Wiley, 2004.

Juran, J.M., and Godfrey, A.B. (eds.), **Juran's Quality Handbook**, 5th ed., McGraw-Hill, 2000.

Kales, P., **Reliability: For Technology, Engineering, and Management**, Prentice Hall, 1997.

Kear, F.W., **Statistical Process Control in Manufacturing Practice**, Marcel Dekker, 1997.

Levinson, W.A., and Tumbelty, F., **SPC Essentials and Productivity Improvement: A Manufacturing Approach**, American Society for Quality Control, 1996.

Montgomery, D.C., **Introduction to Statistical Quality Control**, 6th ed., Wiley, 2008.

Park, S., **Robust Design and Analysis for Quality Engineering**, Chapman & Hall, 1997.

Quensenberry, C.P., **SPC Methods for Quality Improvement**, Wiley, 1997.

Raj, B., Jaykumar, T., and Thavasimuthu, M. (eds.), **Practical Non-Destructive Testing**, 2nd ed., ASM International, 2002.

Ross, P.J., **Taguchi Techniques for Quality Engineering**, 2nd ed., McGraw-Hill, 1996.

Ryan, T.P., **Modern Experimental Design**, Wiley, 2007.

Smith, G.M., **Statistical Process Control and Quality Improvement**, 5th ed., Prentice Hall, 2004.

Taguchi, G., **Introduction to Quality Engineering**, UNIPUB/Kraus International, 1986.

———, **Taguchi on Robust Technology Development: Bringing Quality Engineering Upstream**, ASME Press, 1993.

Tozawa, B., and Bodek, N., **The Idea Generator: Quick and Easy Kaizen**, PCS Press, 2001.

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. Explain why control charts are developed. How are they used?

36.7. What do control limits indicate?

36.8. Define process capability. How is it used?

36.9. What is acceptance sampling? Why was it developed?

36.10. Explain the difference between series and parallel reliability?

36.11. What is meant by six sigma quality?

36.12. Explain the difference between (a) probability and reliability and (b) robustness and reliability?

QUALITATIVE PROBLEMS

36.13. Explain why major efforts are continually being made to build quality into products.

36.14. Give examples of products for which 100% sampling is not possible or feasible.

36.15. What is the consequence of setting lower and upper specifications closer to the peak of the curve in Fig. 36.4?

36.16. Identify several factors that can cause a process to become out of control.

36.17. Describe situations in which the need for destructive testing techniques is unavoidable.

36.18. Which of the nondestructive inspection techniques are suitable for nonmetallic materials? Why?

36.19. What are the advantages of automated inspection? Why has it become an important part of manufacturing engineering?

36.20. Why is reliability important in manufacturing engineering? Give several examples.

36.21. Give examples of the acoustic-impact inspection technique other than those given in the chapter.

36.22. Explain why *GO* and *NOT GO* gages (see Section 35.4.4) are incompatible with the Taguchi philosophy.

36.23. Describe your thoughts regarding the contents of Table 36.1.

36.24. Search the technical literature and give examples of robust design in addition to that shown in Fig. 36.1.

▶**36.25.** What is a Taguchi loss function? What is its significance?

QUANTITATIVE PROBLEMS

36.26. 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.

▶**36.27.** Assume that in Example 36.3 the number of samples was 8 instead of 10. Using the top half of the data in Table 36.4, recalculate the control limits and the standard deviation. Compare your observations with the results obtained by using 10 samples.

▶**36.28.** Calculate the control limits for averages and ranges for (a) number of samples = 8 (b) $\bar{\bar{x}} = 65$, and (c) $R = 6$.

▶**36.29.** Calculate the control limits for (a) number of samples = 6, (b) $\bar{\bar{x}} = 36.5$, and (c) $\text{UCL}_R = 5.75$.

▶**36.30.** In an inspection with a sample size of 12 and a sample number of 40, it was found that the average range was 14 and the average of averages was 80. Calculate the control limits for averages and for ranges.

▶**36.31.** 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

▶**36.32.** The average of averages of a number of samples of size 9 was determined to be 124. The average range was 17.82 and the standard deviation was 4. The following measurements were taken in a sample: 121, 130, 125, 130, 119, 131, 135, 121, and 128. Is the process in control?

SYNTHESIS, DESIGN, AND PROJECTS

36.33. 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?

36.34. 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.

36.35. Survey the available technical literature, contact various associations, and prepare a comprehensive table concerning the life expectancy of various consumer products.

36.36. 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.

36.37. Name several material and process variables that can influence product quality in metal (a) casting, (b) forming, and (c) machining.

36.38. Identify the nondestructive techniques that are capable of detecting internal flaws and those which detect external flaws only.

36.39. Explain the difference between in-process and postprocess inspection of manufactured parts. What trends are there in such inspections? Explain.

36.40. 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.

Manufacturing in a Competitive Environment

PART IX

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 can 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, we can *integrate* and *optimize* various functions and activities that, for a long time, had been separate and distinct entities.

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 as adaptive control, industrial robots, sensor technology, material handling and movement, and assembly systems and how they are all 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 and various enabling technologies as an aid in such activities as design, engineering, manufacturing, and process planning. The chapter also includes discussions on various enabling technologies, such as adaptive control, indusrial robots, sensor technology, 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. Included also is the new concept of holonic manufacturing and the role and significance of communication networks.

The purpose of Chapter 40 is to highlight the importance of the numerous and often complex factors and their interactions that have a major effect 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 economical production; process capabilities; and costs involved, including those of machinery, tooling, and labor.

Automation of Manufacturing Processes

CHAPTER 37

- This chapter describes automation in all aspects of manufacturing processes and operations, by which parts are produced reliably, economically, and accurately at high production rates.
- The chapter 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 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 ends with a comprehensive description of fixturing and assembly systems in manufacturing and their design considerations.

EXAMPLES:

CASE STUDY:

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 dimensions and characteristics. Moreover, each time a different product had 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 on the machines. Furthermore, 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, productivity basically measures operating efficiency. An efficient operation makes optimum use of all resources, such as materials, energy, capital, labor, machinery, and available technology. 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 improving productivity, the important elements have been *mechanization, automation*, and *control* of manufacturing equipment and systems. **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, the worker would still be directly involved in a particular operation and would continually check a machine's overall performance. Consider, for example, the following situations: (a) A cutting tool wears or fractures during a machining operation, (b) a part is overheated during heat treatment, (c) the surface finish of a part begins to deteriorate during grinding, or (d) dimensional tolerances and springback become too large in sheet-metal forming. In all these situations, the operator must intervene and change one or more of the relevant process parameters and machine settings—a task that requires considerable 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. Although there is no precise definition, *automation* generally means "the methodology and system of operating a machine or process by highly *automatic* means" (from the Greek word *automatos*, meaning "self-acting"). Rapid advances in automation and the development of several enabling technologies were then made possible, largely through advances in **control systems**, with the help of increasingly powerful and sophisticated computers and software.

This chapter follows the outline shown in Fig. 37.1. First, it 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 improve handling efficiency.

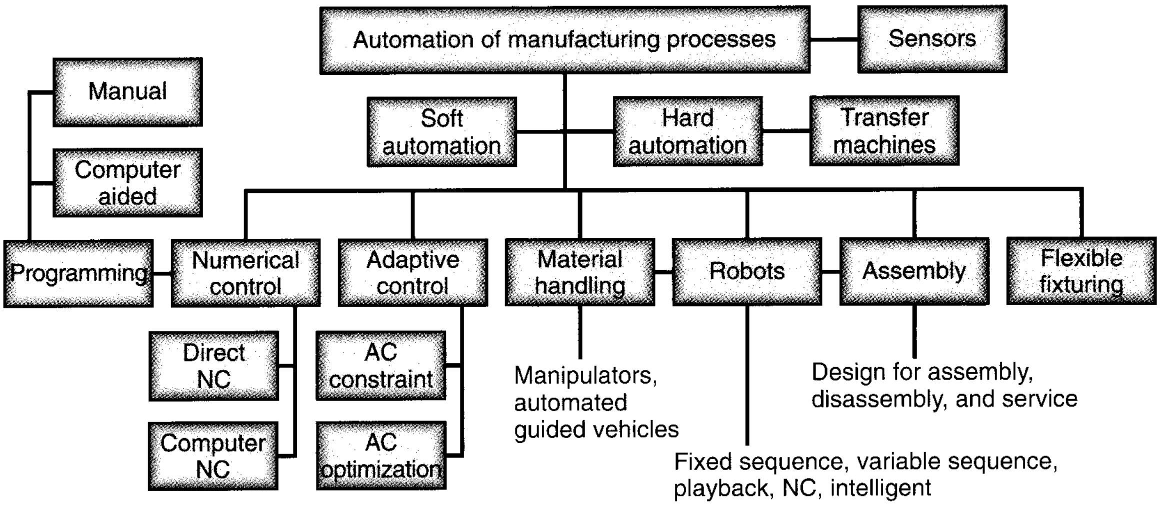


FIGURE 37.1 Outline of topics covered in Chapter 37.

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** are covered as well. These methods enable us to take full advantage of advanced manufacturing technologies, particularly flexible manufacturing systems. The chapter also includes a discussion of the guidelines for **design, for assembly, for disassembly, and for service**, with specific recommendations to improve the efficiency of each of these operations. The final topic of the chapter describes the **economics** of the equipment, processes, and operations covered.

37.2 Automation

Although there have been various definitions, *automation* generally is 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 manufacturing processes and operations. Table 37.1

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

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.

Automation is an *evolutionary* rather than a revolutionary concept. In manufacturing plants, for example, it has been implemented especially in the following basic areas of activity:

- **Manufacturing processes:** Machining, forging, cold extrusion, casting, powder metallurgy, and grinding operations.
- **Material handling and movement:** Materials and parts in various stages of completion (works in progress) 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 manufacture (*in-process inspection*).
- **Assembly:** Individually manufactured parts and components are assembled automatically into subassemblies and then 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*, the *Second Industrial Revolution* having begun in the mid 1950s, with advances in many areas) that automation began to be introduced in the production of goods. 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* products, best represented by the automobile industry, which produced passenger cars at a high production rate and low cost.

The major breakthrough in automation began with numerical control (NC) of machine tools in the early 1950s. 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.

Manufacturing involves various levels of automation, depending on the processes used, the products to be made, and production volumes. Manufacturing systems, in order of increasing automation, include the following classifications (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* (Section 37.3), 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 (Section 37.6).

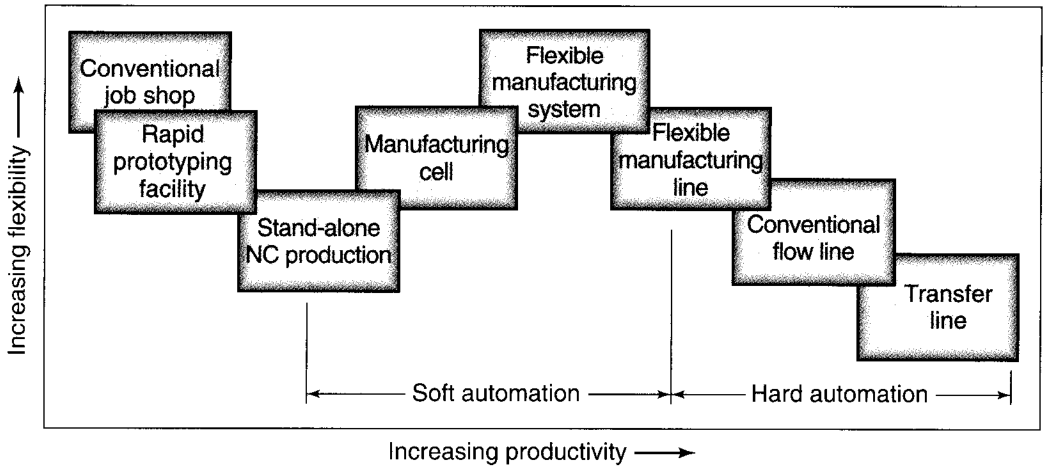


FIGURE 37.2 Flexibility and productivity of various manufacturing systems. Note the overlap between the systems; it is due to the various levels of automation and computer control that are possible in each group. (See also Chapter 39 for details.)

- **Flexible manufacturing systems**: Described in Section 39.3, these systems use computer control of all aspects of manufacturing, the simultaneous incorporation of a number of manufacturing cells, and automated material-handling systems.
- **Flexible manufacturing lines**: These lines organize computer-controlled machinery 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. The manufacturing line is designed with limited or no flexibility, since the goal is to produce a single part.

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 also 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*, described in 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*.

- **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 range from 10 to 100, and the equipment used consists of general-purpose machines and machining centers.

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, missiles, special machinery, dies, jewelry, and orthopedic implants
Batch or high-volume	5000–100,000	Trucks, agricultural machinery, jet engines, diesel engines, computer components, and sporting goods
Mass production	100,000 and over	Automobiles, appliances, fasteners, and food and beverage containers

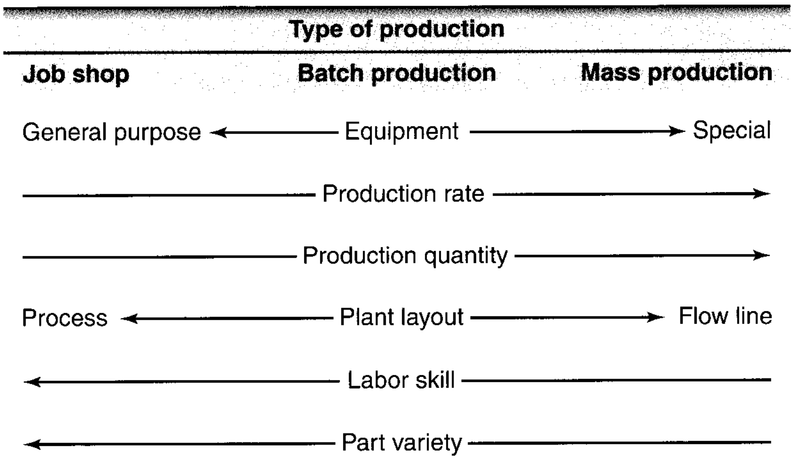


FIGURE 37.3 General characteristics of three types of production methods: job shop, batch, and mass production.

- **Batch production** usually involves lot sizes between 100 and 5000. It utilizes machinery similar to that used for small-batch production, but with specially designed fixtures for higher productivity.
- **Mass production** involves quantities often over 100,000. It requires special-purpose machinery (called **dedicated machines**) and automated equipment for transferring materials and parts in progress. Although the machinery, equipment, and specialized tooling are expensive, both the labor skills required and the labor costs are relatively low. These production systems are organized for a specific type of product; consequently, they lack flexibility.

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 requires at least 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 automated
- Level of skill in the available workforce
- Reliability and maintenance problems that may be associated with automated systems
- Economics of the whole operation.

Because automation generally involves high initial equipment cost and requires a knowledge of operation and maintenance principles, a decision about the implementation of even low levels of automation must involve a careful study of the true 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 skill level required is lower. Conversely, the higher the level of skill available in the workforce, the lower is the need for automation, as is generally the case in developing countries, which often have an abundance of low-skilled labor.

37.2.4 Hard Automation

In *hard automation*, or **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 changed, these machines are specialized and hence lack flexibility. Because such machines are expensive to design and build, their economical use requires the 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, they can easily be regrouped for producing a different part and thus have some adaptability and flexibility.

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, to enable various operations to be performed on the part. Workpieces are transferred by such methods as (a) rails, along which the parts (which usually are placed on *pallets*) are pushed or pulled by various mechanisms (Fig. 37.4a), (b) rotary indexing tables (Fig. 37.4b), and (c) overhead conveyors.

Transfer Lines. A *transfer line,* or **flow line,** in a very large system for producing cylinder heads for engine blocks consisting of a number of transfer machines, is shown in Fig. 37.5. This system is capable of producing 100 cylinder heads per hour. Note the various machining operations performed: milling, drilling, reaming,

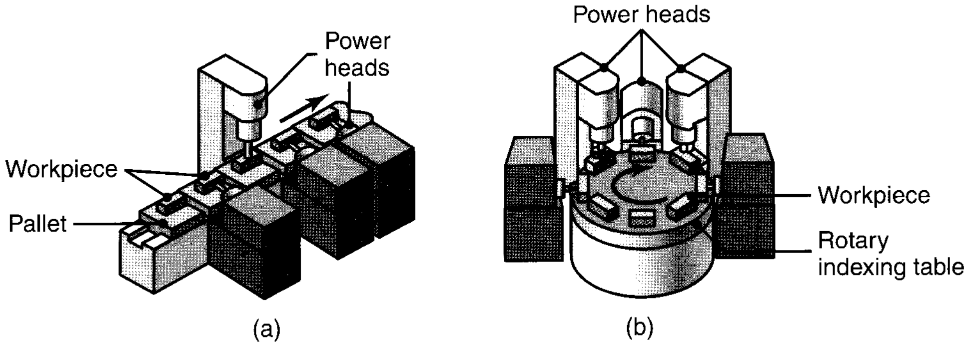


FIGURE 37.4 Two types of transfer mechanisms: (a) straight rails and (b) circular or rotary patterns.

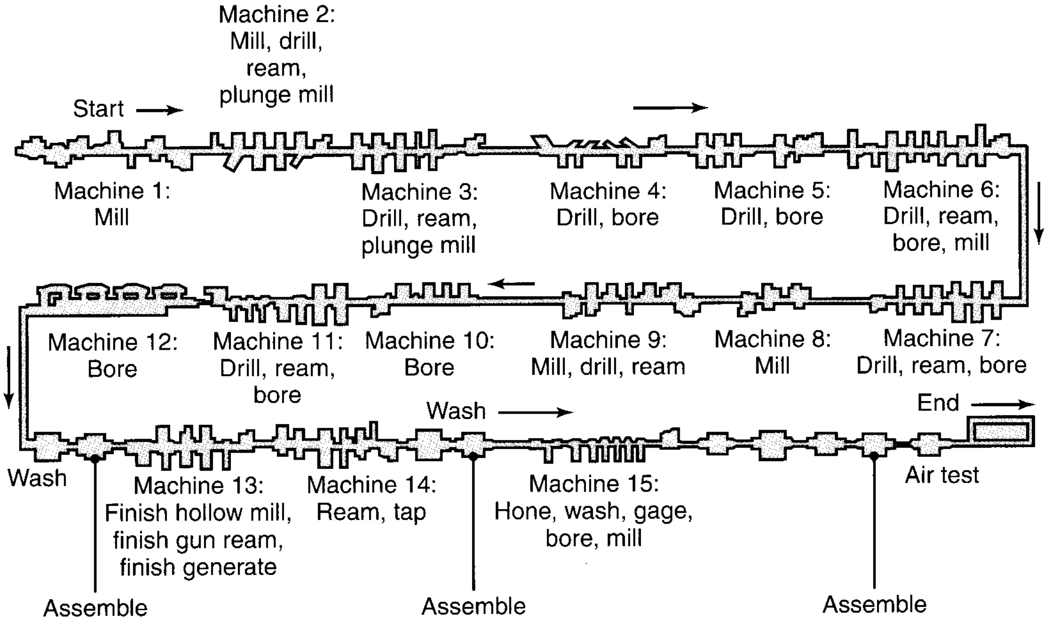


FIGURE 37.5 A large transfer line for producing engine blocks and cylinder heads. *Source*: Courtesy of Ford Motor Company.

boring, tapping, honing, washing, and gaging. The weight and shape of the workpieces influence the arrangement of the individual machines, which is 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. In *soft automation*, also called **flexible** or **programmable automation**, greater flexibility is achieved through the use of computer control of the machine and of its functions; thus, it can produce parts having complex shapes. Soft automation is important because the machines can be reprogrammed easily and readily to produce a part that has a shape or dimensions different from the one produced just prior to it. Further advances in flexible automation include the extensive use of modern computers, leading to the development of **flexible manufacturing systems**, with high levels of efficiency and productivity, as described in Section 39.3.

37.2.6 Programmable Logic Controllers

The control of a manufacturing process in the proper sequence, 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 popular in new installations because of advances in numerical-control machines, PLCs still represent a very large installation base. *Microcomputers* are now used extensively because they are less expensive than PLCs and are easier to program and to network. PLCs are also used in system control with high-speed digital-processing and communication capabilities.

37.2.7 Total Productive Maintenance and Total Productive Equipment Management

Management and maintenance of a wide variety of machines, equipment, and systems are among the important aspects affecting the 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 (a) equipment breakdown and equipment problems; (b) monitoring and improvement of equipment productivity; (c) the implementation of preventive and predictive maintenance; (d) a reduction in setup time, idle time, and cycle time; (e) the full utilization of machinery and equipment and the improvement of their effectiveness; and (f) a reduction in product defects. As expected, teamwork is an important component of this activity and involves the full cooperation of the 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 in the form of numbers and letters into the system. The system automatically interprets these data and converts them to output signals, which, 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 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, using reference points given by any of the three methods shown in the figure. The operator then 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 machined will be different from others.

Now assume further that, during this production run, the order of processing these parts is changed and, in addition, 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, 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.

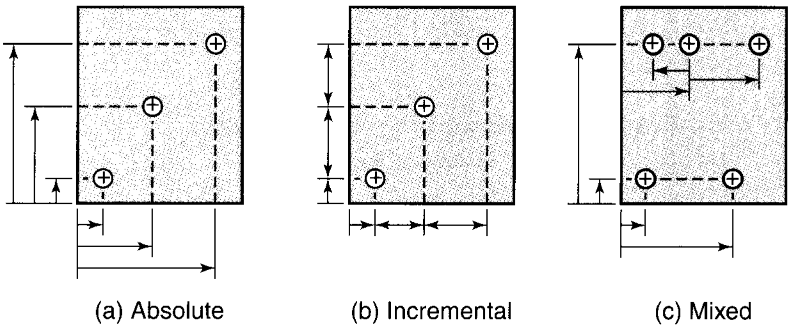

(a) Absolute (b) Incremental (c) Mixed

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.

EXAMPLE 37.1 Historical Origin of Numerical Control

The basic concept behind numerical control apparently was 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) in his attempt to machine complex helicopter blades. The first prototype NC machine was built in 1952 at the Massachusetts Institute of Technology. It was a vertical-spindle, two-axis copy-milling machine retrofitted with servomotors, and the machining operations performed consisted of end milling and face milling 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 experiments, parts were machined successfully, accurately, and repeatedly without 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) yielding 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. Two types of computerized systems were developed—direct numerical control and computer numerical control:

- In **direct numerical control** (DNC), several machines are controlled directly—step by step—by a central mainframe computer. In this system, the operator has 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: If the computer shuts down, 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 system provides large memory and computational facilities and offers flexibility while overcoming the disadvantage of direct numerical control.
- **Computer numerical control** (CNC) is a system in which a control microcomputer 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 used widely today because of the availability of (a) small computers with large memory, (b) low-cost programmable controllers and microprocessors, and (c) program-editing capabilities.

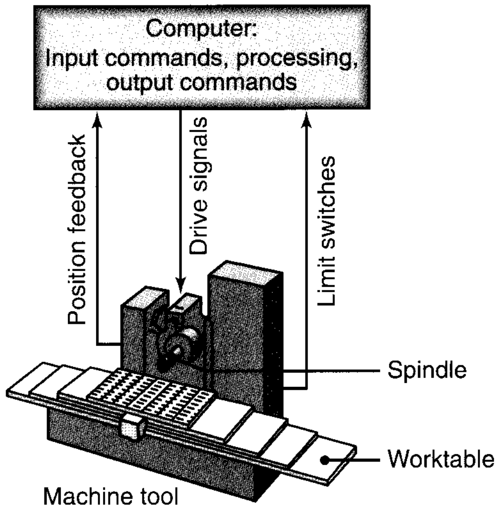


FIGURE 37.7 Schematic illustration of the major components for position control on a numerical-control machine tool.

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

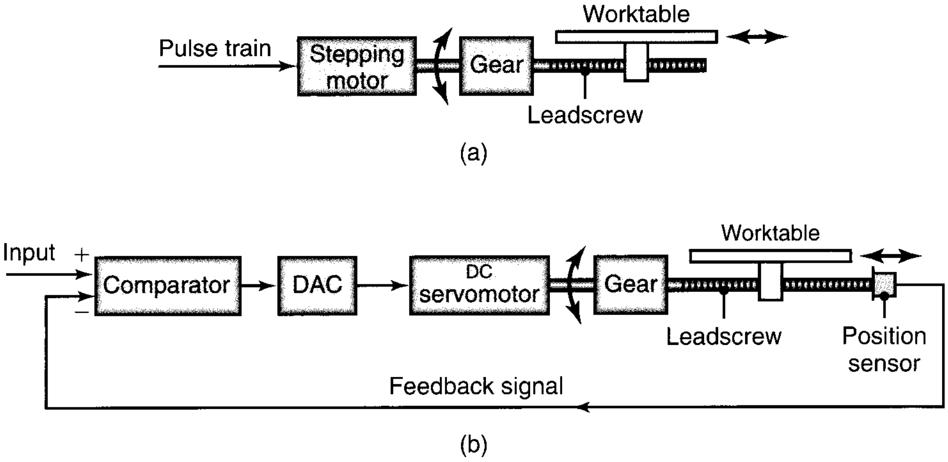


FIGURE 37.8 Schematic illustration of the components of (a) an open-loop and (b) a closed-loop control system for a numerical-control machine. DAC = digital-to-analog converter.

the components involved are as follows:

- **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 on which the workpiece is mounted 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, and the table movements terminate when the proper coordinates are reached.

Position measurement in NC machines is carried out through two methods (Fig. 37.9). In *indirect measuring systems*, *rotary encoders*, or *resolvers*, convert rotary movement to translation. Backlash (the play between two adjacent mating gear teeth), however, can affect measurement accuracy significantly. 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 the more accurate, because the scale is built into the machine and backlash in the mechanisms is not significant.

37.3.3 Types of Control Systems

There are two basic types of control systems in numerical control:

a. 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

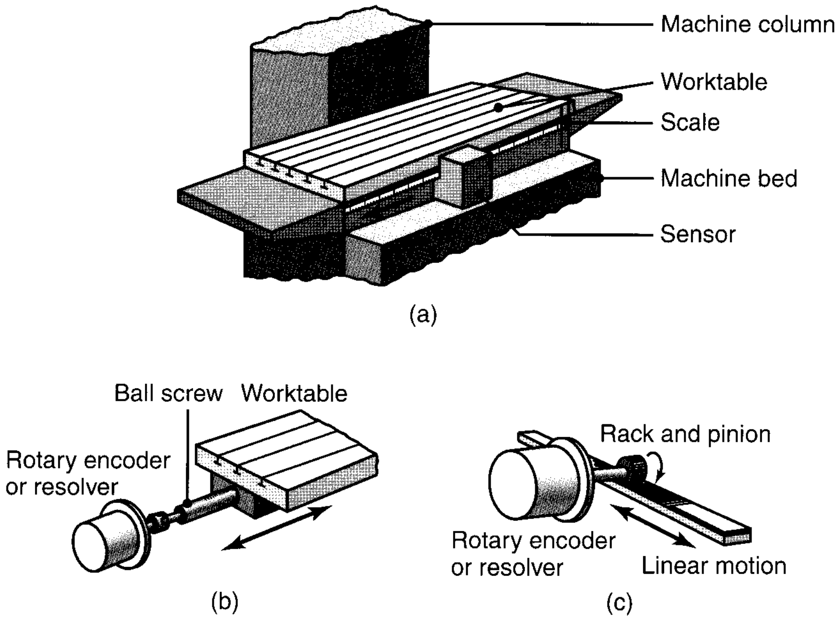


FIGURE 37.9 (a) Direct measurement of the linear displacement of a machine-tool worktable. (b) and (c) Indirect measurement methods.

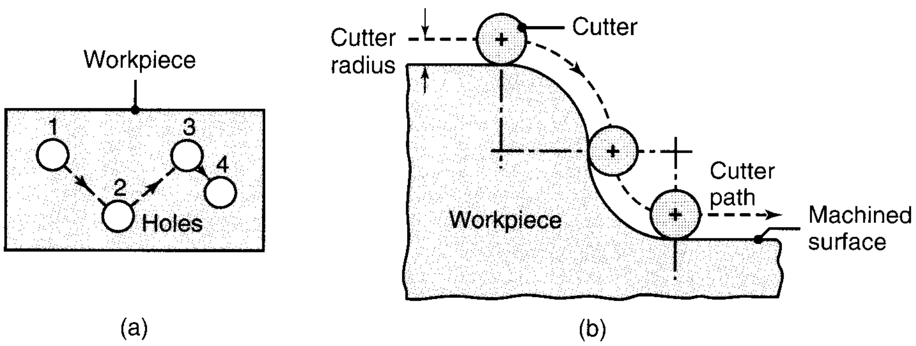


FIGURE 37.10 Movement of tools in numerical-control machining. (a) Point-to-point, in which the drill bit drills a hole at position 1, is retracted and moved to position 2, and so on. (b) Continuous path by a milling cutter. Note that the cutter path is compensated for by the cutter radius. This path also can be compensated for cutter wear.

order to reduce nonproductive time, but then 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).

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.

b. In a **contouring system**, also known as a *continuous-path system*, both the positioning and the operations are performed along controlled paths, but at different velocities. Because the tool acts (i.e., machines, drills, taps, etc.) as it travels along a prescribed path (Fig. 37.10b), accurate control and synchronization

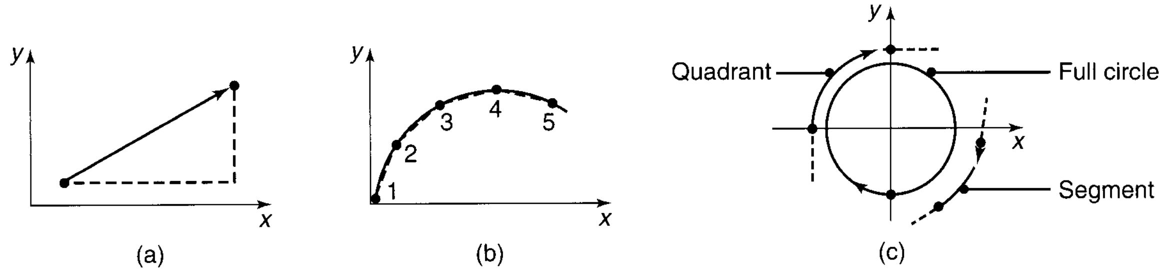


FIGURE 37.11 Types of interpolation in numerical control: (a) linear, (b) continuous path, approximated by incremental straight lines, and (c) circular.

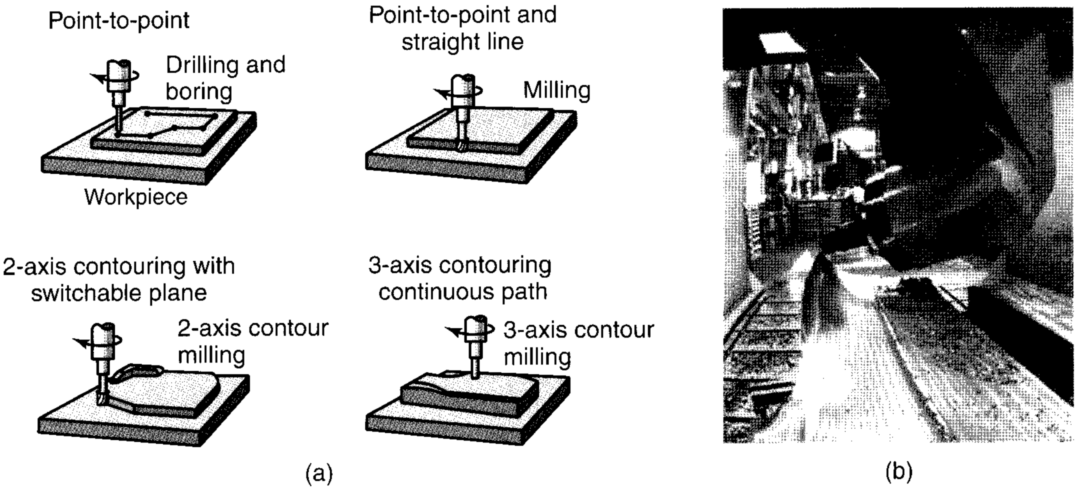


FIGURE 37.12 (a) Schematic illustration of drilling, boring, and milling with various paths. (b) Machining a sculptured surface on a 5-axis numerical-control machine. *Source*: Courtesy of The Ingersoll Milling Machine Co.

of velocities and movements are important. 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. Compensation for different types of tools, for 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. Theoretically, all types of profiles can be produced by this method by making the increments between the points small (Fig. 37.11b), but 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 endpoints, the coordinates of the center of the circle and its radius, and the direction of the tool along the arc.
- In **parabolic interpolation** and **cubic interpolation**, the path is approximated by curves based on higher order mathematical equations. This method is effective

in 5-axis machines and is useful in die-sinking operations, such as for dies for the sheet forming of automotive bodies. These interpolations also are used for the movements of industrial robots, as discussed in Section 37.6.

37.3.4 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.)

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—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 they 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 the relatively high initial cost of the equipment, the need and cost for programming and computer time, and the special maintenance required. Because these machines are complex systems, breakdowns can be costly, so preventive maintenance is essential.

37.3.6 Programming for Numerical Control

A program for numerical control consists of a sequence of directions that causes an NC machine to carry out a certain operation, machining operations of all forms being the most commonly applied. *Programming for NC* may be (a) performed by an internal programming department (most common), (b) done on the shop floor (rare, but minor, changes in G-code are common), or (c) purchased from an outside source.

The program contains instructions and commands: (a) *Geometric instructions* pertain to relative movements between the tool and the workpiece. (b) *Processing instructions* concern spindle speeds, feeds, cutting tools, cutting fluids, and so on. (c) *Travel instructions* pertain to the type of interpolation and to the speed of movement of the tool or the worktable. (d) *Switching instructions* concern the on–off position for coolant supplies, direction or lack of spindle rotation, tool changes, workpiece feeding, clamping, and so on.

Manual part programming begins by calculating the dimensional relationships among the tool, workpiece, and worktable on the basis of the engineering drawings (including CAD) of the part, the operations to be performed, and their sequence. A program sheet is prepared, detailing the information necessary to carry out the particular operation. Then the part program is prepared in accordance with this information. Because the work involved is tedious, time consuming, and uneconomical, manual programming is now used mostly in simple point-to-point applications.

Computer-aided part programming involves special symbolic *programming languages* that determine such coordinate points as corners, edges, and surfaces of the part. A **programming language** is a means of communicating with the computer; it involves the use of symbolic characters. The programmer describes the component to be processed, and the computer converts that description to commands for the NC machine. Several languages are available commercially and are manufacturer specific. Programs for part programming are referred to as macros and are similar to the BASIC programming language.

Complex parts can be machined with the use of graphics-based computer-aided machining programs. A tool path is created in a largely graphic environment that is similar to a CAD program. The machine code (called **G-Code**) is automatically created by the program. This code is valuable for communicating machining instructions to the CNC hardware, but is difficult to edit and troubleshoot without software interpreters. When minor difficulties are encountered, such as (a) using different end-mill diameters than originally programmed or (b) changing the cutting speed to avoid chatter, the machine operator has to to modify the program, which is difficult in G-Code.

In *shop-floor programming*, CNC programming software is used directly on the machine-tool controller. This method allows for higher level geometry and processing information to be sent to the CNC controller. A G-code is then developed by the dedicated computer under control of the machine operator. The advantage is that any changes that are made to the machining program are sent back to the programming group, are stored as a shop-proven design iteration, and can be reused or standardized across parts. The programs must be verified before the machining begins, either by viewing a simulation of the process on a monitor or by first machining the part from an inexpensive material such as aluminum, wood, wax, or plastic.

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 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, we have to steer to avoid potholes by visually and continuously observing the condition of the road directly ahead of the car. Moreover, our body feels the car's rough movements and vibrations; we then react by changing the direction or the speed of the car to minimize the effects of the rough road.

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, the part programmer sets the processing parameters on the basis of the existing knowledge of the workpiece material and relevant data on the particular manufacturing operation. In CNC

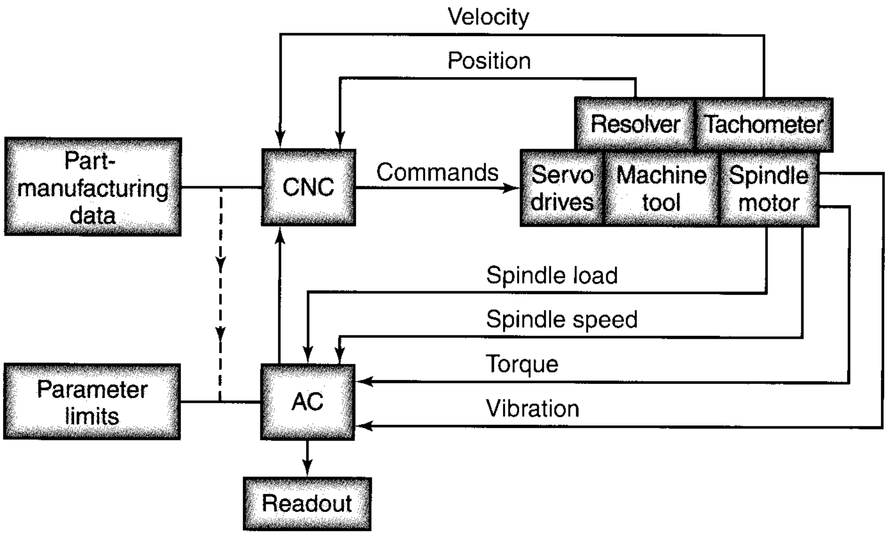


FIGURE 37.13 Schematic illustration of the application of adaptive control (AC) for a turning operation. The system monitors such parameters as cutting force, torque, and vibrations. If these parameters are excessive, it modifies process variables (such as feed and depth of cut) to bring the parameters back to acceptable levels.

machines, on the one hand, these parameters are held constant during a particular process cycle; in AC, on the other hand, the system is capable of automatic adjustments *during* the operation, through closed-loop feedback control (Fig. 37.13). Several adaptive-control systems are available commercially for a variety of applications.

Adaptive Control in Manufacturing. The main purposes of adaptive control in manufacturing are to (a) optimize production rate, (b) optimize product quality, and (c) minimize production cost. The application of AC in manufacturing is particularly important in situations where, for example, quality is not uniform, such as in castings from a poorly controlled process or parts that have been improperly heat treated.

Let's consider a machining operation, such as turning on a lathe (Section 23.2). The adaptive control system senses parameters *in real time*, such as cutting forces, spindle torque, a rise in temperature during machining, tool-wear rate, tool condition, and surface finish produced on the workpiece. The AC system converts this information into commands that then modify the process parameters on the lathe to hold them within certain limits and thus optimize the machining operation.

Those systems which place a constraint on a process variable (such as force, torque, or temperature) are called **adaptive-control constraint** (ACC) systems. Thus, if the thrust force and the cutting force (and hence the torque) increase excessively (due, say, to the presence of a hard region in a casting), the system modifies the cutting speed or the feed in order to lower the cutting force to an acceptable level (Fig. 37.14). Without adaptive control or the direct intervention of the operator (as is the case in traditional machining operations), high cutting forces may cause tool failure or may cause the workpiece to deflect or distort excessively. As a result, workpiece dimensional accuracy and surface finish begin to deteriorate. Those systems which 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.

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

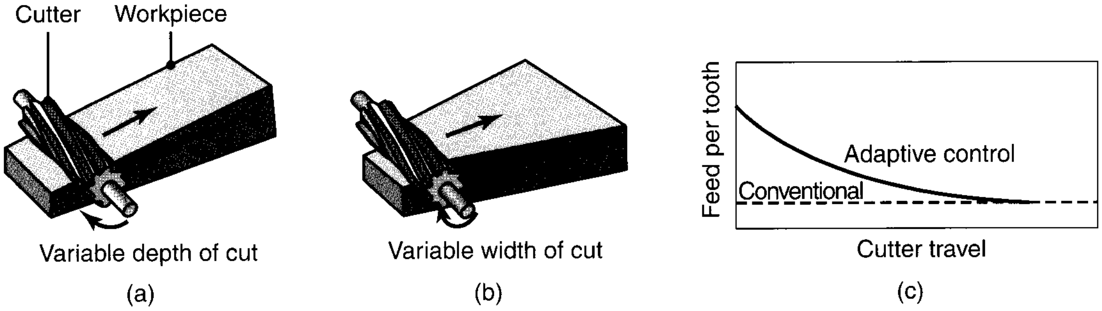


FIGURE 37.14 An example of adaptive control in milling. As the depth of cut (a) or the width of cut (b) increases, the cutting forces and the torque increase. The system senses this increase and automatically reduces the feed (c) to avoid excessive forces or tool breakage in order to maintain cutting efficiency. *Source*: After Y. Koren.

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.

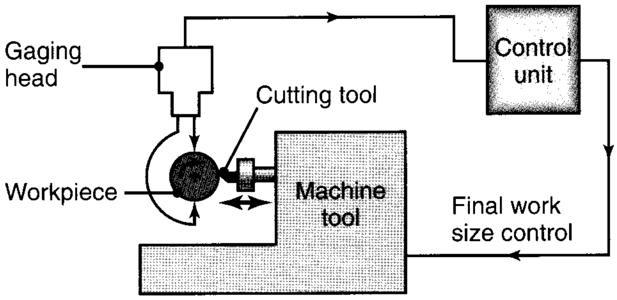


FIGURE 37.15 In-process inspection of a workpiece diameter in a turning operation. The system automatically adjusts the radial position of the cutting tool in order to produce the correct diameter.

For adaptive control to be effective in manufacturing operations, quantitative relationships must 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 (a) to calculate how much of a change in speed or in feed is necessary and (b) to decide whether to increase it or decrease the speed, in order to reduce the tool-wear rate to an acceptable level. The system also should be able to compensate for dimensional changes in the workpiece due to such causes as tool wear and temperature rise (Fig. 37.15).

If the operation is grinding (Chapter 26), the computer software must reflect the desired quantitative relationships among independent process variables (such as wheel and work speeds, feed, and type of wheel) and dependent parameters (such as wheel wear, dulling of abrasive grains, grinding forces, temperature, surface finish, and part deflections). Similarly, in bending sheet metal in a V-die (Section 16.5), data on the dependence of springback (Fig. 16.19) on relevant material and process variables must be stored in the computer memory. It is apparent that, coupled with CNC, adaptive control is a powerful tool in optimizing manufacturing operations.

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. Similarly, tools, molds, dies, and various other equipment and fixtures also are moved around in manufacturing plants. Cutting tools are mounted on lathes, dies are placed in presses or hammers, grinding wheels are mounted on spindles, and parts are mounted on special fixtures for dimensional measurement and inspection.

These materials must be moved either manually or by some mechanical means, and it takes time to transport them from one location to another. **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; furthermore, material handling should be repeatable and predictable. Consider, for example, what happens if a part is loaded improperly into a forging die or onto a machine tool: The consequences of such an action may well be parts that are out of dimensional tolerance, broken dies and tools, and improperly made parts. In addition, such actions can present major safety hazards.

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. For parts requiring multiple operations, equipment should be grouped around the operator or an industrial robot or robots. (See *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.

On the one hand, for small-batch manufacturing operations, raw materials and parts can be handled and transported manually, but this method is generally costly. Moreover, because it involves humans, the practice can be unpredictable, unreliable, and unsafe. On the other hand, 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 to move 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

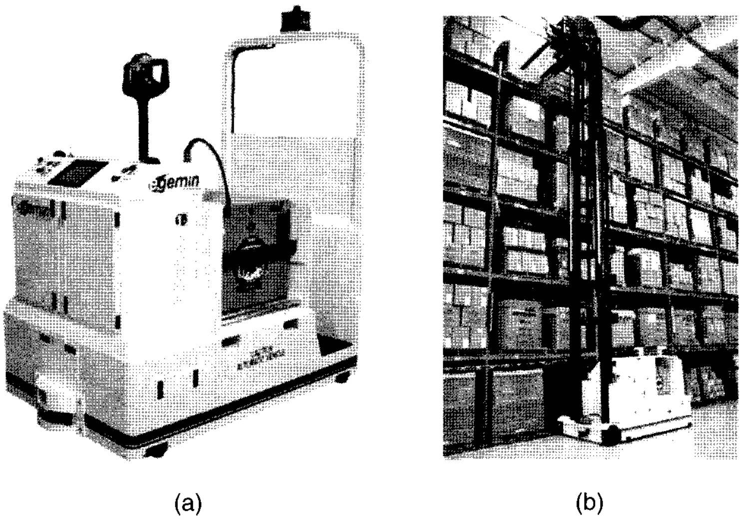


FIGURE 37.16 (a) An automated guided vehicle (Tugger type). This vehicle can be arranged in a variety of configurations to pull caster-mounted cars; it has a laser sensor to ensure that the vehicle operates safely around people and various obstructions. (b) An automated guided vehicle configured with forks for use in a warehouse. *Source*: Courtesy of Egemin, Inc.

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). Automated guided vehicles operate automatically along pathways with in-floor wiring or tapes for optical scanning 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 system.

The movements of AGVs are planned so 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. (See Section 39.5.)

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 follows:

- **Bar coding** is the most widely used system and the least costly.
- **Magnetic strips** constitute the second most common coding system.
- **RF** (radio frequency) **tags** are popular. Although expensive, they do not require the clear line of sight essential in the previous two systems; also, they 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

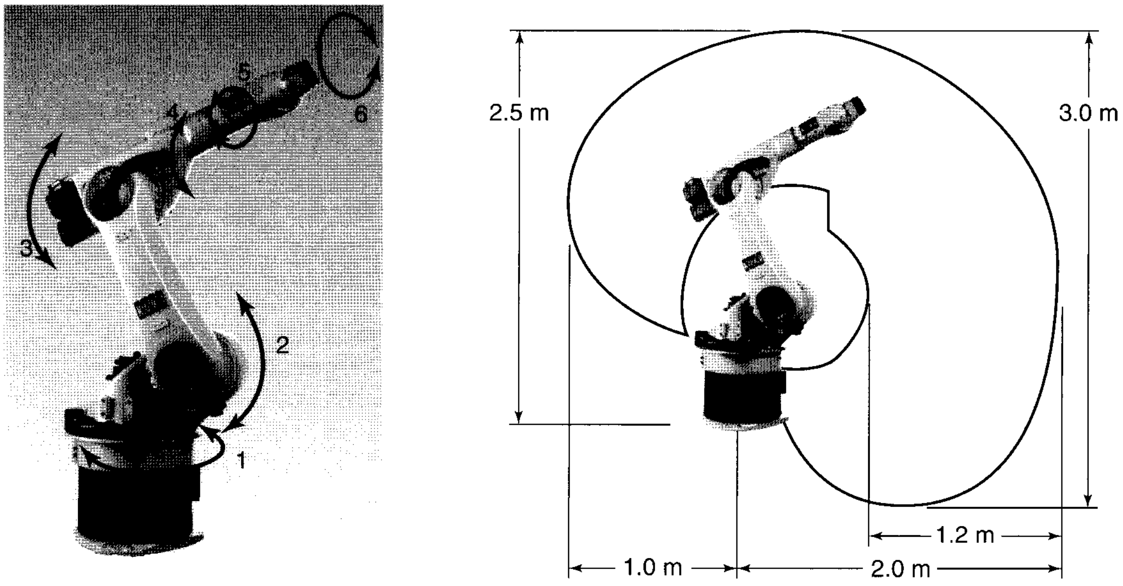
The word **robot** was coined in 1920 by the Czech author K. Čapek in his play *R.U.R.* (Rossum's Universal Robots). It is derived from the word *robota*, meaning "worker." An **industrial robot** has been described by the International Organization for Standardization (ISO) as "a machine formed by a 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 use sensing or adaptation features to take the environment and special circumstances into account.

The first industrial robots were 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 used in hazardous operations, such as the handling of toxic and radioactive materials, and the loading and unloading of hot workpieces from furnaces and in foundries. Simple rule-of-thumb applications for robots are described as the three D's (*dull, dirty*, and *dangerous*; a fourth D—*demeaning*—should also be included) and the three H's (*hot, heavy*, and *hazardous*). As described further in this section, industrial robots are now essential components in 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 unit that provides motions (trajectories) similar to those of a human arm and hand. The



(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 ±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.

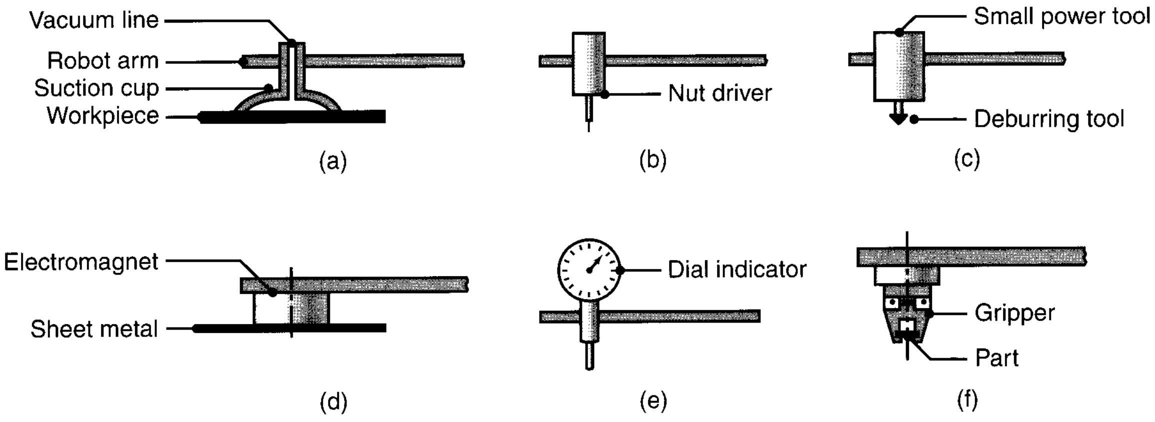


FIGURE 37.18 Types of devices and tools attached to end effectors to perform a variety of operations.

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, and seven degrees of freedom, or *redundant* robots, for special applications also are available.

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
- Spray guns for painting
- Attachments for spot and arc welding and for arc cutting
- Power tools (such as drills, nut drivers, and burrs)
- Measuring instruments.

Mechanical grippers are used the most commonly and are equipped with two or more fingers. **Compliant end effectors** are used to handle fragile materials or to facilitate assembly. These end effectors can use elastic mechanisms to limit the force that can be applied to the workpiece, or they can be designed with a desired 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.

Power Supply. Each motion of the manipulator (linear or rotational) is controlled and regulated by independent actuators that use an electrical, a pneumatic, or a hydraulic power supply. Each source of energy and each type of motor has its own characteristics, 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 this system, 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 in assembly operations for electronic printed circuit boards. 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 as follows (Fig. 37.19):

- **a.** *Cartesian*, or *rectilinear*;
- **b.** *Cylindrical*;
- **c.** *Spherical*, or *polar*;
- **d.** *Articulated*, *revolute*, *jointed*, or *anthropomorphic*.

Robots may be attached permanently to the floor of a plant, they may move along overhead rails (*gantry robots*), or they may 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 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 and sequence of the motions and can repeat them continually without any further action or guidance by the operator. Another type is the **teach pendant**, which 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 reenacted automatically by the robot whenever they are needed.

Numerically Controlled Robot. The numerically controlled robot is programmed and operated much like a numerically controlled machine. The robot 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 controls: point to point and

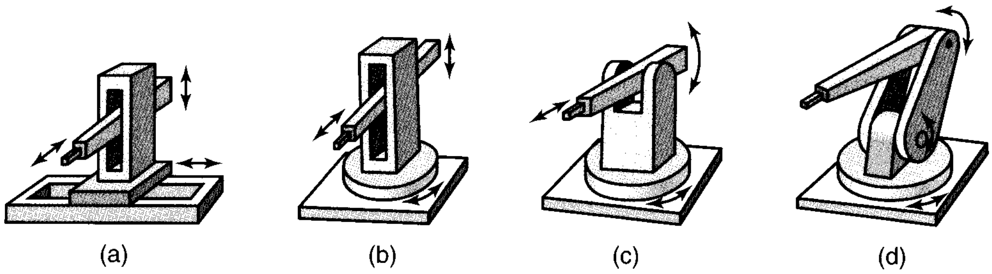


FIGURE 37.19 Four types of industrial robots: (a) Cartesian (rectilinear), (b) cylindrical, (c) spherical (polar), and (d) articulated (revolute, jointed, or anthropomorphic).

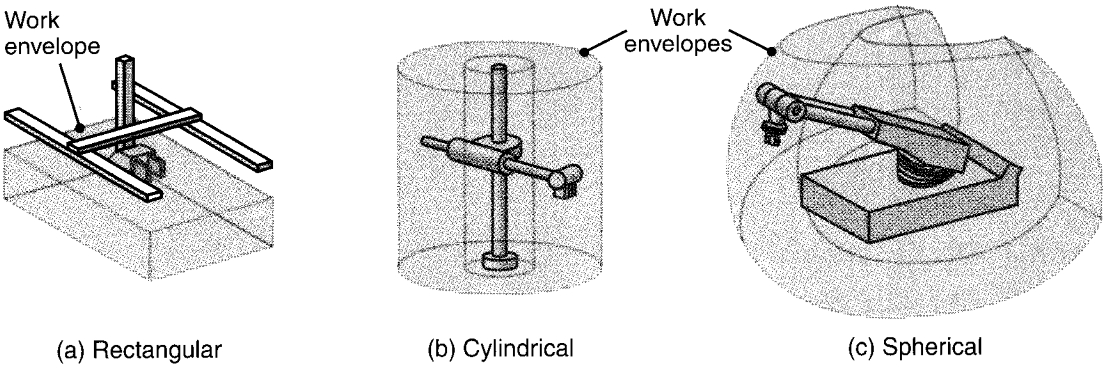


FIGURE 37.20 Work envelopes for three types of robots. The choice depends on the particular application. (See also Fig. 37.17b.)

continuous path. *Point-to-point* robots are easy to program and have a higher load-carrying capacity and a larger **work envelope**, which is the maximum extent or reach of the robot hand or working tool in all directions (Fig. 37.20). *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. It 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. Because its operation is very complex, powerful computers are required to control this type of robot.

Developments in intelligent robots include:

- Behaving more like humans, 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.
- Transporting a part to a machine for further processing or inspection.
- Assembling components into subassemblies or 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 with robots, thereby improving quality and reducing scrap losses. Some examples are (a) casting and molding operations in which molten metal, raw materials, lubricants, 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; (c) forming operations in which parts are loaded and unloaded from presses and various other types of metalworking machinery.
- Spot welding unitizes automobile and truck bodies, producing welds of good quality (Fig. 37.21a). Robots also perform other, similar operations, such as arc welding, arc cutting, and riveting.

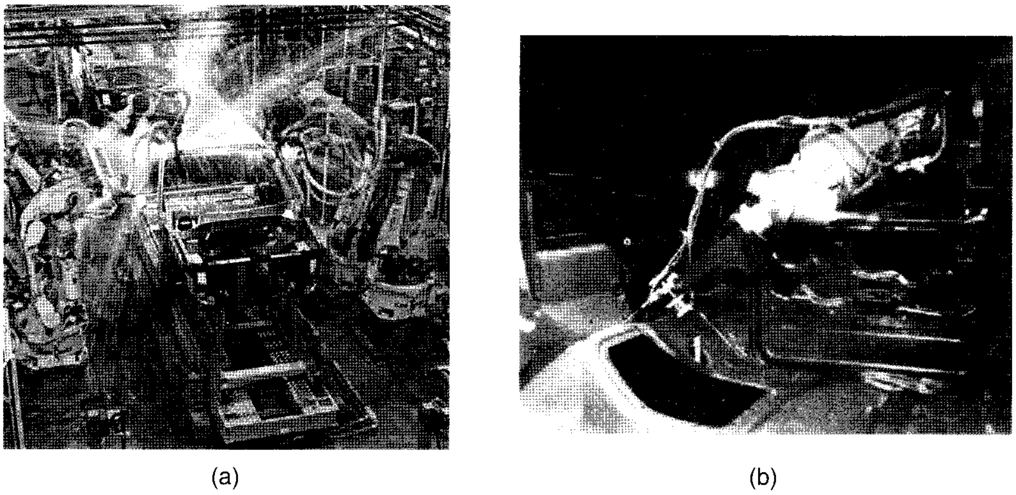

(a) (b)

FIGURE 37.21 Examples of industrial robot applications. (a) Spot welding automobile bodies with industrial robots. (b) Sealing joints of an automobile body with an industrial robot. *Source*: (a) Courtesy of Ford Motor Co. (b) Courtesy of Cincinnati Milacron, Inc.

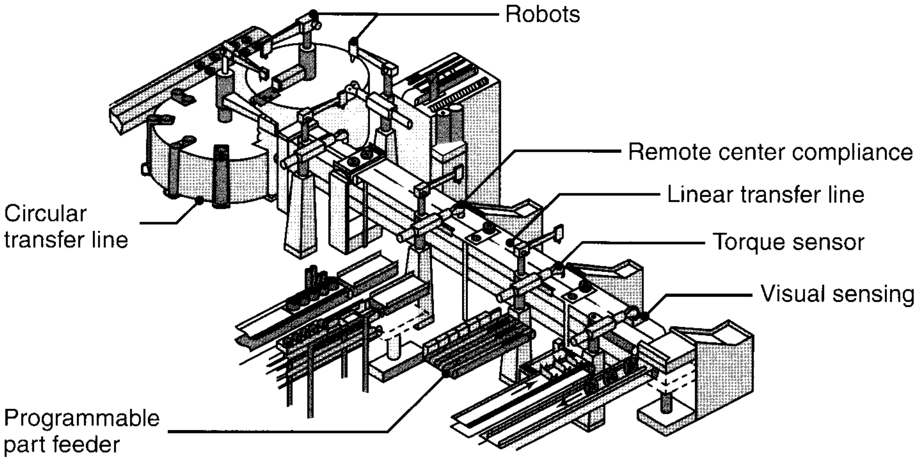


FIGURE 37.22 Automated assembly operations using industrial robots and circular and linear transfer lines.

- Operations such as deburring, grinding, and polishing can be done by using appropriate tools attached to the end effectors.
- Applying adhesives and sealants, as in the automobile frame shown in Fig. 37.21b.
- 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. 37.22).
- Inspection and gaging at speeds much higher than those which can be achieved by humans.

Robot Selection. Factors that influence the selection of robots in manufacturing are as follows:

- Load-carrying capacity
- Work envelope (see Figs. 37.17b and 37.20)
- Speed of movement
- Reliability
- Repeatability
- Arm configuration
- Degrees of freedom
- The control system
- Program memory.

Economics. In addition to the technical factors, cost and benefit considerations are major aspects of robot selection and their use. The increasing availability and reliability, and the reduced costs, of sophisticated, intelligent robots are having a major economic 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 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.

CASE STUDY 37.1 Robotic Deburring of a Blow-molded Toboggan

Roboter Technologie produces high-quality toboggans and car seats made of plastic by injection molding (described in 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, but 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 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 with a KUKA KR-15 robot, which carries out the cutting and deburring operations in a single step while also compensating for the plastic shrinkage (Fig. 37.23).

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 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 40–50 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 39.5).

The robot was successfully applied in a hazardous and dirty environment and has eliminated

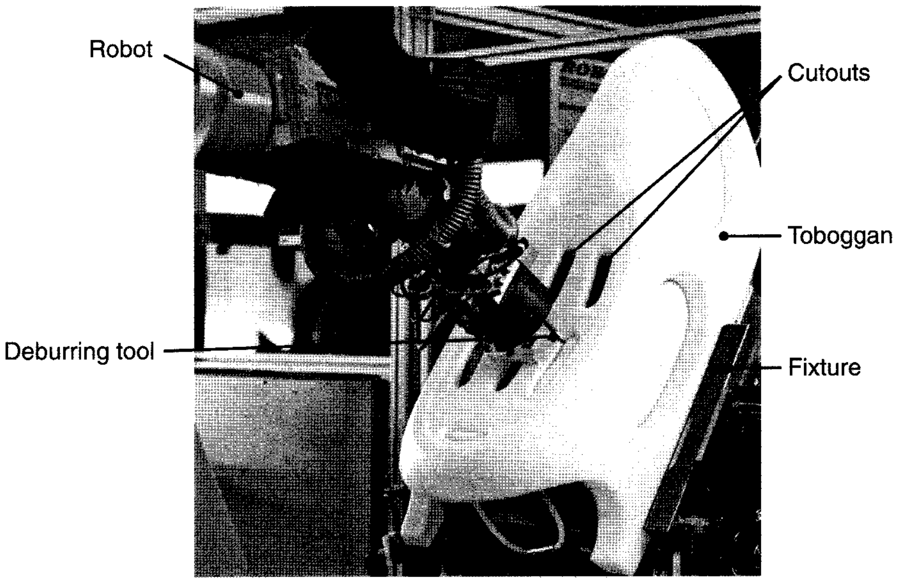


FIGURE 37.23 Robotic Deburring of a Blow-Molded Toboggan. *Source*: Courtesy of Kuka Robotics, Inc., and Roboter Technologie, GmbH. (See also Fig. 26.34.)

tasks that were associated with work-related injuries incurred by fume exposure and stresses on the wrist in manual deburring. In addition, because of the higher quality of deburring and the elimination of rejected parts, the robot cell paid for itself within 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, 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 work-table movements, (b) pressure and temperature gages with automatic shutoff features, and (c) governors on engines to prevent excessive speed of operation. Sensor technology is now an important aspect of manufacturing processes and systems, and is essential for proper data acquisition, monitoring, communication, and computer control of machines and systems (Fig. 37.24). Sensors are essential to the control of intelligent robots and are being developed with capabilities that resemble those of humans. (Such sensors are called *smart sensors.*)

Because they convert one quantity to another, sensors are often referred to as *transducers. Analog sensors* produce a signal, such as voltage, that is

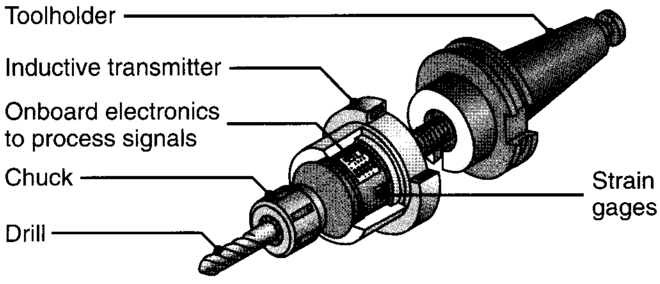


FIGURE 37.24 A toolholder equipped with thrust-force and torque sensors (and thus a smart toolholder), capable of continuously monitoring the cutting operation. Such toolholders are necessary for the adaptive control of manufacturing operations. *Source*: Courtesy of Cincinnati Milacron, Inc.

proportional to the measured quantity. *Digital sensors* have numeric or digital outputs that can be transferred to computers directly. **Analog-to-digital converters** (ADCs) are available for interfacing analog sensors with computers.

37.7.1 Sensor Classification

Sensors that are of greatest interest in manufacturing operations are generally classified as follows:

- **Mechanical** sensors measure such quantities as position, shape, velocity, force, torque, pressure, vibration, strain, and mass.
- **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, chemical, 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 take appropriate action, usually by electromechanical means.

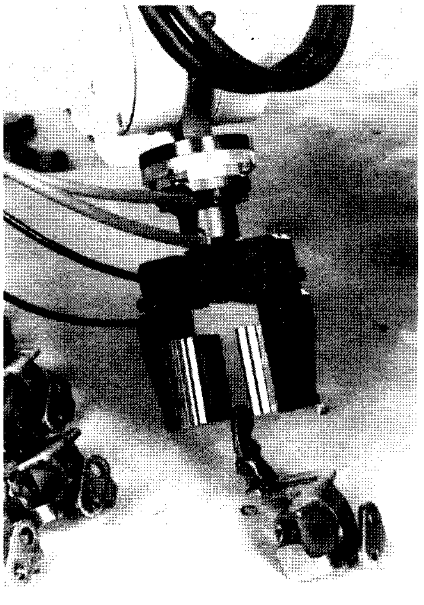


FIGURE 37.25 A robot gripper with tactile sensors. In spite of their capabilities, tactile sensors are used less frequently because of their high cost and their low durability in industrial environments. *Source*: Courtesy of Lord Corporation.

Tactile Sensing. Tactile sensing involves the continuous sensing of variable contact forces, commonly by an array of sensors (Fig. 37.25). Such a system is capable of operating within an arbitrary three-dimensional space. Fragile parts, such as eggs, thin glass bottles, 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).

The force that is sensed is monitored and controlled through closed-loop feedback devices. Compliant grippers with force feedback capabilities and sensory perception are complicated 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, we can sense rope slipping from our hands.

Visual Sensing. In visual sensing, cameras optically sense the presence and shape of the object (Fig. 37.26). This system, called *machine vision* or *computer vision*, utilizes a microprocessor that processes the image, usually within less than one second. The image is then measured and the

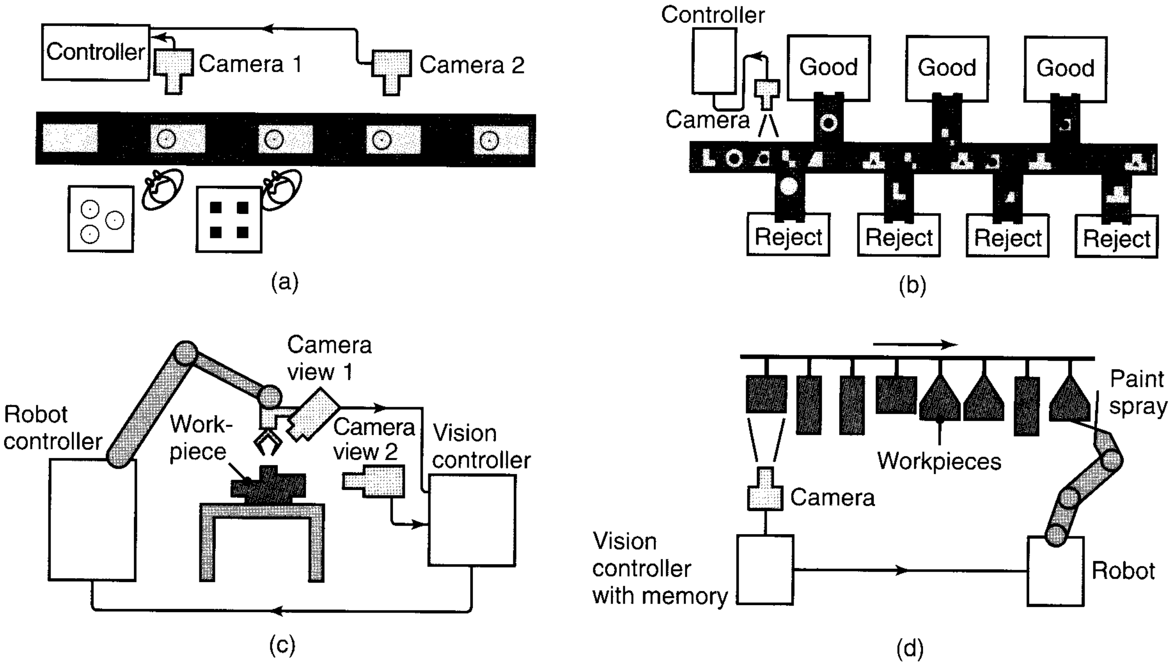


FIGURE 37.26 Examples of machine-vision applications. (a) In-line inspection of parts. (b) Identification of parts with various shapes and inspection and rejection of defective parts. (c) Use of cameras to provide positional input to a robot relative to the workpiece. (d) Painting parts having different 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.

measurements are digitized (referred to as **image recognition**). There are two basic systems of machine vision.

In a **linear-array** system, only one dimension is sensed, such as the presence of an object or some feature on its surface. By contrast, **matrix arrays** sense two or even three dimensions and are capable of detecting, for example, whether or not a component is inserted properly in a printed circuit or a solder joint has the proper dimensions (known as *assembly verification*). When used in automated inspection systems, these sensors also have the capability to detect cracks and flaws.

Machine vision is particularly useful for inaccessible parts, in hostile manufacturing environments, for measuring a large number of small features, and in situations where physical contact may cause damage to a part. Applications of machine vision include (a) online, real-time inspection in sheet-metal stamping lines and (b) sensors for machine tools that can sense tool offset and tool breakage, verify part placement and fixturing, and monitor surface finish.

Applications. Several applications of machine vision in manufacturing are shown in Fig. 37.26. With visual sensing capabilities, end effectors are able to pick up parts and grip 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 (a) the particular quantity to be measured or sensed, (b) the sensor's interaction with other components in the system, (c) the sensor's expected service life, (d) its required level of sophistication, (e) difficulties associated with its use, (f) the power source, and (g) 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. For example, in machining, a computer chip with sensors can be programmed to turn a machine tool off when a cutting tool fails. Likewise, a smart sensor can stop a mobile robot 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 a number of sensors in such a manner that the 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 we take such a common event for granted, it can readily be seen that drinking from a cup involves data input from our eyes, lips, tongue, and hands. Through our five basic senses (sight, hearing, smell, taste, and touch), there is real-time monitoring of relative movements, positions, and temperatures. For example, if the coffee is too hot, the hand movement of the cup toward the lip is slowed and adjusted accordingly. Note that the fingers and the hand already sense the temperature and that this signal has become an input into the conrol system.

Another example of sensor fusion is a machining operation in which a set of different, but integrated, sensors monitors (a) the dimensions and surface finish of the workpiece as it is being machined, (b) forces, vibrations, chatter, and tool wear and failure, (c) spindle power, and (d) increases in temperature in various regions of the tool–workpiece interface.

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. For this application, receiving redundant data from different sensors is essential. It can be seen that receiving, integrating, and processing all data from various sensors can be a complex problem.

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.

EXAMPLE 37.2 Special Applications of Sensors

Three special applications of sensors are described in this example:

1. Fiber-optic sensors are being developed for gas-turbine engines. With a size of the head of a pin, these sensors are installed in critical locations in the engine, and they monitor conditions such as temperature, pressure, and the flow of gases. Continuous monitoring of the signals from these sensors helps detect possible engine problems and also provides the necessary data for improving the efficiency of the engines.

Source: Courtesy of Prime Research.

2. It has been determined that, in the United States, more than a quarter of the passenger-vehicle tires are underinflated, a condition that can lead to tire separation and blowouts. Tire-pressure remote sensors have been developed and are installed in some automobiles. Because underinflated tires can become hot due to excessive flexing and internal friction, a special rubber has been developed that changes its color from black to red when its temperature rises above 77°C. Thus, a small strip of this material is built into the sidewalls of tires and would be visible as a red ring when a tire gets too hot.

   Source: After F. Kelley and B. Rosenbaum.

3. Electronic sensors have been developed to distinguish basic tastes, such as the sourness of milk and the bitterness of medications. Coffee, however, has subtle taste variations that are difficult to detect by these sensors, including sensory characteristics such as caffeine level, acidity, sourness, and bitterness. Every year, $70 billion worth of coffee is sold around the world, and the quality of coffee beans traditionally has been checked and ensured by human tasters.

   A new "electronic tongue" sensor has been developed, consisting of 10 gold electrodes covered with thin films of an electrically conducting polymer. When this sensor is immersed in liquid coffee, various molecules in the coffee are absorbed by the films, changing their electrical properties.

   Source: After L.H. Mattoso and *Technology Review*.

37.8 Flexible Fixturing

In describing *work-holding devices* for the manufacturing operations discussed throughout this book, the words *clamp*, *jig*, and *fixture* often were used interchangeably and sometimes in pairs, such as in *jigs and fixtures*. **Clamps** are simple multifunctional work-holding devices, and **jigs** have various reference surfaces and points for accurate alignment of parts or tools for processing. **Fixtures** generally are designed for specific purposes. Other work-holding devices are *chucks*, *collets*, and *mandrels*, many of which are usually operated manually. 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 certain ranges of dimensions; and (c) various devices and fixtures called *dedicated fixtures* are designed and made for specific workpiece shapes and dimensions and for specific tasks. If the workpiece has curved surfaces, it is possible to shape the contacting surfaces of the jaws themselves by machining them (these jaws are known as *machinable jaws*) to conform to the workpiece surfaces.

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 behind what are 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 often is used for small or moderate lot sizes, especially when the cost of dedicated fixtures and the time required to produce them may be difficult to justify. Complex workpieces can be located within machines through fixtures produced quickly from standard components and then can 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.

Several other standard components, such as locating pins, adjustable stops, workpiece supports, V-blocks, clamps, and springs, can be mounted onto a *base plate* or block to quickly produce a fixture. By computer-aided fixture-planning techniques, such fixtures can be designed, assembled, and modified with the use of industrial robots. Compared with dedicated fixturing, modular fixturing has been shown to be low in cost, to have a shorter lead time, to have more easily repaired components, and to possess more intrinsic flexibility of application.

Tombstone Fixtures. Also referred to as **pedestal-type fixtures**, tombstone fixtures have between two and six vertical faces (hence, they resemble *tombstones*) onto which workpieces can be mounted. Tombstone 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 tombstone to begin work on other surfaces. These fixtures allow feeding more than one part into a machine, but are not as flexible as other fixturing systems. 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 kind of system is shown in Fig. 37.27. 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 to the workpiece 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 in a medium:

a. 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 (like the wooden stick in a popsicle), a process that is similar to *insert molding* (Sections 11.3.5 and 19.3). After setting, the solidified lead block is clamped in a simple fixture. 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.

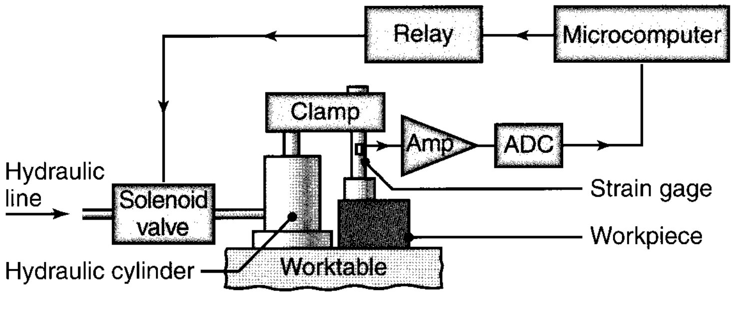


FIGURE 37.27 Schematic illustration of an adjustable-force clamping system. The clamping force is sensed by the strain gage, and the system automatically adjusts this force. *Source*: After P.K. Wright.

b. The supporting medium is a *magnetorheological* (MR) or *electrorheological* (ER) *fluid*. In the MR application, the particles are ferromagnetic or paramagnetic nanoparticles 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. Afterwards, the workpiece is retrieved by removing the external magnetic field. The process is particularly suitable for nonferrous metal parts. In the ER application, the fluid is a suspension of fine dielectric particles in a liquid of low dielectric constant. After the application of an electrical field, the liquid becomes a solid.

37.9 Assembly Systems

The individual parts and components produced by various manufacturing processes have to be *assembled* into finished products. Traditionally, assembly has involved much manual work and thus has contributed significantly to product cost. The total assembly operation is usually broken into individual assembly operations (*subassemblies*), with an operator assigned to carry out each step.

Assembly costs are typically 25 to 50% of the total cost of manufacturing, with the percentage of workers involved in assembly operations ranging from 20 to 60%. In the electronics industries, typically some 40 to 60% of total wages are paid to assembly workers. 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 been improved upon vastly over the years. The first application of large-scale efficient assembly was the assembly of flywheel magnetos for the Model T Ford. This experience eventually led to mass production of the whole 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, we know that although they all have the same *nominal* dimensions, some balls in a lot will be smaller than others 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. 37.28, 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,

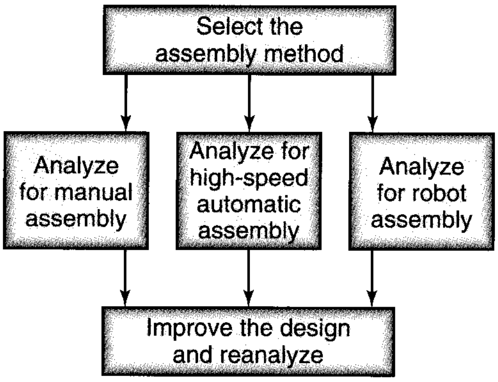


FIGURE 37.28 Stages in the design-for-assembly analysis. *Source*: After G. Boothroyd and P. Dewhurst.

the aligning and placing of a simple square peg into a square hole involving small clearances can be difficult in automated assembly—yet the human hand is capable of doing this simple operation with relative ease.

2. **High-speed automated assembly** utilizes *transfer mechanisms* designed specially for assembly. Two examples of such assembly are shown in Fig. 37.29, in which individual assembly is carried out on products that are *indexed* for proper positioning.
3. In **robotic assembly**, one or two general-purpose robots operate at a single workstation (Fig. 37.30), or the robots operate at a multistation assembly system.

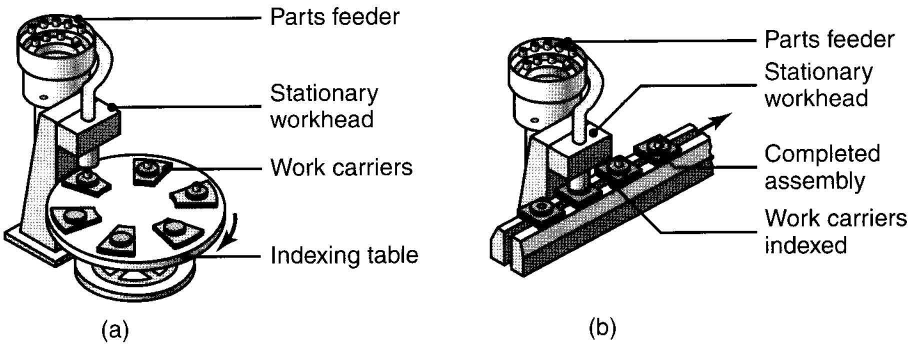


FIGURE 37.29 Transfer systems for automated assembly: (a) rotary indexing machine and (b) in-line indexing machine. *Source*: After G. Boothroyd.

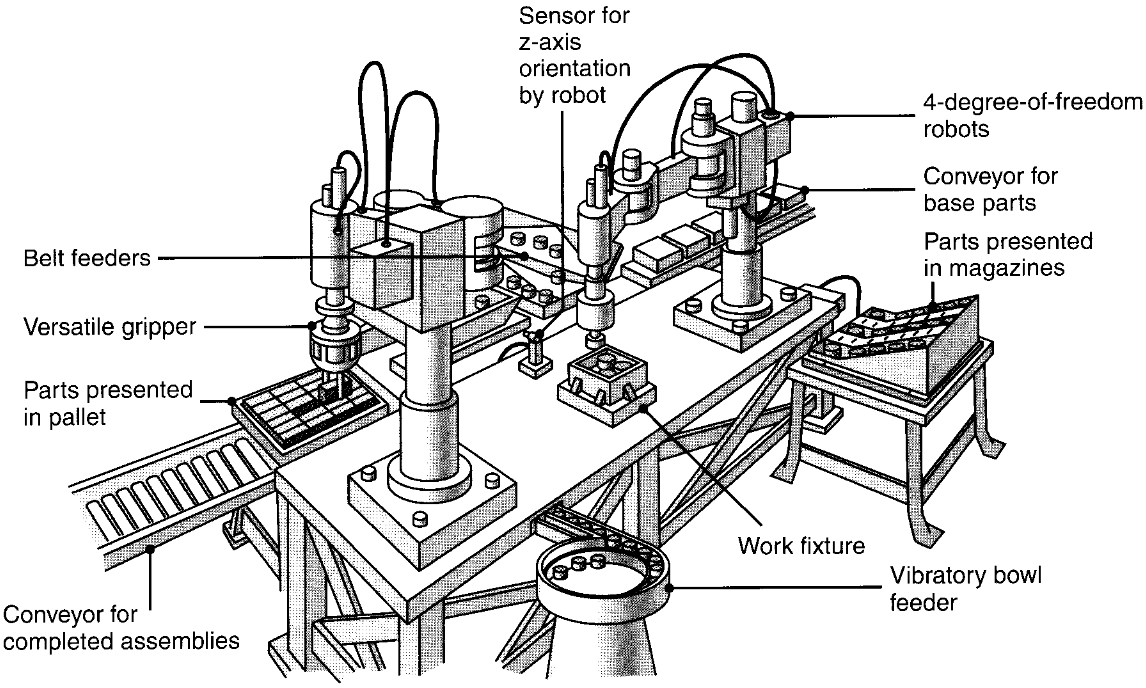


FIGURE 37.30 A two-arm robot assembly station. *Source*: After G. Boothroyd and P. Dewhurst.

There are three basic types of assembly systems:

a. Synchronous systems. In these *indexing* systems, individual parts and components are supplied and assembled at a constant rate at fixed individual stations. The rate of movement of the product in the 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. 37.29. These systems can operate in either a fully automatic mode or a semiautomatic mode. Note, however, that a breakdown of one station will shut down the whole assembly operation.

The *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. 37.31. Orienting parts properly and avoiding jamming are essential in all automated assembly operations.

b. Nonsynchronous systems. Each station in this system operates independently, and any imbalance is accommodated in storage (*buffer*) between stations. The station continues operating until the next buffer is full or the previous buffer is empty. Furthermore, if, for some reason, one station becomes inoperative, the

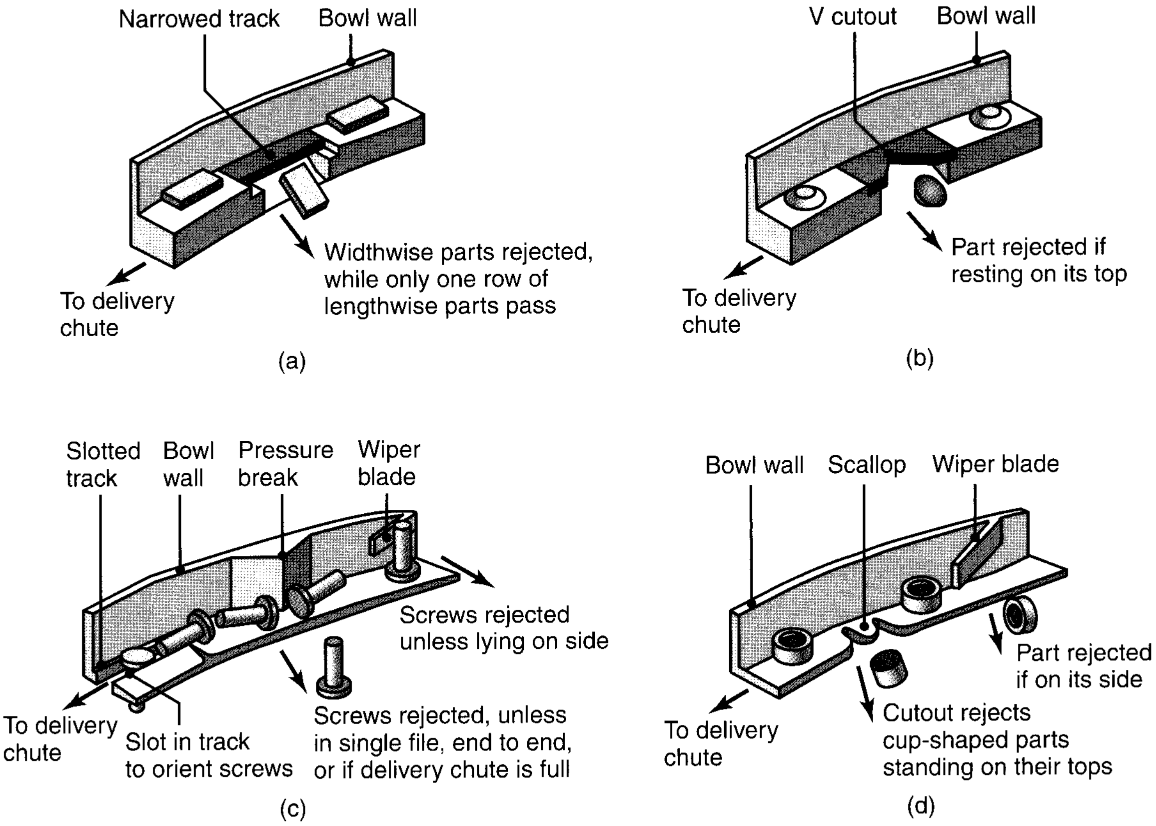


FIGURE 37.31 Examples of guides to ensure that parts are properly oriented for automated assembly. *Source*: After G. Boothroyd.

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. If the times required for the individual assembly operations vary significantly, the output will be constrained by the slowest station.

c. **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.

Flexible Assembly Systems. Assembly systems generally are set up for a specific product line. However, they can be modified for increased flexibility in order to assemble product lines that have a variety of product models. Such *flexible assembly systems* (FAS) utilize 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 manufacturing processes and systems, design of the devices and systems is an integral part of the topics described. These topics require special design considerations, as described here.

37.10.1 Design for Fixturing

The proper design, construction, and operation of flexible work-holding devices and fixtures are essential to the operation and efficiency of advanced manufacturing systems. The major design issues involved are as follows:

- Work-holding devices must position the workpiece automatically and accurately. They must maintain its location precisely and with sufficient clamping force to withstand the particular manufacturing operation. Fixtures also should be able to accommodate parts repeatedly in the same position.
- The fixtures must have sufficient stiffness to resist, without excessive distortion, the normal and shear stresses 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 stick to the wet surfaces due to surface-tension forces.
- A flexible fixture should be able to accommodate 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 thermoplastic or rubber parts, or (c) has a relatively soft coating on its contacting surfaces.
- Clamps and fixtures should have low profiles to avoid collision with cutting tools. The avoidance of collisions is also an important factor in programming tool paths in machining operations.

- Flexible fixturing must meet special requirements in manufacturing cells and flexible manufacturing systems.
- Workpieces should 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 machinery.

37.10.2 Design for Assembly, Disassembly, and Servicing

Design for Assembly. While product design for manufacture has been a matter of major interest for some time, *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, because the human senses guide the hands properly. However, in *high-speed automated assembly*, automatic handling generally requires that parts be separated from the bulk, conveyed by hoppers or vibratory feeders (Fig. 37.31), and assembled in their proper locations and orientations.

General guidelines for design for assembly may be summarized as follows:

1. Reduce the number and variety of parts in a product. Simplify the product design and incorporate multiple functions into a single part. 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. Design parts for easy insertion.
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, and 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. Parts should be inserted from a single direction, preferably vertically (from above) to take advantage of gravity. Assembly from two or more directions can be difficult.
7. Products should be designed, or existing products redesigned, so that there are no physical obstructions to the free movement of the parts during assembly. For example, sharp external and internal corners 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.

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. Significant effort has been directed towards the development of analytical and computer-based tools to estimate the efficiency of assembly operations. These tools provide a basis for comparisons of designs and the objective selection of design attributes that make assembly easier.

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 be evaluated with respect to design-for-assembly. It has been noted that products which are in need of redesign to facilitate assembly usually have assembly efficiencies around 5 to 10%, while well-designed parts have assembly efficiencies around 25%. (Note that assembly efficiencies near 100% are unlikely to be achieved in practice.)

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. Consider, for example, the difficulties one has in removing certain components from under the hood of some automobiles; similar difficulties exist in the disassembly of several 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 objects.

An important aspect of design for disassembly is how, after its life cycle, a product is to be taken apart for *recycling*, especially to salvage its more valuable components. For example, note that, depending on their design and location, the type of tools used, and whether manual or power tools are used, (a) rivets may take longer to remove than screws or snap fits and that (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 has to be studied and measured. Although that time depends on the manner in which disassembly is done, some examples are as follows: cutting wire at 0.25 second, disconnecting wire at 1.5 seconds, effecting snap fits and clips at 1 to 3 seconds, and loosening screws and bolts at 0.15 to 0.6 second per revolution.

Design for Servicing. Design for servicing is essentially based on the concept that the elements which 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 operations. Because all production systems are essentially a combination of machines and people, important factors influencing the final decisions include (a) the type and cost of machinery, equipment, and tooling; (b) the cost of operation of the machinery; (c) the skill level and amount of labor required; and (d) the production quantity desired. Recall also that lot size and production rate greatly influence the economics of production.

Small quantities per year can be manufactured in job shops. However, the type of machinery in job shops generally requires skilled labor to operate it, and the production quantity and rate are low; as a result, the cost per part can be high. At the other extreme is the production of very large quantities of parts, 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 labor skill required and the labor costs are relatively low because of the high level of automation implemented. However, these production systems are organized for a specific type of product and, hence, 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. Generally, the higher the level of skill available in the workforce, the lower is the need for automation, provided that higher labor costs are justified and that there are sufficient qualified workers available. Conversely, if a manufacturing facility already has been automated, the skill level required is relatively lower.

Furthermore, some products—especially those requiring extensive assembly—have a large labor component; thus, their production is labor intensive. Examples of labor-intensive products are aircraft, software, bicycles, pianos, furniture, toys, shoes, textiles, and garments. Labor intensiveness is a major reason that so many consumer products are now made or assembled in countries where labor costs are low (Table I.4).

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 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 these modern technologies; a wide variety of sensors based on various principles have been developed and installed.
- Other advances include flexible fixturing and automated assembly techniques that reduce the need for worker intervention and that lower manufacturing costs. The effective and economic implementation of these techniques requires that design for assembly, disassembly, and servicing be recognized as an important factor in the total design and manufacturing process.
- The efficient and economic implementation of these techniques requires that design for assembly, disassembly, and servicing be recognized as an important factor in the total design and manufacturing process.

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
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

Amic, P.J., **Computer Numerical Control Programming**, Prentice Hall, 1996.

Asfahl, C.R., **Robots and Manufacturing Automation**, 2nd ed., Wiley, 1992.

Boothroyd, G., **Assembly Automation and Product Design**, 2nd ed., Marcel Dekker, 1992.

Boothroyd, G., Dewhurst, P., and Knight, W., **Product Design for Manufacture and Assembly**, 2nd ed., Marcel Dekker, 2001.

Busch-Vishniac, I., **Electromechanical Sensors and Actuators**, Springer, 1999.

Craig, J.J., **Introduction to Robotics: Mechanics and Control**, 3rd ed., Prentice Hall, 2004.

Fraden, J., **Handbook of Modern Sensors: Physics, Designs, and Applications**, 3rd ed., Springer-Verlag, 2004.

Galbiati, L.J., **Machine Vision and Digital Image Processing Fundamentals**, Prentice Hall, 1997.

Gibbs, D., and Crandell, T., **CNC: An Introduction to Machining and Part Programming**, Industrial Press, 1991.

Hornberg, A., **Handbook of Machine Vision**, Wiley, 2007.

Ioannu, P.A., **Robust Adaptive Control**, Prentice Hall, 1995.

Jain, R. (ed.), **Machine Vision**, Prentice Hall, 1995.
Lynch, M., **Computer Numerical Control for Machining**, McGraw-Hill, 1992.
Mitchell, H.B., **Multi-Sensor Data Fusion: An Introduction**, Springer, 2008.
Molloy, O., Warman, E.A., and Tilley, S., **Design for Manufacturing and Assembly: Concepts, Architectures and Implementation**, Kluwer, 1998.
Nof, S.Y. (ed.), **Handbook of Industrial Robotics**, Wiley, 1999.
Rehg, J.A., **Introduction to Robotics in CIM Systems**, 5th ed., Prentice Hall, 2002.
Steger, E.R., Ulrich, M., and Wiedemann, C., **Machine Vision Algorithms and Applications**, Wiley, 2008.
Stenerson, J., and Curran, K.S., **Computer Numerical Control: Operation and Programming**, 3rd ed., Prentice Hall, 2006.
Valentino, J.V., and Goldenberg, J., **Introduction to Computer Numerical Control**, 4th ed., Regents/Prentice Hall, 2008.
Zhang, J., **Practical Adaptive Control: Theory and Applications**, VDM Verlag, 2008.
Zuech, N., **Understanding and Applying Machine Vision**, 2nd ed., Marcel Dekker, 1999.

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. Describe the principle of numerical control of machines.

37.4. Explain open-loop and closed-loop control circuits.

37.5. Describe the principle and purposes of adaptive control.

37.6. What factors have led to the development of automated guided vehicles?

37.7. Describe the features of an industrial robot. Why are these features necessary?

37.8. List and describe the principles of various types of sensors.

37.9. Describe the concept of design for assembly. Why has it become an important factor in manufacturing?

37.10. Is it possible to have partial automation in assembly? Explain.

37.11. Explain the advantages of flexible fixturing.

37.12. How are robots programmed to follow a certain path?

QUALITATIVE PROBLEMS

37.13. Why is automation generally regarded as evolutionary rather than revolutionary?

37.14. Are there activities in manufacturing operations that cannot be automated? Explain.

37.15. Explain the factors that have led to the development of numerical control.

37.16. Giving specific examples, discuss your observations concerning Fig. 37.2.

37.17. What are the relative advantages and limitations of the two arrangements for power heads shown in Fig. 37.4?

37.18. Discuss methods of online gaging of workpiece diameters in turning operations other than that shown in Fig. 37.15.

37.19. Are drilling and punching the only applications for the point-to-point system shown in Fig. 37.10a? Explain.

37.20. What determines the number of robots in an automated assembly line such as that shown in Figs. 37.22 and 37.30?

37.21. Describe situations in which the shape and size of the work envelope of a robot (Fig. 37.20) can be critical.

37.22. Explain the difference between an automated guided vehicle and a self-guided vehicle.

37.23. Explain why sensors have become so essential in the development of automated manufacturing systems.

37.24. Table 37.2 shows a few examples of typical products for each category. Add several other examples to the table.

37.25. Describe applications of machine vision for specific parts that are similar to the examples shown in Fig. 37.26.

37.26. Sketch the workspace (envelope) of each of the robots shown in Fig. 37.19.

QUANTITATIVE PROBLEMS

37.27. 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 16.53, and estimate the assembly efficiency for each design.

37.28. 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.

37.29. 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.

▶**37.30.** Review Example 14.1, and develop open- and closed-loop control system equations for the force if the coefficient of friction is μ.

SYNTHESIS, DESIGN, AND PROJECTS

37.31. Refer to Part III of this book, and give an example of a metal-forming operation that is suitable for adaptive control.

37.32. Describe possible applications for industrial robots not discussed in this chapter.

37.33. Design two different systems of mechanical grippers for two widely different applications.

37.34. Give some applications for the systems shown in Figs. 37.26a and c.

37.35. For a system similar to that shown in Fig. 37.27, design a flexible fixturing setup for a lathe chuck.

37.36. Give examples of products that are suitable for the three types of production shown in Fig. 37.3.

▶**37.37.** Describe situations in which tactile sensors would not be suitable. Explain why.

▶**37.38.** Are there situations in which machine vision cannot be applied properly and reliably? Explain.

37.39. Choose one machine each from Parts II through IV, and design a system in which sensor fusion can be used effectively.

37.40. 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.

37.41. 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.

37.42. 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.

37.43. 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.

37.44. 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.

37.45. 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.

37.46. Obtain an old toaster and disassemble it. Explain how you would go about reassembling it by automated assembly.

37.47. 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.

Computer-Aided Manufacturing

CHAPTER 38

- Computers have fundamentally and pervasively changed the product design and manufacturing process; powerful computer software is now available to assist and integrate all engineering tasks.
- This chapter opens with a discussion of computer-aided design, in which the graphic description of parts is performed through 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.

38.1 Introduction

The importance of product quality was emphasized in Chapter 36, along with the necessity for the commitment of a company's management to total quality control. 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 integrated properly, rather than if they were treated as separate entities. 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 can be described in a CAD program, which is in itself a fairly complex computer program. The manufacture of this part can be achieved, for example, by programming it into G-code (see Section 37.3.6), which uses another fairly complex computer program to translate instructions into machine actions. However, 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 possible.

38.2 Manufacturing Systems

As we have seen throughout various chapters, 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 adjustments can be made.

For example, the supply of a particular raw material may be reduced significantly due to, say, global demands—especially by developing countries—strife, strikes, or geopolitical reasons. Because the raw material cost now will rise, alternative materials have to be considered and selected. This 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. For instance, the material selected may not be as easy to form, machine, or weld, and thus product integrity may suffer during its processing.

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 on a relatively short lead time while minimizing large expenditures in new machinery and tooling.

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. Furthermore, it is not easy to correctly predict and control some of these variables, because (a) machine-tool characteristics, their performance, and their response to random external disturbances cannot be precisely modeled, (b) raw-material costs are difficult to predict accurately, and (c) human behavior and performance are even more difficult to model. In spite of the difficulties, much progress continues to be made in modeling and simulating manufacturing systems.

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. CIM is a *methodology*, and its effectiveness critically depends on the use of a *large-scale integrated communications system* consisting of computers, machines, and their controls, all described in 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-sized companies.

Comprehensive and long-range strategic planning covering all phases of the operation is essential in order to fully benefit from CIM. 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 of CIM. Computer-integrated manufacturing systems comprise the following *subsystems*, which are integrated into a whole (Fig. 38.1):

- **a.** Business planning and support,
- **b.** Product design,
- **c.** Manufacturing process planning,
- **d.** Process automation and control,
- **e.** Production-monitoring systems.

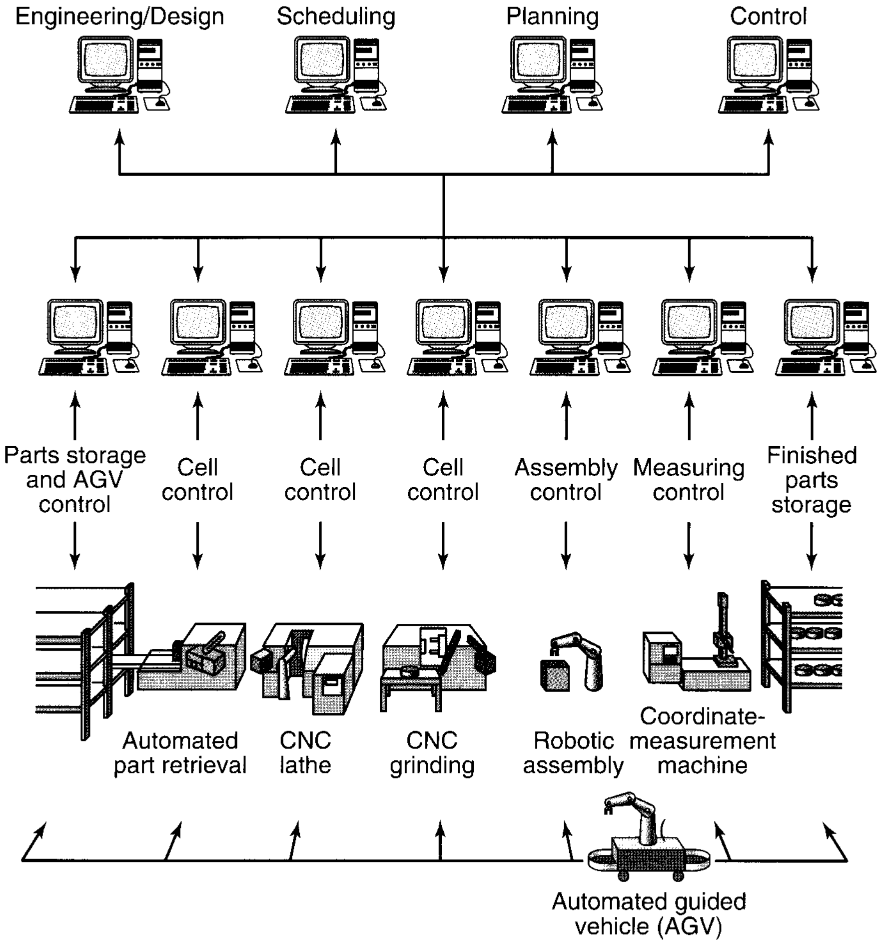


FIGURE 38.1 A schematic illustration of a computer-integrated manufacturing system. The manufacturing cells and their controls shown at the lower left are described in Section 39.2. *Source*: After U. Rembold.

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**: These include activities such as forecasting, scheduling, material-requirements planning, invoicing, and accounting.
- **Business-execution functions**: Includes production and process control, material handling, testing, and inspection of the system.

If CIM is implemented properly, its major benefits are as follows:

- 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 Database

An efficient computer-integrated manufacturing system requires a single *database* to be shared by the entire manufacturing 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 is stored in computer memory and recalled or modified as necessary, either by individuals in the organization or by the CIM system itself.

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**: Revision level and part number—for example, to establish the history of a part; this history can include descriptions or keywords to assist in retrieving data.
- **Production data**: Manufacturing processes used.
- **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, using 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 are also 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, 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, of course, 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 computers to create design drawings and product models. (See also Fig. I.11 in the General Introduction.) Computer-aided design is usually associated with **interactive computer graphics**, known as a **CAD system.** These systems are powerful tools and are used in the design and geometric modeling of components and products. The designer can easily conceptualize the part to be designed on a computer monitor and can consider alternative designs or modify a particular design to quickly respond to specific design requirements.

There are several powerful commercially available programs to aid designers in geometry description and engineering analysis, such as SolidWorks, ProEngineer, CATIA, AutoCAD, Solid Edge, and VectorWorks. The software can help identify potential problems, such as excessive loads, deflections, or interference at mating surfaces, when these problems are 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 designer can analyze the manufacturing economics of alternative designs.

Computer-aided engineering (CAE) allows 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 circuits and various electronic devices.

38.4.1 Exchange Specifications

Because of the availability of a wide variety of CAD systems with different characteristics 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 basically it has become a standard because of its long-term success. DFX is, however, limited to transferring geometric information only. Similarly, stereolithography (STL; Section 20.3.2) 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 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.

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) 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 increasingly popularity of PDES and STEP have led to less use of IGES.

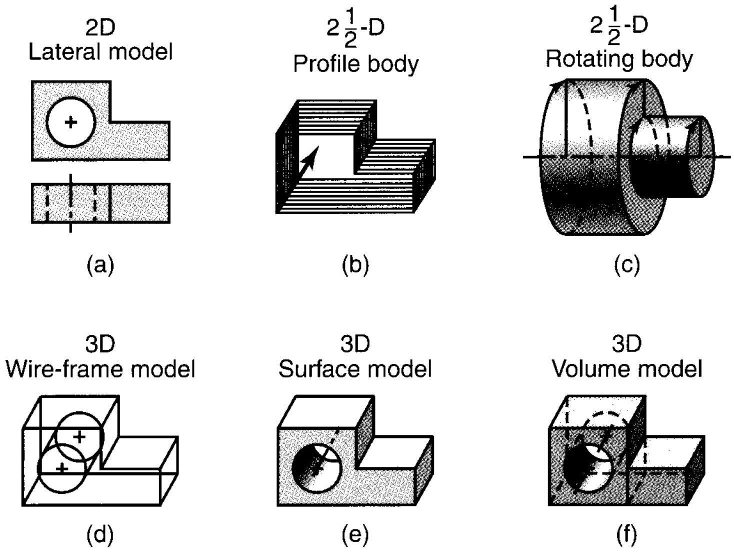


FIGURE 38.2 Various types of modeling for CAD.

38.4.2 Elements of CAD Systems

The design process in a CAD system consists of four stages.

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. A two-and-one-half-dimensional image can be obtained by a *translational sweep*—that is, 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.

2. In the **surface model**, all visible surfaces are shown. Surface 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 the 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 kind 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 describe the interior volume. Solid models can be constructed from *swept volumes* (Figs. 38.2b and c) or by the techniques shown in Fig. 38.4. In **boundary representation** (B-rep), surfaces are combined to develop a solid model (Fig. 38.4a). In **constructive solid geometry** (CSG), simple shapes such as spheres, cubes, blocks, cylinders, and cones (called **primitives of solids**) are combined to develop a solid model (Fig. 38.4b).

   Computer programs allow the user to select any combination of these primitives and their sizes and combine them into the desired solid model. Although solid models have such advantages as ease of design analysis and ease of preparation for manufacturing the part, they require more computer memory and processing time than the wire-frame and surface models shown in Fig. 38.2.

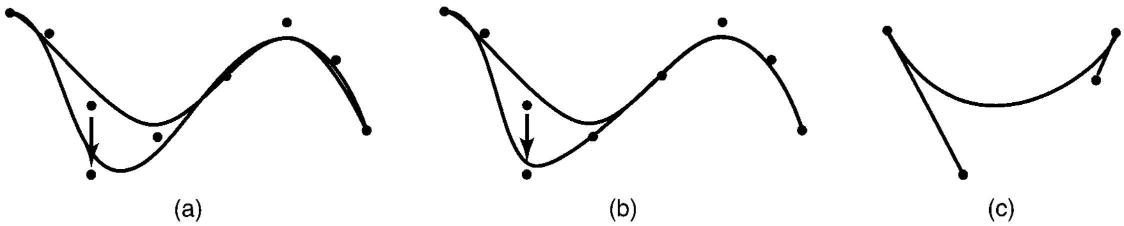


FIGURE 38.3 Types of splines. (a) A Bezier curve passes through the first and last control point, but generates a curve from the other points. Changing a control point modifies the entire curve. (b) A B-spline is constructed piecewise, so that changing a vertex affects the curve only in the vicinity 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.

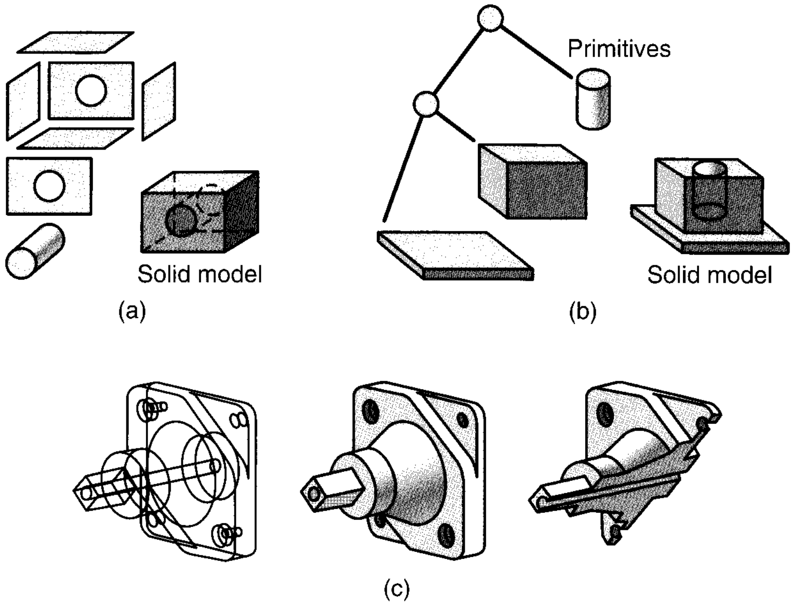


FIGURE 38.4 (a) Boundary representation of solids, showing the enclosing surfaces of the solid model and the generated solid model. (b) A solid model represented as compositions of solid primitives. (c) Three representations of the same part by CAD. *Source*: After P. Ranky.

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. 38.5). Whenever a change is made, the part is re-created from these definitions. This feature allows for simple and straightforward updates and changes to be made to the models.

The **octree representation** of a solid object is shown in Fig. 38.6; it is a three-dimensional analog to pixels on a television screen. Just as any area can be broken down into quadrants, any volume can be broken down into octants, which are then identified as solid, void, or partially filled. Partially filled *voxels* (from *vo*lume pi*xels*) are broken

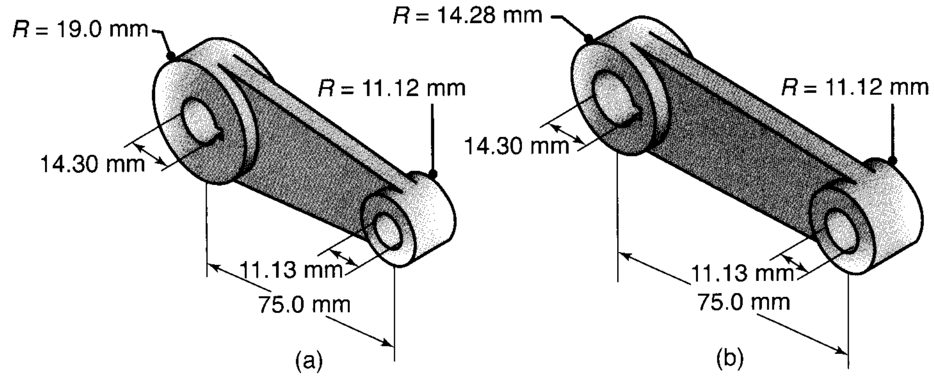


FIGURE 38.5 An example of parametric design. Dimensions of part features can be modified easily to quickly obtain an updated solid model.

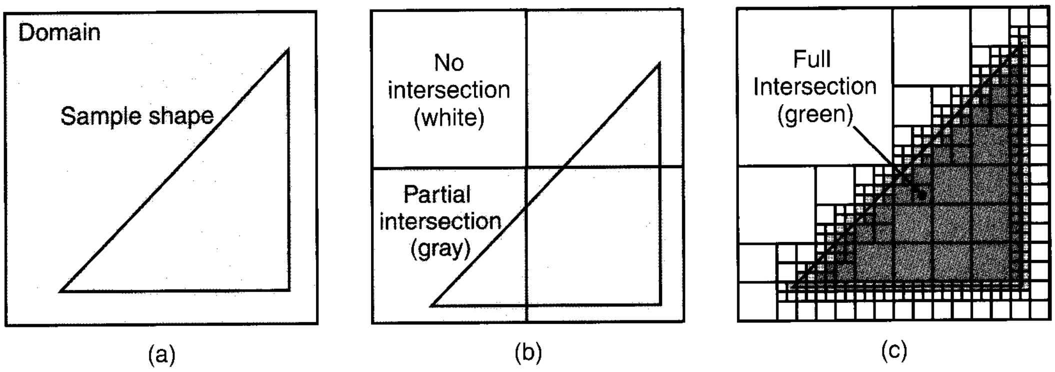


FIGURE 38.6 The octree representation of a solid object. Any volume can be broken down into octants, which are then identified as solid, void, or partially filled. Shown is a two-dimensional version (or quadtree) for the representation of shapes in a plane.

into smaller octants and are reclassified. With increasing resolution, exceptional part detail can be achieved. This process may appear to be somewhat cumbersome, but it allows for accurate description of complex surfaces. It is used particularly in biomedical applications, such as modeling bone geometries.

A **skeleton** (Fig. 38.7) is commonly used for kinematic analysis of parts or assemblies. A skeleton is the family of lines, planes, and curves that describe a part, but without the detail of surface models. Conceptually, a skeleton can be constructed by fitting the largest circle (or sphere for three-dimensional objects) within the geometry. The skeleton is the set of points that connect the centers of the circles or spheres. A continuing area of research involves using skeleton models instead of conventional surface or solid models.

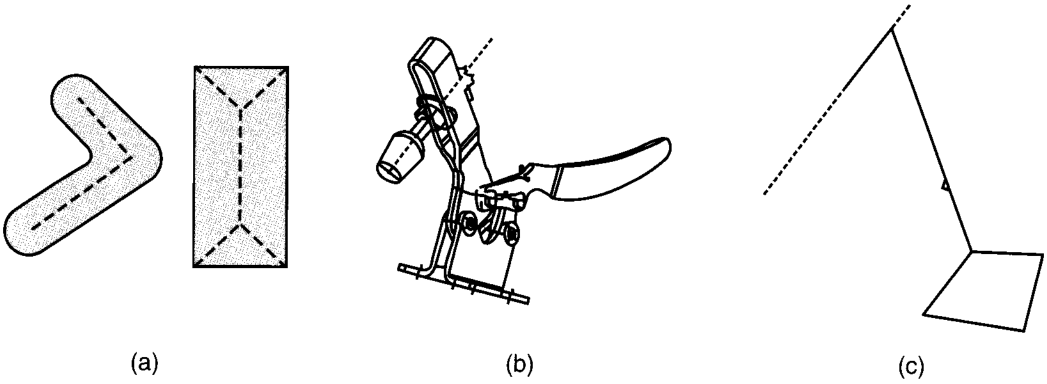


FIGURE 38.7 (a) Illustration of the skeleton data structure for solid objects. The skeleton is the dashed line in the object interior. (b) General view of a clamp. (c) Skeleton model. *Source*: S.D. Lockhart and C.M. Johnson, *Engineering Design Communication*, Prentice Hall, 2000.

CASE STUDY 38.1 CAD Model Development for Automotive Components

In the automotive industry, it is especially important to have a detailed CAD model of a particular component in the product database in order to ensure that all those who will be working on it have all the data they need to perform their tasks. Special care is taken to build very precise CAD models of the automobile components that passengers will see and interact with on a regular basis. Examples of such components are outer-body panels, handles, seats, and the instrument panel (Fig. 38.8). The quality of the *visible* (Class I) surfaces has a major impact on overall vehicle quality and customer's perception of the look and feel of the automobile.

Two-dimensional Concept Sketches

Stylists with a background and experience in industrial design or art first develop two-dimensional concepts through a series of sketches. These sketches most frequently are drawn by hand, although software may

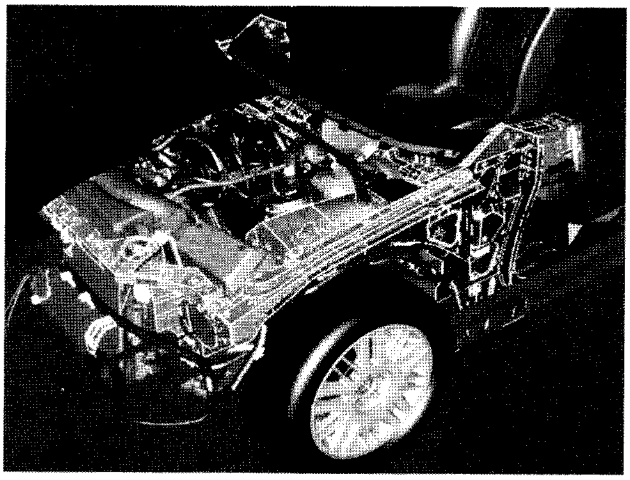


FIGURE 38.8 Every vehicle component, from body panels to knobs on the instrument panel, has a solid model associated with it. *Source*: Courtesy of Ford Motor Company.

be used instead, especially if the stylist starts with a photograph or a scanned drawing that needs to be modified. Concept sketches provide an overall feel for aesthetics, frequently are very detailed, and show texture, color, and the locations where individual surfaces on a vehicle should meet.

Most often, stylists are given a set of packaging constraints, such as (a) how a component should be assembled together with other components, (b) what the size of the component should be, and (c) what the size and shape of any structure lying behind the visible surfaces should be. The time involved in producing a series of such concept sketches for an individual component or a set of components typically ranges from a few days to several weeks.

Three-dimensional Surface Model

As a concept is being reviewed and refined, several highly accurate surface models of the component are constructed. To start the surface model, a computer-controlled optical scanner scans a conceptualized clay model, producing a *cloud of points* organized along the scan lines. Depending on the size of the component, a point cloud may consist of hundreds of thousands to millions of points. The point cloud is read into point-processing software (such as Paraform or ICEM Surf)[1] to further organize the points and filter out noise. Scanning can take anywhere from several hours to a day to be completed. If, however, a digital three-dimensional clay model of the component is already available, it is converted into a point cloud organized into scan lines without the need for physical scanning and point-cloud postprocessing.

Next, the scan lines from the point cloud are used to construct mathematical surfaces with software such as ICEM Surf and Alias/Wavefront StudioTools.[2] To construct the surfaces, free-form NURBS curves that interpolate or approximate the scan lines are first constructed. Next, a NURBS surface patch is fit through the curves. An individual surface patch then models a small region of a single component's face. Several patches are constructed and joined smoothly at common edges to form the entire face.

Faces join each other at common edges to model an entire component. A great deal of experience is required to determine how to divide a face into a collection of patches that can be fit with the simplest low-order surfaces possible and still meet smoothly at common boundaries. A surfacing specialist performs this task in cooperation with the stylist, to ensure that the surfaces are of high quality and that they capture the stylist's intent. A single component may take as long as a week to model.

Surface models are passed along to the various downstream departments to be used for tooling design, feasibility checks, analysis, and the design of *nonvisible* (Class II) surfaces. As the design evolves, the dimensional tolerances on the surfaces are gradually tightened.

In designing outer-body panels, a major milestone is finalizing what is called *first flange and fillet*, shown in Fig. 38.9, in which the edges of the body panels are turned under or hemmed (see Fig. 16.23c) to provide a flange to connect to the inner-body panel. (See Fig. 16.25.) The shape of the flange and fillet affects the overall aesthetics and shape of the body panels; hence, the flange and fillet must be carefully designed. After the shape has been decided, the inner-body panel designs can be completed.

Surface Verification

Once the final surface model is completed, it must be verified and evaluated for surface quality and aesthetics. NC tool paths (Section 37.3.3) are generated automatically from the surfaces, either from within the surfacing software or through specialized machining software. The tool paths then are used to CNC machine the surfaces in clay. If the component is small, then, instead of machining a clay model, an STL file can be generated and the mock-up can be built on a rapid prototyping machine (see Chapter 20). It takes anywhere from several hours to several days to machine a component and perform any hand finishing that may be required.

Clay models can be coated with a thin layer of latex material and painted to make them look more realistic and to help evaluate surface quality. The coating may be modified to improve surface smoothness between patches or to change the way light reflects off the surface of the model. All changes made to the clay model must be translated back to the digital surface model, either through rescanning and refitting the surface patches or by tweaking the shape of the existing patches. At this stage of surface verification, only very minor changes to within ±0.5 mm can be made.

---

[1] Paraform is a product of Paraform, Inc.; ICEM Surf is a product of ICEM Technologies, Inc.

[2] StudioTools is a product of Alias/Wavefront.

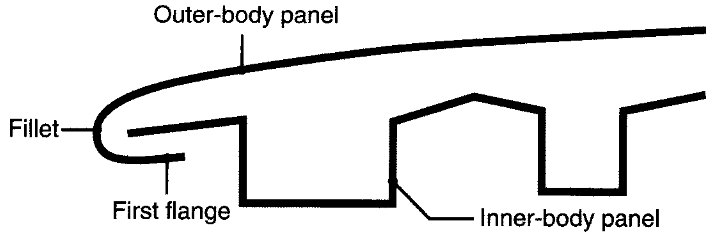


FIGURE 38.9 First flange and fillet (exaggerated).

Solid-model Construction

After the surface model is finalized, it is used to develop a solid model. For sheet-metal components (such as body panels), body specialists offset the surfaces to form a solid. For other components (such as instrument panels, door handles, and wheels), the specialists add manufacturing features to the surface model (such as flanges, bosses, and ribs). Collaborating with manufacturing engineers, they determine what features must be added and where they should be placed to ensure that the component can be fabricated from the desired materials and at the target cost.

In the process of making a solid model, it may be discovered that the surface model has to be modified because of (a) changes in packaging constraints, (b) a component failing to meet minimal manufacturing requirements, or (c) surfaces not matching properly at common edges with the appropriate smoothness. These changes are communicated back to the surface modeler and the stylist so that the surfaces can be modified and reverified. Finally, the solid model is entered into a product database, where it is available to suppliers and engineers for further analysis and manufacturing.

Source: Courtesy of A. Marsan and P. Stewart, Ford Motor Company.

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 consist of analyzing, for example, stresses, strains, deflections, vibrations, heat transfer, temperature distribution, or dimensional tolerances. Various sophisticated 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 done, 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 used to check for any interference or excess gap between various components. This 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 the design review and evaluation 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 that are mass produced according to a given design specification or are identical to 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 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, computer-aided design and computer-aided manufacturing 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 need to reenter the data on part geometry manually. The database developed during CAD is stored and further processed by CAM into the necessary data and instructions for purposes such 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:

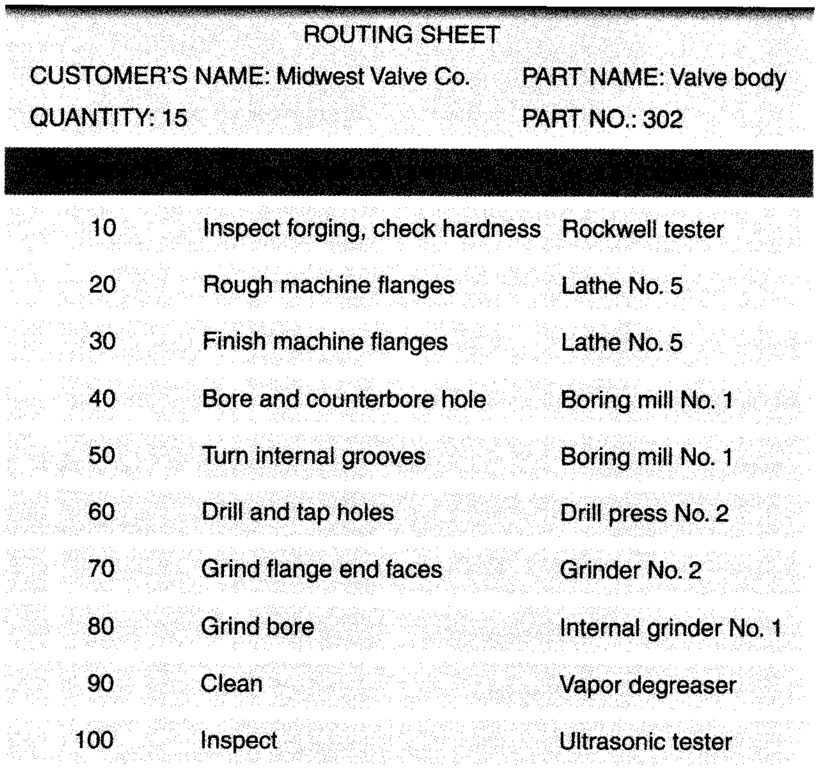
- Programming for numerical control and industrial robots.
- Design of dies and molds for casting in which, for example, shrinkage allowances are preprogrammed.
- Dies for metalworking operations, such as complex dies for sheet forming and progressive dies for stamping.
- Design of tooling and fixtures 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. For example, the two-engine Boeing 777 passenger airplane 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 (which was 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.



ROUTING SHEET

CUSTOMER'S NAME: Midwest Valve Co. PART NAME: Valve body
QUANTITY: 15 PART NO.: 302

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 38.10 An example of a simple routing sheet. These *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. The current practice is to store all relevant data in computers and to affix to the part a bar code that serves as a key to the database of parts information.

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 on a **routing sheet** (Fig. 38.10).

When performed manually, this task is highly labor intensive and time consuming; equally important, it relies heavily on the experience of the process planner. A modern routing sheet now has both the relevant data stored in computer memory and a **bar code** or other identification affixed to the part.

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 it requires extensive software and good coordination with CAD/CAM (as well as with other aspects of integrated manufacturing systems described throughout the rest of this chapter), 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 manufactured. On the basis of its shape and its manufacturing characteristics, a search for a standard plan is 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.

The variant-process plan includes information such as the types of tools and machines required, the sequence of manufacturing 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. In the variant system, 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, one is made for the new part 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 the part shape and dimensions; process capabilities; selection of manufacturing methods, machinery, and tools; and the sequence of operations to be performed. These capabilities of computers (known as **expert systems**) are described in Section 39.8.

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 to 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, or 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.

Effective implementation of ERP can be a challenging task because of the following factors:

- 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.

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:

a. It is a model of a specific operation, intended to determine the viability of a process or to optimize and improve its performance.

b. 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 with various mathematical schemes. (See, e.g., Fig. 10.17). 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 (**process simulation**) that are available commercially and are inexpensive.

The simulation of an entire manufacturing system involving multiple processes and equipment helps plant engineers to organize machinery and to identify critical machinery elements. In addition, such models can assist manufacturing engineers with scheduling and routing, by *discrete-event simulation*. Commercially available software packages often are used for these simulations, altough dedicated software programs written for a specific company also can be developed.

EXAMPLE 38.1 Simulation of Plant-scale Manufacturing

Several examples and case studies presented in this book have focused upon the simulation of individual manufacturing processes. The availability of low-cost, high-performance computer systems and the development of advanced software have allowed the simulation of entire manufacturing systems and have led to the optimization of manufacturing and assembly operations.

As an example, *Digital Manufacturing Hub* software (Delmia Corporation) allows the simulation of manufacturing processes in three dimensions, including (a) the use of human manikins, to identify

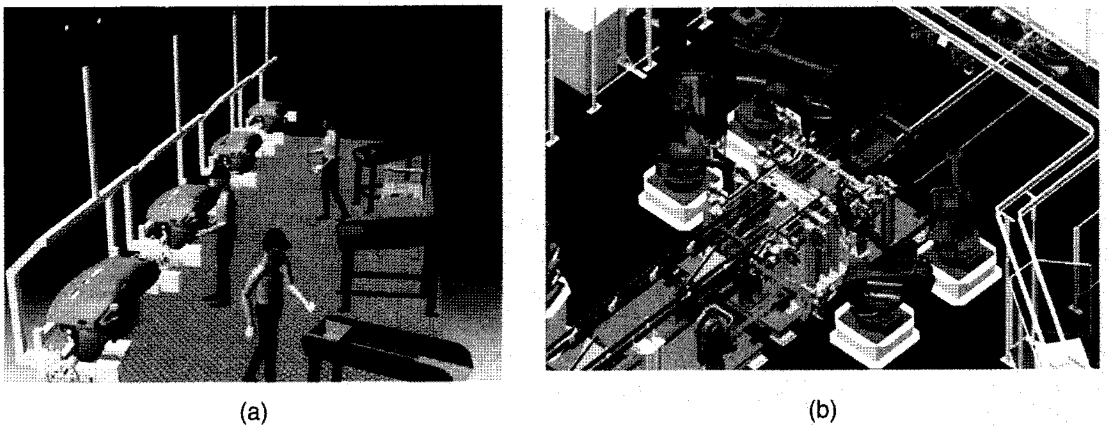

(a) (b)

FIGURE 38.11 Simulation of plant-scale manufacturing operations. (a) The use of virtual manikins to evaluate the required motions and efficiency in manually assembling an automotive dashboard. (b) A robot welding line, where interactions of multiple robots and workpieces can be simulated to detect and avoid collisions and improve productivity. *Source*: Courtesy of Delmia Corporation.

safety hazards, manufacturing problems, or bottlenecks; (b) improving machining accuracy; or (c) optimizing tooling organization. (See Fig. 38.11a.) Since simulation can be performed prior to building an assembly line, it can significantly reduce development times and cost.

For example, Fig. 38.11b illustrates a simulation of a robotic welding line in an automotive plant, where the motions of a robot can be simulated and collisions between neighboring robots or other machinery detected in a virtual environment. The program can then be modified to prevent such collisions before the line is actually put into operation. While this example is a powerful demonstration of the utility of system simulation, a more common application is to optimize the sequence of operations and organization of machinery to reduce manufacturing costs.

The software also has the capability of conducting ergonomic analysis of various operations and machinery setups and, therefore, of identifying bottlenecks in the movement of parts, equipment, or personnel. The bottlenecks can then be relieved by the process planner by adjusting the automated or manual procedures at these locations. Using such techniques, a Daimler-Chrysler facility in Rastatt, Germany, was able to balance its production lines so that each worker is productive an average of 85% to 95% of the time.

Another application of systems simulation is the planning of manufacturing operations to optimize production and to prepare for *just-in-time production*, described in Section 39.5. For example, if an automobile manufacturer needs to produce 1000 vehicles in a given time, production can be optimized by using certain strategies, such as distributing the number of vehicles with sunroofs throughout the day or grouping vehicles by color so that the number of paint changes in paint booths is minimized. With respect to just-in-time production (Section 39.5), software such as that produced by ILOG Corporation can plan and schedule plant operations far enough in advance to allow ordering materials as they are needed, thus eliminating costly stockpiled inventory.

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. These characteristics of similar parts, as illustrated in Fig. 38.12, clearly suggest that 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.

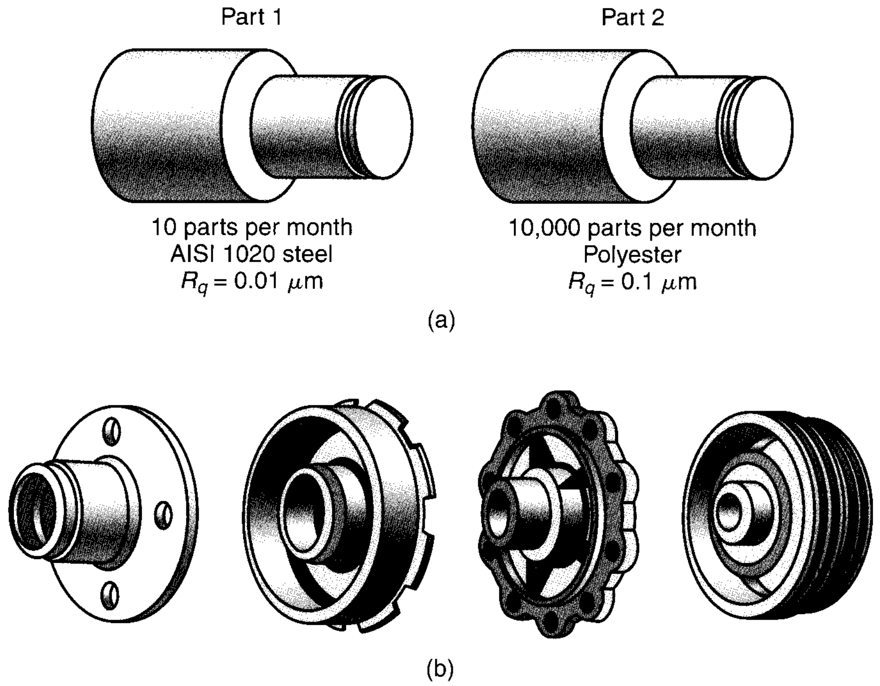


FIGURE 38.12 Grouping parts according to their (a) geometric similarities and (b) manufacturing attributes.

Moreover, questions also must be raised as to why a particular product should have so many different sizes of fasteners.

A pump, for example, can be broken down into its basic components, such as the motor, housing, shaft, flanges, and seals. Note that, in spite of the variety of pumps a company manufactures, each of these components is basically the same in terms of its design and manufacturing characteristics. Consequently, all shafts can be placed in one family of shafts, and so on. Group technology becomes especially attractive because of the ever-greater variety of products, which are often produced in batches, that are available to consumers. Since nearly 75% of manufacturing today is batch production, improving the efficiency of batch production thus becomes important.

Plant Layout. A traditional product flow in batch manufacturing, called **functional layout**, is shown in Fig. 38.13a. Note that machines of the same type are arranged in groups—that is, groups of lathes, milling machines, drill presses, and grinders. Note also 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 not efficient. A more efficient product flow line to take advantage of group technology is the **group layout** (Fig. 38.13b). (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 summarized as follows:

- Group technology makes possible the standardization of part designs and the minimization of design duplication. New part designs can be developed from

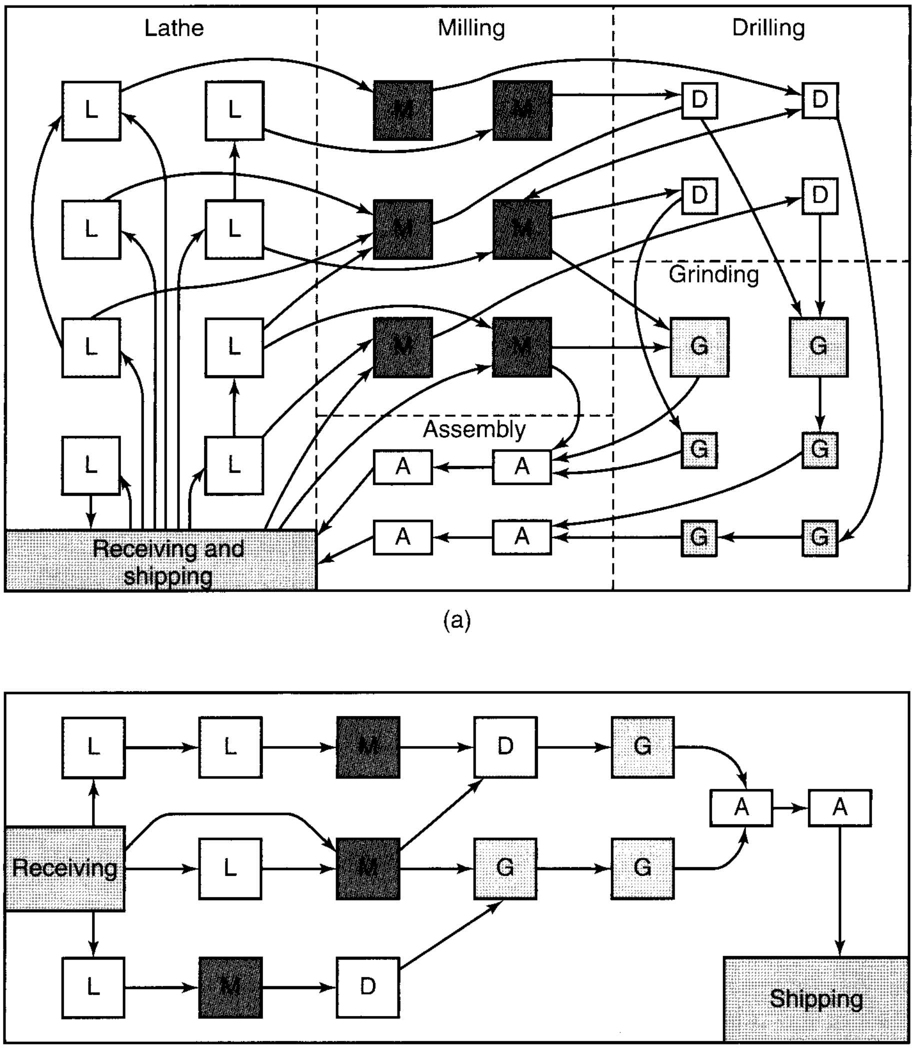


FIGURE 38.13 (a) Functional layout of machine tools in a traditional plant. Arrows indicate the flow of materials and parts in various stages of completion. (b) Group-technology (cellular) layout. *Legend:* L = lathe, M = milling machine, D = drilling machine, G = grinding machine, A = assembly. *Source:* After M.P. Groover.

similar, yet previously used, designs; in this way, a significant amount of time and effort can be saved. The product designer quickly can determine whether data on a similar part already exists in the computer files.

- 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 are standardized and scheduled more efficiently, orders are grouped for more efficient production, and 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. Programming for NC is automated more fully.

- With the implementation of CAD/CAM, cellular manufacturing, and CIM, group technology is capable of greatly improving productivity and reducing costs in 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. 38.12.)

Design Attributes. These attributes pertain to similarities in geometric features and consist of the following:

- 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, including the shape, the dimensions, and other geometric features of the part, manufacturing and design attributes are interrelated. The manufacturing attributes of a part consist of the following:

- 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. The coding can be done simply by viewing the shapes of the parts in a generic way and then classifying the parts accordingly, into, 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:

a. **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.

b. **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.

c. **Decision-tree coding.** This type of coding, also called **hybrid coding**, is the most advanced and combines both design and manufacturing attributes (Fig. 38.14).

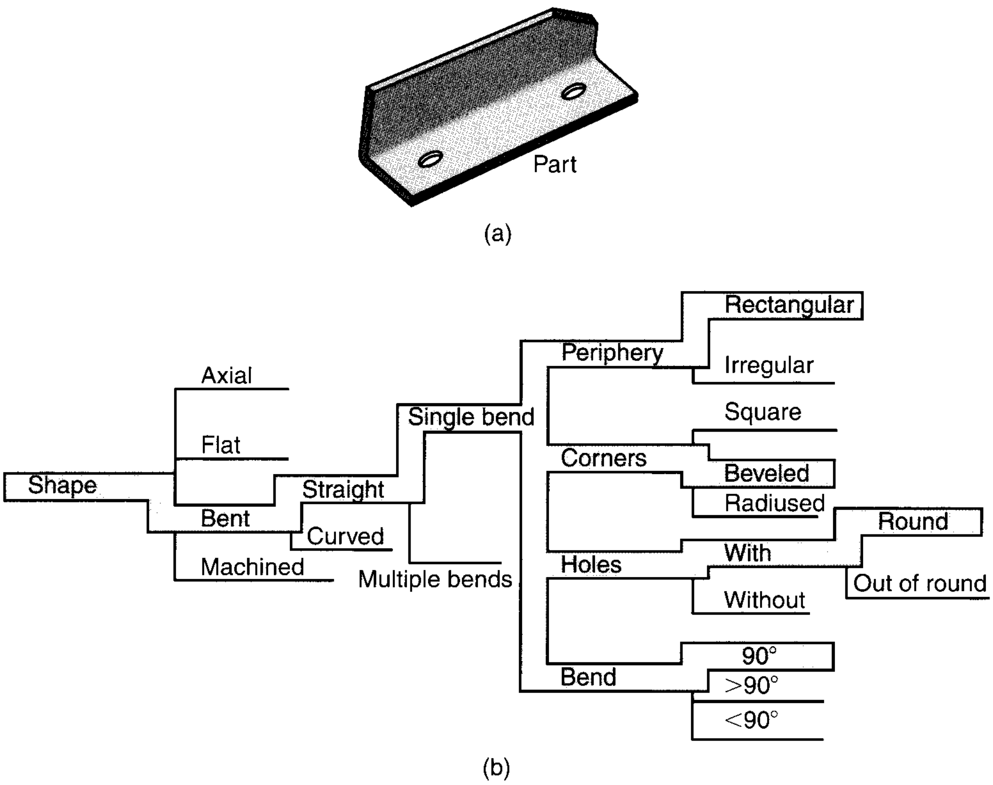


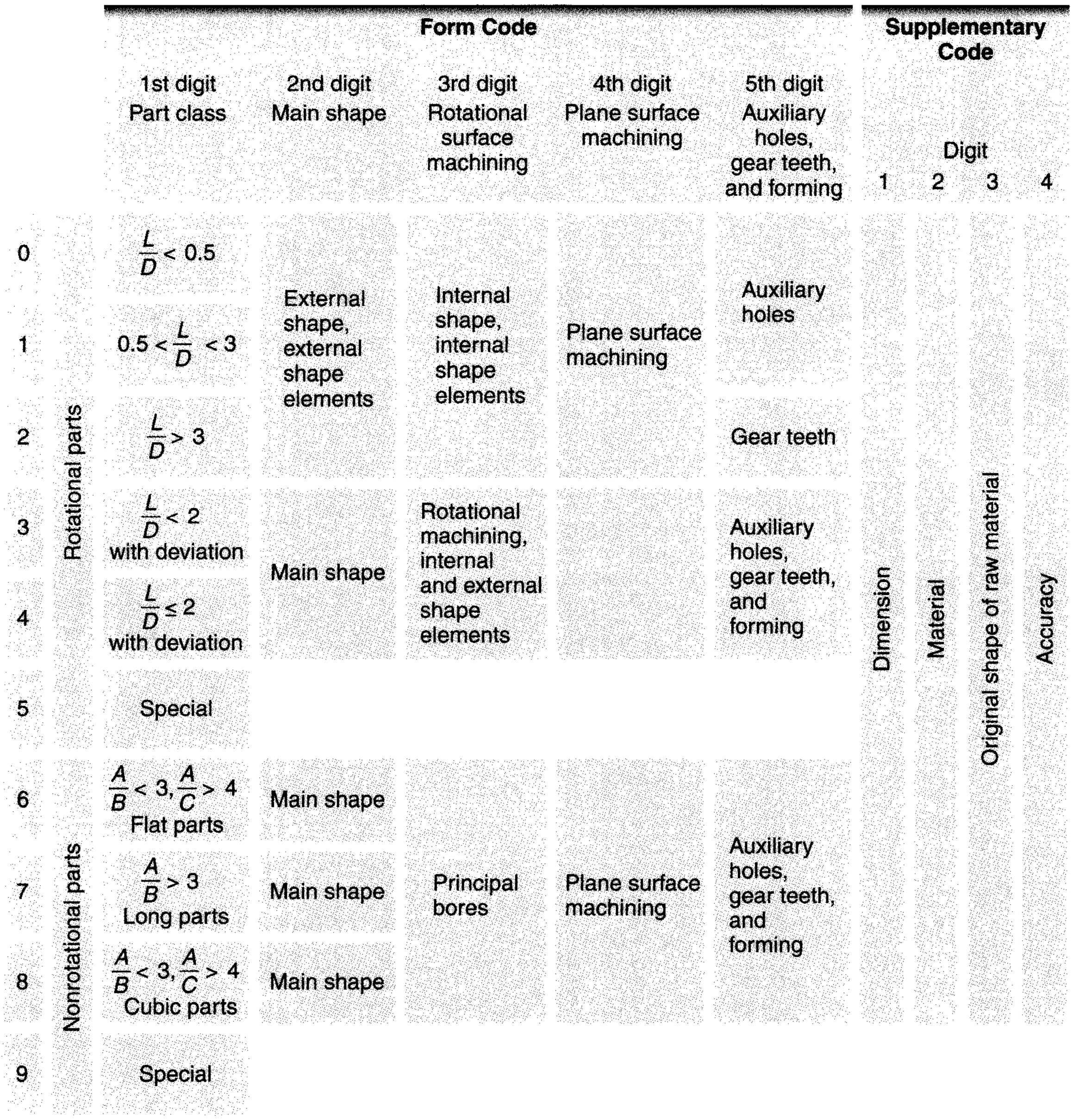
FIGURE 38.14 Decision-tree classification for a sheet-metal bracket. *Source*: After G.W. Millar.

38.8.4 Coding Systems

Three major industrial coding systems are described here:

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. 38.15). Four additional codes (1234) may be used to identify the type and sequence of production operations 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.



		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	Digit 2	Digit 3	Digit 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 38.15 Classification and coding system according to Opitz, consisting of a form code of five digits and a supplementary code of four digits.

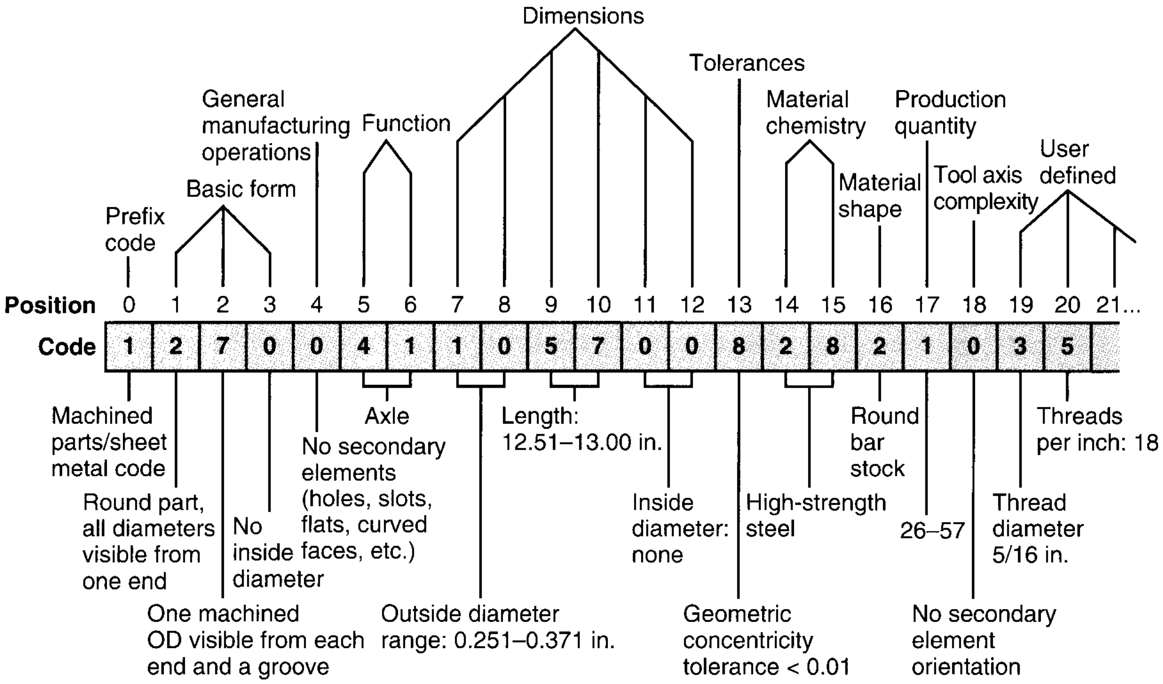


FIGURE 38.16 Typical multiClass code for a machined part. *Source*: Courtesy of Organization for Industrial Research. (Note: to convert in. to mm, multiply by 25.4.)

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 38.17 The structure of a KK-3 system for rotational components. *Source*: Courtesy of Japan Society for the Promotion of Machine Industry.

2. The **multiClass system** was developed to help automate and standardize several design, production, and management functions. It involves up to 30 digits (Fig 38.16). 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. It uses a 21-digit decimal system. This 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. 38.17.

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, and improving the control of both manufacturing and the management functions of an organization.
- Advanced software and powerful computers allow the description of part geometry in several different formats, including wire-frame, octree, surface models, solid models, skeletons, and boundary representations.
- Computers are also used to simulate manufacturing operations and systems, as well as 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. A number of group-technology coding systems are available.

KEY TERMS

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
Modeling
Octree representation
Paperless design
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, 2003.

Burbidge, J.L., **Production Flow Analysis for Planning Group Technology**, Oxford, 1997.

Chang, T.-C., Wysk, R.A., and Wang, H.P., **Computer-Aided Manufacturing**, 3rd ed., Prentice Hall, 2005.

Foston, A.L., Smith, C.L., and Au, T., **Fundamentals of Computer-Integrated Manufacturing**, Prentice Hall, 1991.

Giarratano, J.C., and Riley, G.D., **Expert Systems: Principles and Programming**, 4th ed., Course Technology, 2005.

Groover, M., **Automation, Production Systems and Computer-Integrated Manufacturing**, 3rd ed., Prentice Hall, 2008.

Gu, P., and Norrie, D.H., **Intelligent Manufacturing Planning**, Chapman & Hall, 1995.

Jackson, P., **Introduction to Expert Systems**, 3rd ed., Prentice Hall, 1998.

Louis, R.S., **Integrating Kanban with MRP II: Automating a Pull System for Enhanced JIT Inventory Management**, Productivity Press, 2005.

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, 2005.

Schodeck, D., Bechthold, M., Griggs, J.K., Kao, K., and Steinberg, M., **Digital Design and Manufacturing: CAD/CAM Applications in Architecture and Design**, Wiley, 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 are the differences between the terms "computer-aided" and "computer-integrated"?

38.5. What are the advantages of CAD systems over traditional methods of design? Are there any limitations?

38.6. What do the following abbreviations mean: NURB, DAS, DFX, PDES?

38.7. Describe the purposes of process planning. How are computers used in such planning?

38.8. Describe the features of a routing sheet. Why is it necessary?

38.9. What is group technology? Why was it developed? Explain its advantages.

38.11. Describe what is meant by the term "manufacturing system." What are its benefits?

38.12. What does classification and coding mean in group technology?

QUALITATIVE PROBLEMS

38.13. Describe your observations regarding Figs. 38.1 and 38.2.

38.14. Give examples of primitives of solids other than those shown in Figs. 38.4a and b.

38.15. Describe your understanding of the octree representation in Fig. 38.6.

38.16. Explain the logic behind the arrangements shown in Fig. 38.13b.

38.17. What are the advantages of hierarchical coding?

38.18. Referring to Fig. 38.3, explain the advantages of a third-order piecewise Bezier curve over a B-spline or a conventional Bezier curve.

38.19. Describe situations that would require a design change at its larger end of the part in Fig. 38.5.

38.20. Describe your thoughts on the differences between e-commerce and traditional business practices.

SYNTHESIS, DESIGN, AND PROJECTS

38.21. How would you describe the principle of computer-aided manufacturing to an older worker in a manufacturing facility who is not familiar with computers?

38.22. Review various manufactured parts illustrated in this book, and group them in a manner similar to those shown in Fig. 38.12. Explain.

38.23. Think of a simple part and make a decision-tree chart similar to that shown in Fig. 38.14.

38.24. Review the machine arrangements in Fig. 38.13 and suggest changes that may improve the flow of materials and parts.

38.25. Think of a simple product and make a routing sheet, similar to that shown in Fig. 38.10. If the same part is given to another person, what is the likelihood that the routing sheet developed will be the same? Explain.

38.26. Review Fig. 38.10, 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.

Computer-Integrated Manufacturing Systems

CHAPTER 39

- This chapter describes how computer systems and communications networks affect product development and manufacturing through the integration of all of their activities.
- The chapter begins by explaining the principles of manned and untended manufacturing cells and their features, and how cells can be integrated into flexible manufacturing systems.
- The new concept of holonic manufacturing and its applications are reviewed.
- Just-in-time production, lean manufacturing, and communication systems are then described.
- The chapter ends with a discussion of artificial intelligence and expert systems as applied to manufacturing.

EXAMPLES:

39.1 Introduction

This chapter focuses on the **computer integration of manufacturing activities.** *Integration* means that manufacturing processes, operations, and their management are 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 rapidly responding to changes in product types and fluctuating demands, as well as ensuring *on-time delivery* of products to the customer. Failure of on-time delivery in a highly competitive global environment can upset management plans and production schedules and, consequently, can have major adverse effects on a company's operations.

This chapter describes the key elements that enable the execution of the functions necessary to achieve flexible manufacturing. The chapter begins with *cellular manufacturing,* which is the basic unit of flexibility in the production of goods. It shows that manufacturing cells can be broadened into *flexible manufacturing systems,* with major implications for the production capabilities of an operation. *Holonic manufacturing* is then described, which is a new concept of how manufacturing units can be organized to achieve higher efficiency in operations.

The important concept of *just-in-time production* is examined, in which parts are produced "just in time" to be made into subassemblies, assemblies, and final products. This method eliminates the need for inventories (which can be 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 *communication networks* constitute a critical component of the overall operation. The chapter concludes with a review of *artificial intelligence*, which consists of expert systems, natural-language processing, machine vision, artificial neural networks, and fuzzy logic.

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 (called a *group-machine cell*), with each machine performing a different operation on a part. The machines can be modified, retooled, and regrouped for different product lines within the same family of parts.

Cellular manufacturing has 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 this 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 attended (*manned*) 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 located centrally in the cell.

Flexible Manufacturing Cells. Recall that, in view of rapid changes in market demand and the need for more product variety in smaller quantities, flexibility in manufacturing operations is highly desirable. 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). An example of an attended *flexible manufacturing cell* (FMC) that involves machining operations is illustrated in Fig. 39.1. Note that an automated guided vehicle moves parts between machines and inspection stations; machining centers fitted with automatic tool changers and tool magazines have an ability to perform a wide variety of operations. (See Section 25.2.)

A computer-controlled inspection station with a coordinate-measuring machine can similarly inspect dimensions on a wide variety of parts. Thus, the organization of these machines into a cell can allow the successful 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 parts. (The actual delay is the time required to download new machining instructions.)

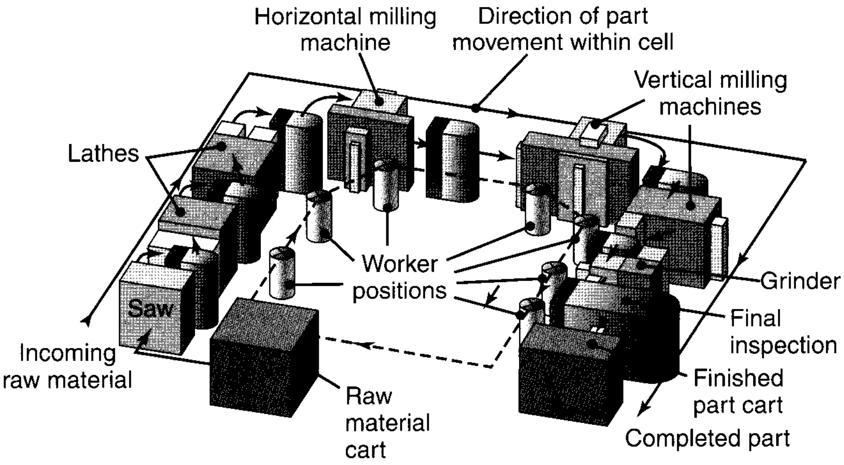


FIGURE 39.1 Schematic illustration of a manned flexible manufacturing cell, showing various machine tools and an inspection station. *Source*: After JT. Black.

Flexible manufacturing cells are usually *unattended* or *unmanned*; their design and operation are thus more exacting than those for other cells. The selection of machines and industrial robots, including the types and capabilities of end effectors and their control systems, is critical to the proper functioning of the FMC. The likelihood of a significant change in demand for part families should be considered during the design of the cell, in order to ensure that the equipment involved has the necessary flexibility and capacity. As with other flexible manufacturing systems (described in Section 39.3), the cost of FMCs is very high. However, this disadvantage is outweighed by increased productivity (at least for batch production), 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. The machines may be arranged along a line or in a U-shape, an L-shape, or a loop. Selecting the best machine and material-handling equipment arrangement also involves taking into account such factors as the production rate, the type of product, and its shape, size, and weight. The cost of flexible cells can be high, but this disadvantage is outweighed by increased productivity, flexibility, and controllability.

EXAMPLE 39.1 Manufacturing Cells in a Small Machine Shop

What follows is an actual example of the application of the manufacturing-cell concept in a small shop. Company A has only 10 employees, 11 milling machines, and 11 machining centers. The machines are set up in cells (milling cells and turning cells). The machines in the cells are arranged so as to allow an operator to machine a part in the most efficient and precise manner. Each cell allows the operator to monitor the performance of the machines in the cell.

Over 1200 different product lots have been produced over the years, with quantities ranging from 1 part to as many as 35,000 parts of the same design. The parts are inspected as they are produced. Each employee in the shop is involved in the programming and the running of the machines and in the in-process inspection of parts.

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). First 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. All of these are interfaced with a central computer. Different computer instructions can be downloaded for each successive part passing through a particular workstation. The system can handle a *variety of part configurations* and produce them *in any order*. A general view of an FMS installation in a plant is shown in Fig. 39.3.

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 materials, (c) measurement and inspection, and (d) assembly. The most common applications of FMS to date have been in machining and assembly operations.

An 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 a lower unit cost.
- Direct labor and inventories are reduced or eliminated.

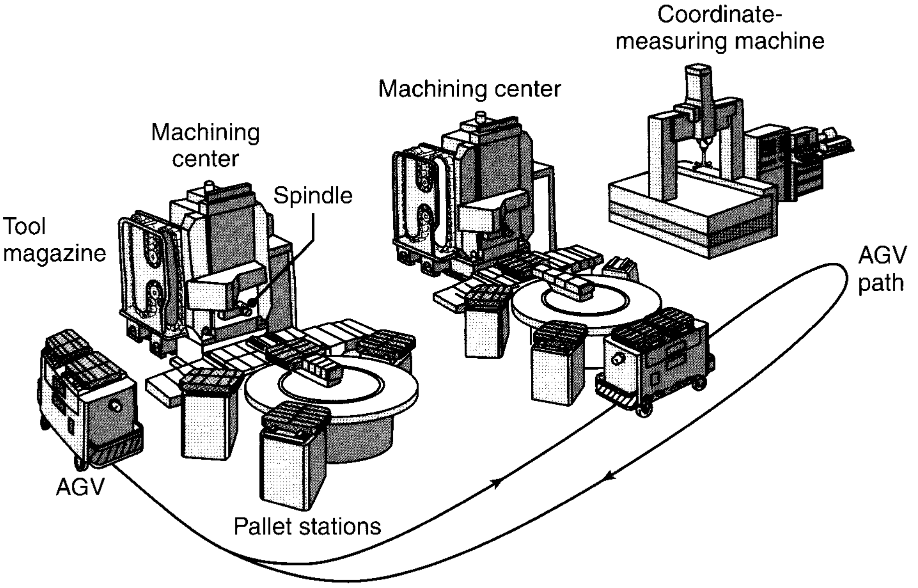


FIGURE 39.2 A schematic illustration of a flexible manufacturing system, showing machining centers, a measuring and inspection station, and automated guided vehicles. *Source*: After JT. Black.

- The lead times required for product changes are shorter.
- Because the system is self-correcting, production is more reliable and product quality is uniform.

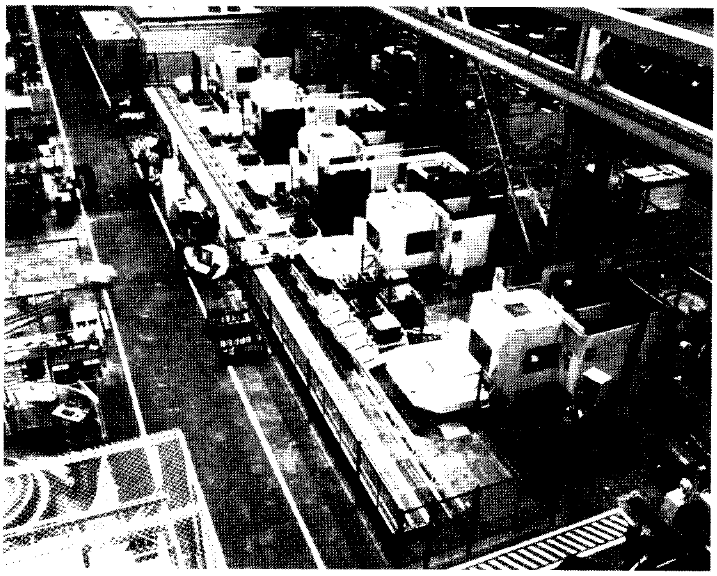


FIGURE 39.3 A general view of a flexible manufacturing system in a plant, showing several machining centers and automated guided vehicles moving along the white line in the aisle. *Source*: Courtesy of Cincinnati Milacron, Inc.

Elements of FMS. The basic elements of a flexible manufacturing system are (a) workstations and cells, (b) automated handling and transport of materials and parts, and (c) control systems. The workstations are arranged to yield the greatest efficiency in production with an orderly flow of materials and parts in progress through the system.

The types of machines in workstations depend on the type of production. For example, for machining operations, they usually consist of a variety of 3- to 5-axis machining centers, CNC lathes, milling machines, drill presses, and grinders. Also included are various other pieces of equipment, such as that for automated inspection (including coordinate-measuring machines), assembly, and cleaning. Other types of manufacturing operations suitable for FMS are sheet-metal forming, punching and shearing, and forging. FMS also may incorporate furnaces, various machines, trimming presses, heat-treating facilities, and cleaning equipment.

Because of the flexibility of FMS, material-handling systems are very important. These systems are controlled by a central computer, and their operations are performed by automated guided vehicles, conveyors, and various transfer mechanisms. 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* usually are moved on specially designed **pallets**. Parts having *rotational symmetry*, such as those used in turning operations, usually are moved by robots and various mechanical devices.

Scheduling. Because FMS are a major capital investment, efficient machine utilization is essential. Machines must not stand idle; consequently, proper scheduling and process planning are crucial. Scheduling for FMS is *dynamic*, unlike that in job

TABLE 39.1

Comparison of General Characteristics of Transfer Lines and Flexible Manufacturing Systems (FMS)

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

shops, where a relatively rigid schedule is followed to perform a set of operations. The scheduling system in flexible manufacturing specifies the types of operations to be performed on each part and identifies the machines or manufacturing cells on which these operations are to take place. Dynamic scheduling is capable of responding to quick changes in product type; thus, it is responsive to real-time decisions.

Because of the flexibility of FMS, no setup time is wasted in switching between manufacturing operations. However, the characteristics, performance, and reliability of each unit in the system must 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 before any final decision is made. This analysis must include such factors as the costs of capital, energy, materials, and labor; the expected markets for the products to be manufactured; and any anticipated fluctuations in market demand and product type. An additional important consideration is the time and effort required for installing and debugging the 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 of 15,000 to 35,000 aggregate parts per year. For individual parts with the same configuration, production may 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 machining centers (Chapter 25).

EXAMPLE 39.2 Flexible Manufacturing Systems in Large and Small Companies

Because of the advantages of FMS technology, many manufacturers have long considered implementing a large-scale system in their facilities. However, after detailed review, and on the basis of the experience of other companies, many companies 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 easier to control than an FMS.

When FMS became an established alternative, the expectations were high, and in some cases extensive computerization led to much confusion and inefficiency in company operations. Particularly for smaller companies, important considerations include not only the fact that a large capital investment and major hardware and software acquisitions are necessary, but also the fact that the efficient operation of a large FMS requires the extensive training of personnel. Often, the surprising result is found that an FMS leads to a manufacturing enterprise that is less lean (see Section 39.6).

There are several examples of the successful and economically viable implementation of an FMS in a large company. The results of a survey of 20 such operating systems in the United States indicated significant improvements over previous methods. Some systems are now capable of economically producing lot sizes of even one part. In spite of the high cost, the system has paid for itself in a number of companies.

39.4 Holonic Manufacturing

Holonic manufacturing is a new concept describing a unique organization of manufacturing units. The word *holonic* is from the Greek *holos* (meaning "whole") and the suffix *on* (meaning "a part of"). Thus, each component in a holonic manufacturing system (at the same time) is an independent entity (or *whole*) and a subservient *part* of a hierarchical organization. We describe this system here because of its potential beneficial impact on computer-integrated manufacturing operations.

Holonic organizational systems have been studied since the 1960s, and there are a number of examples in biological systems. Three fundamental observations 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-systems 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 both (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 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 between holons, as opposed to traditional hierarchical control, where individual processing power is essential. A large number of specific arrangements and software algorithms have been proposed for holarchical systems, but a detailed description of these is beyond the scope of this book. However, 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. For example, available holons may 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 breakdown, the unavailability of a particular machine, or 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 between 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 seamlessly integrate into a system.

39.5 Just-in-time Production

The *just-in-time* (JIT) *production* concept originated in the United States, but 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 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 the production is matched with demand for the final assembly of products.

There are no stockpiles and the ideal production quantity is one. JIT is also called *zero inventory, stockless production*, and *demand scheduling*. Moreover, parts are inspected as they are manufactured and are used within a short 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 of those operations and resources which do not add value are eliminated. This approach emphasizes (a) pride and dedication in producing high-quality products, (b) the elimination of idle resources, and (c) teamwork among workers, engineers, and management to quickly solve any problems that arise during production or assembly.

The ability to detect production problems as the parts are being made 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 inventories 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 can 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 JIT concept requires the timely delivery of all supplies and parts from outside sources and from other divisions of a company; thus, it 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. This approach requires reliable suppliers, close cooperation and trust between the company and its vendors, and a reliable system of transportation. Also important for smoother operation is a reduction in the number of its suppliers. In one example, an Apple Computer plant reduced the number of its suppliers from 300 to 70.

Advantages of JIT. Summarized here are the major advantages of just-in-time production:

- 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, just-in-time production has resulted in reductions of 20 to 40% in product cost, 60 to 80% in inventory, up to 90% in rejection rates, 90% in lead times, and 50% in scrap, rework, and warranty costs. Increases of 30 to 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 largely replaced by bar-coded plastic 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 contain information on the type of part, the location where the card was issued, the part number, and the 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.

39.6 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 or by considering a customer's product. (See also Section 40.9.1.) The goal of any organization is to produce a product that a customer wants, at a desired price, capability, location, time,

volume, etc. Providing the wrong good or service produces waste, even if it is provided efficiently. It is important to identify *all* of a manufacturer's 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 an activity:

   a. Clearly adds value.
   b. Adds no value, but cannot be avoided.
   c. Adds no value, but can be avoided.

2. **Identify value streams.** The *value stream* is the set of all actions required to bring a product to fruition, including
   a. Product design and development tasks, involving all actions from concept, to detailed design, to production launch;
   b. Information management tasks, involving order taking, a detailed schedule, and delivery;
   c. Physical production tasks, by means of which raw materials progress to a finished product in the hands of the customer.

   It has often been noted that no one person can be accurately described as a manufacturer of cars, boats, or airplanes. Organizations or systems create these products, and systems can become large and unwieldy, with tasks that do not create value. By identifying value streams, those tasks can be identified and eliminated.

3. **Make the value stream flow.** It has been noted that flow is easiest to achieve in mass production, but it is more difficult for small-lot production. Production in batches should be avoided, so just-in-time approaches (Section 39.5) are essential. The solution in such cases is to use manufacturing cells (Section 39.2), where minimum time and effort are required to switch from one product to another and a product being manufactured encounters continuous flow.

   In addition to just-in-time approaches, establishing product flow through factories requires the following:
   - Eliminating waiting time. Waiting time may be caused by unbalanced workloads, unplanned maintenance, or quality problems. Therefore, the efficiency of workers must be maximized at all times.
   - Eliminating unnecessary processes and steps, because they represent costs.
   - Minimizing or eliminating product transportation, because it represents an activity that adds no value. This waste can be either eliminated, by, for example, forming machining cells, or minimized, with, for instance, better plant layouts.
   - Performing time and motion studies to identify inefficient workers or unnecessary product motions.
   - Eliminating part defects.

4. **Establish pull.** Pull-type and push-type systems are described in Section 39.5. 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 physical production. Indeed, 90% reductions in physical production have been noted in some cases. Under these circumstances, it is possible to establish pull manufacturing, where products are produced upon order by a customer, 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, firms encounter an initial benefit, referred to as *kaikaku*, or "radical improvement."

39.7 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. Various network layouts (Fig. 39.4) 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. A number of advanced network protocols, including *ipV6* and *Internet2*, are under development, but have not yet been applied to manufacturing. Controlling access to the network is important; otherwise collisions can occur when several workstations are transmitting simultaneously. Continuous scanning of the transmitting medium is essential.

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. Other access-control methods are the *token ring* and *token bus*—in each of which a

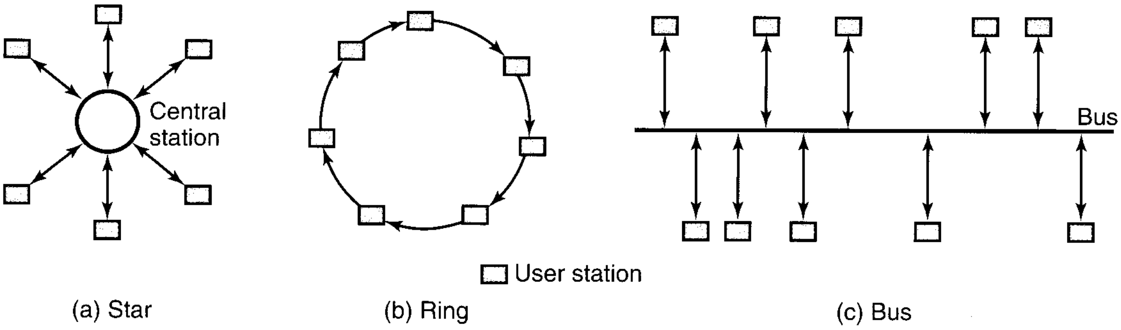


FIGURE 39.4 Three basic types of topology for a local area network (LAN). (a) The *star* topology is suitable for situations that are not subject to frequent configuration changes. All messages pass through a central station. Telephone systems in office buildings usually have this type of topology. (b) In the *ring* topology, all individual user stations are connected in a continuous ring. The message is forwarded from one station to the next until it reaches its assigned destination. Although the wiring is relatively simple, the failure of one station shuts down the entire network. (c) In the *bus* topology, all stations have independent access to the bus. This system is reliable and is easier than the other two to service. Because its arrangement is similar to the layout of the machines in the factory, its installation is relatively easy, and it can be rearranged when the machines are rearranged.

token (special message) is passed from device to device. Only the device that has the token is allowed to transmit, while all of the other devices only receive.

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, and 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.

Each cell's computers have their own specifications and proprietary standards, and they typically cannot communicate with others beyond the cell, unless they all are equipped with custom-built interfaces. This situation created *islands of automation*; in some cases, up to 50% of the cost of automation was related to overcoming difficulties in the communications between individual manufacturing cells and other parts of the organization.

The existence of automated cells that could function independently from each other (i.e., without a common base for information transfer) led to the need for a *communications standard* to improve both communications and the efficiency of computer-integrated manufacturing. The first step toward standardization began in 1980. After considerable effort, and on the basis of existing national and international standards, a set of communications standards known as the **manufacturing automation protocol** (MAP) was developed.

The International Organization for Standardization (ISO)/Open System Interconnect (OSI) reference model is accepted worldwide. The ISO/OSI model has a hierarchical structure in which communication between two users is divided into seven layers (Fig. 39.5). Each layer has a special task:

- Data transmission, by mechanical and electronic means
- Error detection and correction
- Correct transmission of the message
- Control of the dialog between users
- Translation of the message into a common syntax
- Verification that the data transferred has been understood.

The operation of this system is complex. Basically, each standard-sized chunk of message or data to be transmitted from User A to User B moves sequentially through the successive layers at A's end from Layer 7 to Layer 1. More information

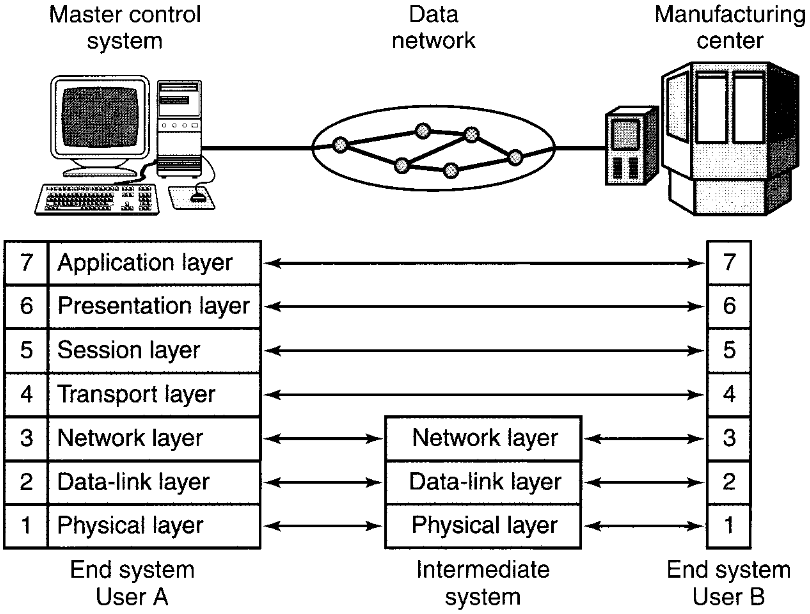


FIGURE 39.5 The ISO/OSI reference model for open communication. *Source*: After U. Rembold.

is added to the original message as it travels through each layer. The complete *packet* is transmitted through the physical communications medium to User B and then moves through the layers (from 1 to 7) at B's end. The transmission takes place through coaxial cable, fiber-optic cable, microwaves, and similar devices.

Communication protocols have been extended to office automation as well, with the development of **technical and office protocol** (TOP), which is based on the ISO/OSI reference model. In this way, total communication (MAP/TOP) is established among the factory floor and offices at all levels of an organization. A common practice is the use of **Internet tools** (hardware, software, and protocols) within a company to link all departments and functions into a self-contained and fully compatible **Intranet.** Several tools for implementing this link are available commercially; they are inexpensive and are easy to install, integrate, and use.

39.8 Artificial Intelligence

Artificial intelligence (AI) is that part of computer science which is concerned with systems that exhibit some characteristics usually associated with intelligence in human behavior, such as learning, reasoning, problem solving, recognizing patterns, and understanding language. The goal of AI is to *simulate* such human 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 a major effect on the design, automation, and overall economics of the manufacturing operation, largely because of advances in computer memory expansion (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 personal computers. Thus, AI has become accessible to office desks and shop floors.

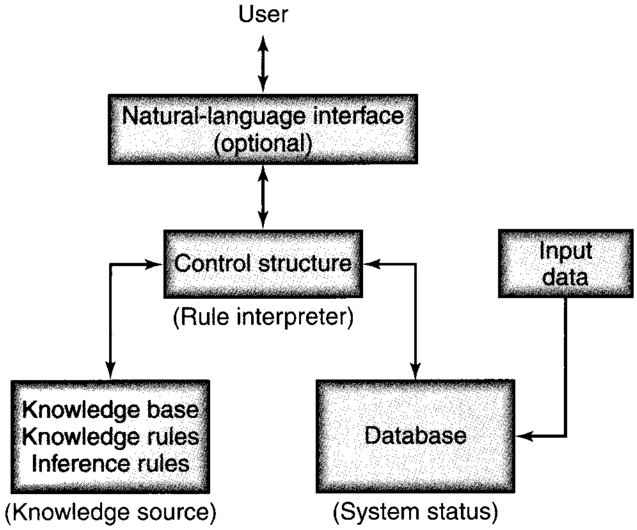


FIGURE 39.6 Basic structure of an expert system. The knowledge base consists of knowledge rules (general information about the problem) and inference rules (the way conclusions are reached). The results may be communicated to the user through the natural-language interface.

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 the use of **knowledge-based** and **inferential** procedures (Fig. 39.6). The goal of an expert system is to conduct an intellectually demanding task in the way that a human expert would.

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 facts, data, definitions, and assumptions. They also have the capability of adopting a **heuristic** approach—that is, making good judgments on the basis of discovery and revelation and making high-probability guesses, just as a human expert would. The knowledge base is expressed in computer code, usually in the form of **if-then rules**, and can generate a series of questions. The mechanism for using these rules to solve problems is called an **inference engine**. 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) a great deal of knowledge and (b) a mechanism for manipulating that knowledge to create solutions. Because of the difficulty involved in accurately modeling the many years of experience of an expert or a team of experts, and because of 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.

Expert systems operate on a real-time basis, and their short reaction times provide rapid responses to problems. The programming languages most commonly used for these applications are C++, Java, and Prolog; other languages also are available. An important development is expert-system software **shells** or **environments**, also called **framework systems**. These software packages are essentially expert-system outlines that allow a person to write specific applications to suit special needs. Writing these programs requires considerable experience and time.

Several expert systems have been developed that utilize computers with various capabilities and for such specialized applications as the following:

- Problem diagnosis in various types of machines and equipment, and the determination of corrective actions
- Modeling and simulation of production facilities
- Computer-aided design, process planning, and production scheduling
- Management of a company's manufacturing strategy.

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 are in various stages of development. These interfaces allow a user to obtain information by entering English or other language commands in the form of simple, typed questions.

Software shells are available, and they are used in such applications as scheduling material flow in manufacturing and analyzing information in databases.

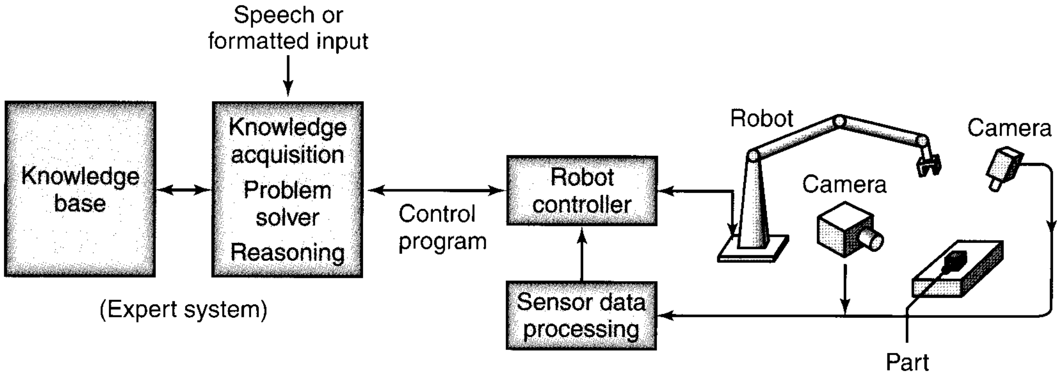


FIGURE 39.7 Expert system as applied to an industrial robot guided by machine vision.

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. The basic features of machine vision were described in Section 37.7.1. Computers and software, implementing artificial intelligence, are combined with cameras and other optical sensors (Fig. 39.7). These machines then perform such operations as inspecting, identifying, and sorting parts, as well as guiding robots (*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 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 senses (sight, hearing, smell, taste, and touch) simultaneously (called *data fusion*) and in real time. 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, and 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 now is believed that human learning is accomplished by changes in the strengths of these signal connections between 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. Although still controversial, the opinion of many is that true artificial intelligence will evolve only through advances in ANN.

Fuzzy Logic. An element of AI having important applications in control systems and pattern recognition is *fuzzy logic*, also called *fuzzy models*. Introduced in 1965 and based on the observation that people can make good decisions on the basis of imprecise and nonnumeric information, fuzzy models are mathematical means of representing vagueness and imprecise information—hence the term "fuzzy."

These models have the ability to recognize, represent, manipulate, interpret, and utilize data and information that are vague or lack precision. Fuzzy models deal with reasoning and decision making at a level higher than neural networks. Typical linguistic examples are the words: *few, very, more or less, small, medium, extremely,* and *almost all.*

Fuzzy technologies and devices have been developed (and successfully applied) in areas such as robotics and motion control, image processing and machine vision, machine learning, and the design of intelligent systems. Some applications are in (a) the automatic transmission in automobiles; (b) a washing machine that automatically adjusts the washing cycle for load size, type of fabric, and amount of dirt; and (c) a helicopter that obeys vocal commands to go forward, up, left, and right, and to hover and land.

39.9 Economic Considerations

The economic considerations in implementing the various computer-integrated activities described in this chapter are critical in view of the many complexities and the high costs involved. Flexible manufacturing system installations are very capital intensive; consequently, a thorough cost–benefit analysis must be conducted before a final decision is made. This analysis should include such factors as the following:

- The cost of capital, energy, materials, and labor.
- Expected markets for the products to be produced.
- Anticipated fluctuations in market demand and in the type of product.
- The time and effort required for installing and debugging the system.

Typically, an FMS system can take two to five years to install and at least six months to debug. Although FMS requires few, if any, machine operators, the personnel in charge of the total operation must be trained and highly skilled. These personnel 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 to 35,000 aggregate 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 (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 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. With the extensive use of computers and the rapid developments in sophisticated software, many aspects of manufacturing—including product designs, their analysis, and their simulation—are now highly detailed and thorough.

- 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.
- 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
Cellular manufacturing
Communications network
Communications standard
Computer-integrated manufacturing systems
Environments
Ethernet
Expert systems
Flexible manufacturing
Flow
Framework systems
Fuzzy logic
Holonic manufacturing
If-then rules
Inference engine
Internet tools
Just-in-time production
Kanban
Knowledge engineering
Knowledge-based system
Lean manufacturing
Local area network
Machine vision
Manufacturing cell
Natural-language processing
Pallet
Pull system
Push system
Technical and office protocol
Value
Waste
Wireless networks
Zero inventory

BIBLIOGRAPHY

Allen, J., Robinson, C., and Stewart, D., **Lean Manufacturing: A Plant Floor Guide**, Society of Manufacturing Engineers, 2001.

Baudin, M., **Lean Assembly**, Productivity Press, 2002.

Bedworth, D.D., Henderson, M.R., and Wolf, P.M., **Computer-Integrated Design and Manufacturing**, McGraw-Hill, 1991.

Biekert, R., **CIM Technology—Fundamentals and Applications**, Goodheart-Willcox, 1998.

Black, JT., and Hunter, S.L., **Lean Manufacturing Systems and Cell Design**, Society of Manufacturing Engineers, 2003.

Bodek, N., **Kaikaku: The Power and Magic of Lean**, PCS Press, 2004.

Cheng, T.C.E., and Podolsky, S., **Just-in-Time Manufacturing: An Introduction**, 2nd ed., Chapman & Hall, 1996.

Connor, G., **Lean Manufacturing for the Small Shop**, Society of Manufacturing Engineers, 2001.

Hales, H.L., Anderson, B., and Fillmore, W., **Planning Manufacturing Cell Textbook**, Society of Manufacturing Engineers, 2002.

Hitomi, K., **Manufacturing Systems Engineering**, 2nd ed., Taylor & Francis, 1996.

Irani, S.A. (ed.), **Handbook of Cellular Manufacturing Systems**, Wiley, 1999.

Jordan, J., and Michel, F., **Lean Company: Making the Right Choices**, Society of Manufacturing Engineers, 2001.

Krishnamoorty, C.S., and Rajeev, S., **Artificial Intelligence and Expert Systems for Engineers**, CRC Press, 1996.

Leondes, C.T. (ed.), **Fuzzy Logic and Expert Systems Applications**, Academic Press, 1998.

Liebowitz, J. (ed.), **The Handbook of Applied Expert Systems**, CRC Press, 1997.

Liker, J., **The Toyota Way**, McGraw-Hill, 2003.

Monden, Y., **Toyota Production System: An Integrated Approach to Just-in-Time**, 3rd ed., Institute of Industrial Engineers, 1998.

Morgan, J.M., and Liker, J.K., **The Toyota Product Development System**, Productivity Press, 2006.

Ortiz, C.A., **Kaizen Assembly: Designing, Constructing, and Managing a Lean Assembly Line**, CRC Press, 2006.

Rehg, J.A., **Computer-Integrated Manufacturing**, Prentice Hall, 1994.

———, **Introduction to Robotics in CIM Systems**, 3rd ed., Prentice Hall, 1997.

Russell, S., and Norvig, P., **Artificial Intelligence: A Modern Approach**, 2nd ed., Prentice-Hall, 2002.

Shingo, S., **A Study of the Toyota Production System**, Productivity Press, 1989.

Singh, N., and Rajamani, D., **Cellular Manufacturing Systems: Design, Planning and Control**, Chapman & Hall, 1996.

Vajpayee, S.K., **Principles of Computer-Integrated Manufacturing**, Prentice Hall, 1998.

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 are the advantages of a communications standard?

39.7. What are the differences between ring and star networks?

39.8. Describe your understanding of holonic manufacturing.

39.9. What is Kanban? Explain.

39.10. What is lean manufacturing?

39.11. Describe the elements of artificial intelligence. Is machine vision a part of it? Explain.

QUALITATIVE PROBLEMS

39.12. In what ways have computers had an impact on manufacturing? Explain.

39.13. What advantages are there in viewing manufacturing as a system? What are the components of a manufacturing system?

39.14. Discuss the benefits of computer-integrated manufacturing operations.

39.15. Why is just-in-time production required in lean manufacturing?

39.16. Would machining centers be suitable in just-in-time production? Explain.

39.17. Give an example of a push system and of a pull system. Indicate the fundamental difference between the two methods.

39.18. Is there a minimum to the number of machines in a manufacturing cell? Explain.

39.19. Are robots always a component of a FMC? Explain.

39.20. Are there any disadvantages to zero inventory? Explain.

39.21. Give examples in manufacturing processes and operations in which artificial intelligence could be effective.

SYNTHESIS, DESIGN, AND PROJECTS

39.22. 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.

39.23. 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.

39.24. Describe your opinions concerning the voice-recognition capabilities of future machines and controls.

39.25. Can a factory ever be completely untended? Explain.

39.26. 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.

39.27. 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?

39.28. 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.

39.29. It has been suggested by some that artificial intelligence systems ultimately will be able to replace the human brain. Do you agree? Explain.

39.30. 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.

CHAPTER 40

Product Design and Process Selection in a Competitive Environment

- Manufacturing high-quality products at the lowest possible cost is critical in a global economy.
- This chapter discusses the many interrelated factors in product design, development, and manufacturing.
- The chapter begins with a discussion of product design and life-cycle considerations in design and manufacturing.
- Material and process selection, together with their effects on design and manufacturing, are then described, followed by a discussion of the important 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 relationships among numerous factors. It was indicated throughout this text that

1. Product design and selection of materials and manufacturing processes are interrelated, and
2. Designs are periodically modified to,
    a. Improve product performance,
    b. Strive for zero-based rejection and waste,
    c. Make products easier and faster to manufacture,
    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, and 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. Recall also that world-class status, like product quality, is not a fixed target for a company

EXAMPLES:

CASE STUDY:

to reach, but rather a *moving target*, rising to higher and higher levels as time passes (also known as *continued improvement*).

Although the *selection of materials* for products traditionally has required much experience, several databases and expert systems are now available that greatly facilitate the selection process that is aimed at meeting specific requirements. Also, in reviewing the materials used in existing products (from simple hand tools to automobiles and aircraft), there are numerous opportunities for the *substitution of materials* for improved performance and, especially, cost savings.

In the production phase, it is imperative that the *capabilities of manufacturing processes* be properly assessed as an essential guide to the ultimate selection of an appropriate process or sequence of processes. As described throughout this book, there usually is more than one method of manufacturing a product, its components, and its subassemblies.

Increasingly important are the *life-cycle assessment* and *life-cycle engineering* of products, services, and systems, particularly regarding their potentially adverse 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 a company 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 important overall manufacturing cost factors. It also introduces cost-reduction methods, including *value analysis*, which is a powerful tool to evaluate the cost of each manufacturing step relative to its contribution to a product's value.

40.2 Product Design

Those aspects which are relevant to *design for manufacture and assembly* (DFMA), as well as to competitive manufacturing, have been highlighted throughout various chapters of 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 a number of 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 and, consequently, reducing product cost.

Product Design Considerations. In addition to the design guidelines we have given 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 the following:

- Have all alternative designs been thoroughly investigated?
- Can the design be simplified and the number of its components minimized without adversely affecting its intended functions and performance?
- Can the design be made smaller and lighter?
- Are there unnecessary features in the product or some of its components, and if so, can they be eliminated or combined with other features?
- Have modular design and building-block concepts been considered for a family of similar products and for servicing and repair, upgrading, and installing options?
- Are the specified dimensional tolerances and surface finish unnecessarily tight, thereby significantly increasing product cost, and if so, can they be relaxed without any adverse effects?

TABLE 40.1

References to Various Topics in This Book (Page numbers are in parentheses)

Material Properties

Tables 2.1 (57), 2.2 (59), 2.3 (62), and Figs. 2.4, 2.6, 2.7, 2.8, 2.15, 2.16, 2.17, 2.29
Tables 3.1 (89), 3.2 (90), and Figs. 3.1, 3.2, 3.3
Tables 5.2 (139), 5.4 (141), and 5.5 (144)
Tables 6.3 through 6.10 (153–162)
Tables 7.1 (172), 7.2 (180), 7.3 (186)
Tables 8.1 (199), 8.2 (202), 8.3 (206)
Tables 9.1 (218), 9.2 (220), 9.3 (228), and Figs. 9.3, 9.5, 9.7
Table 10.1 (248)
Table 11.3 (281)
Tables 12.3 (304), 12.4 (305), 12.5 (305), and Fig. 12.4
Tables 16.2 (392), 16.3 (398), 16.4 (409), and Fig. 16.14
Tables 17.3 (455), 17.4 (455), 17.5 (456), and Fig. 17.10
Table 20.2 (529)
Tables 22.1 (593), 22.2 (594), 22.3 (594), 22.5 (600), and Figs. 22.1, 22.9
Table 26.1 (721)
Table 32.3 (931)

Manufacturing Characteristics of Materials

Table I.3 (16)
Table 4.1 (120)
Table 5.8 (147)
Table 6.2 (152)
Tables 12.1 (297), 12.6 (308)
Table 14.3 (348)
Table 16.3 (398), and Fig. 16.33
Tables 17.1 (447), 17.2 (453)
Tables 21.1 (559), 21.2 (571)
Fig. 22.2

Dimensional Tolerances and Surface Finish

Table 11.2 (261)
Table 23.1 (617), and Figs. 23.13, 23.14
Fig. 25.16
Fig. 27.4
Figs. 35.19, 35.20
Figs. 40.4, 40.5

Capabilities of Manufacturing Processes

Tables 11.1 (259), 11.2 (261)
Table III.1 (315)
Table 14.1 (337), 14.4 (353)
Table 16.1 (383)
Section 17.7 and Fig. 17.14
Table 18.1 (466)
Tables 19.1 (485), 19.2 (521)
Table 20.1 (528)
Tables 23.1 (617), 23.6 (627), 23.8 (635), 23.10 (644)
Tables 26.3 (735), 26.4 (736)
Table 27.1 (761)
Tables 28.1 (801), 28.2 (809), 28.3 (810), and Fig. 28.19
Tables 29.1 (847), 29.3 (856)
Table VI.1 (864)
Table 30.1 (866)
Table 32.4 (932)
Table 34.1 (989)
Table 37.2 (1056), and Fig. 37.3
Table 39.1 (1121)
Tables 40.3 (1146), 40.5 (1153), and Figs. 40.2, 40.3, 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 16.13

General Costs and Economics

Tables I.5 (30), I.6 (33), I.7 (34), and Section I.10
Table 6.1 (152)
Section 12.4
Section 14.9
Section 16.15
Table 17.6 (461) and Section 17.8
Table 19.2 (521) and Section 19.16
Section 25.8
Section 26.9 and Fig. 26.35
Section 27.11
Section 31.8
Section 32.7
Section 37.11
Section 39.9
Table 40.6 (1157) and Section 40.9

- Will the product be difficult or excessively time consuming to assemble and disassemble for maintenance, servicing, or recycling of some or all of its components?
- Is the use of fasteners minimized, including their quantity and variety?
- Have environmental considerations been taken into account and incorporated into product design and material and process selection?

- Have green design and life-cycle engineering principles been applied, including recycling and cradle-to-cradle considerations?
- Can any of the design activities be outsourced?

40.2.1 Product Design and Quantity of Materials

Depending on the particular product, the cost of materials can become a significant portion of the total cost. Although material costs cannot be reduced below the often-fluctuating market level, reductions can be made in the *quantity* of the materials used in each of the components of a product. The wide use of available techniques, such as minimum-weight design; design optimization; and computer-aided design, manufacturing, and assembly, as well as the availability of vast resources on materials and their characteristics, have greatly facilitated design analysis, material selection and process, and overall optimization.

Significant reductions in the quantity of materials purchased can be achieved by reducing the component's volume or using materials with higher strength-to-weight or stiffness-to-weight ratios. 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 (such as I-beams and channels) or by using tubular or hollow components instead of solid bars.

Implementing such design changes may, however, present significant challenges in manufacturing. Consider, for example, the following:

a. 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).

b. Forging of thin sections requires high forces, due to friction, and especially in hot forging, due to rapid chilling of tin regions (Section 14.3).

c. Impact extrusion of thin-walled parts can be difficult, especially when high dimensional accuracy and symmetry are required (Section 15.4.1).

d. The formability of sheet metal may be reduced as sheet thickness decreases; it also can lead to buckling of the part under the high compressive stresses developed in the plane of the sheet during forming (Section 16.3).

e. Machining and grinding of thin workpieces may lead to part distortion, poor dimensional accuracy, and vibration and chatter (Section 26.5); consequently, advanced machining processes have to be considered (Chapter 27).

f. Welding thin sheets or slender structures can cause significant distortion due to thermal gradients developed during welding (Section 30.10).

Conversely, making parts with thick cross sections can have their own adverse effects. Consider, for example, the following:

a. In processes such as die casting (Section 11.4.5) and injection molding (Section 19.3), the production rate 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 mold.

b. Porosity can develop in thicker regions of castings, unless controlled (Fig. 10.14).

c. The bendability of sheet metals decreases as their thickness increases (Section 16.5).

d. In powder metallurgy, there can be significant variations in density and, hence, properties, throughout parts with varying thicknesses (Section 17.6).

e. Welding thick sections can present problems in the quality of the welded joint (Section 30.9).

f. In die-cast parts, thinner sections will have a higher strength per unit thickness (because of the smaller grain size developed), compared with thicker sections (Section 11.4.5).

g. Processing plastic parts requires increased cycle times as their thickness or volume increases; this is because of the longer time required for the parts to cool sufficiently to be removed from the molds (Chapter 19).

EXAMPLE 40.1 An Application of Design for Manufacturing and Assembly

The redesign of the pilot's instrument panel for a military helicopter, built by McDonnell-Douglas, was considered with a view toward reducing the number of parts in the panel (and thus also its weight) and the time required for its fabrication and assembly. The components of the panel consisted of sheet metal, extrusions, and rivets.

Using DFMA software and analyzing the panel in detail, it was estimated that the redesign would lead to the following changes: (a) the number of parts, from 74 down to 9; (b) the panel weight, from 3.00 kg to 2.74 kg; (c) fabrication time, from 305 hrs to 20 hrs; (d) assembly time, from 149 hrs to 8 hours; and (e) total production time, from 697 hrs to 181 hrs. It also was estimated that, as a result of design modifications, cost savings would be 74%. On the basis of these results, other components of the instrument panel were subjected to such analysis as well, resulting in similar savings.

40.3 Product Quality and Life Expectancy

Product quality and the techniques involved in quality assurance and control are described in detail in Chapter 36. Recall that the word *quality* is difficult to define precisely, partly 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:

- It satisfies the needs and expectations of the customer.
- It has a pleasing appearance and handles well.
- It has high reliability and functions safely over its intended life.
- It is compatible with and responsive to the customer's capabilities and working environment.
- Installation, maintenance, and future improvements are easy to perform and at low cost.

A major priority in product quality is the concept of **continuous improvement**, as exemplified by the Japanese term **kaizen**, meaning *never-ending improvement.* Note, however, that the level of quality a manufacturer chooses to impart to a particular product depends on the particular market for which the product is intended. For example, low-quality, low-cost products have their own market niche, including what are commonly referred to as dollar stores. Conversely, there always is a market for high-quality, expensive products, such as a Rolls-Royce automobile, a gold and diamond-studded wristwatch, high-performance sports equipment, and a high-precision machine tool.

40.3.1 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 major influence on customer satisfaction.

- An incremental improvement in quality vis-a-vis the additional costs involved must be carefully investigated.
- There must be a certain limit on how much should be spent on quality improvements.
- Because quality can be rather subjective, all changes to be made must be critically evaluated.

Although customer satisfaction is a qualitative factor and is difficult to include in calculations, satisfaction is increased and customers are more likely to be retained (and become repeat customers) when there are no defects in products.

On the one hand, high-quality products do not necessarily cost more. For example, in industries making computer chips and computer hardware, the ROQ is minimized while the aim is to approach zero defects. (See also *six sigma*, Section 36.7.2.) On the other hand, there are other products, such as ordinary door hinges, water faucets, and hubcaps, for which the additional cost involved in eliminating the final few defects can be unnecessarily high. It also is important to consider 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.

Life Expectancy of Products. The *average life expectancies* of products are given in Tables I.4 and 36.1. As expected, life expectancies within each group of products can vary significantly; the variations will depend on the materials and production processes employed.

40.4 Life-cycle Assessment and Sustainable Manufacturing

Life-cycle assessment (LCA) is defined, according to the ISO 14000 standard (Section 36.6.3), 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." The **life cycle** involves consecutive and interlinked stages of a product or a service, from the very beginning to its disposal or recycling, and includes the following:

a. Extraction of natural resources.
b. Processing of raw materials.
c. Manufacturing of products.
d. Transportation and distribution of the product to the customer.
e. Use, maintenance, and reuse of the product.
f. Recovery, recycling, and reuse of the components of the product, or else their disposal, including metalworking fluids, cleaning solvents, and various liquids used in heat-treating and plating processes.

All of these factors are basically applicable to any type of product. Recall that each type of product has its own metallic and nonmetallic materials, processed into individual components and assembled; thus, each product has its own life cycle. Moreover, (a) some products, particularly those made of paper, cardboard, inexpensive plastic, and glass, are intentionally made to be disposable, but nonetheless are recyclable, and (b) numerous other products are completely reusable.

Life-cycle Engineering. The major aim of *life-cycle engineering* (LCE) is to consider *reusing* and *recycling* the components of a product, beginning with the earliest stage: product design. Life-cycle engineering is also called **green design** or **green engineering.** The considerations involved include environmental factors, optimization, and numerous technical factors regarding each component of a product.

Although life-cycle analysis and engineering are comprehensive and powerful tools, their implementation can be challenging, time consuming, and costly, largely because of uncertainties (regarding materials, processes, long-term effects, costs, etc.) in the input data and the time required to collect reliable data to properly assess the often complex interrelationships among all the components of the whole system. Various software is being developed to expedite these analyses, particularly for the chemical and process industries, because of the higher potential for environmental and ecological damage in their operation. Examples of such software include FeaturePlan and Teamcenter, which runs in a ProEngineer environment.

Cradle-to-cradle Production. In examining the importance of product life-cycle considerations, the principles of *cradle-to-grave* and *cradle-to-cradle* production were described in some detail in Section I.4. The case study that follows illustrates an application of this type of production.

CASE STUDY 40.1 Automobile Tires: From Cradle-to-grave to Cradle-to-cradle

Automobiles, buses, trucks, tractors, and motorcycles are pervasive in modern society. They are depended on for personal transport as well as for commerce, and they bring products to markets that are accessible to consumers.

One major environmental concern associated with such vehicles is caused by the need to periodically purchase new tires. Even though tires have surprisingly high wear resistance, they eventually (after about 65,000 km to 100,000 km) are not suited for further safe operation on vehicles. Traditionally, tires have been removed and replaced with new ones, with the discarded tires typically dumped in landfills. This practice has now become an environmental hazard as the number of discarded tires continues to grow. Each year in the United States, there are around 300 million tires discarded, mostly in landfills.

The traditional model of manufacturing a tire and taking into account its use and disposal is a classic example of *cradle-to-grave* production. However, tires do not fit into the biological *recycling* paradigm, because they do not readily break down into nutrients for organisms and thus take several decades to decompose in landfills. Although tires are flammable, their combustion produces harmful gases and particulates, making incineration an unviable option for disposal. As a result, the traditional life cycle has led to an accumulation of discarded tires.

The product life cycle for tires is being transformed by PermaLife products into a *cradle-to-cradle* model. The company takes discarded tires and processes them as follows:

- In a cryogenic freezing operation, the tire is separated into its major components: rubber, and steel and fiber reinforcement. The rubber is subjected to a temperature below its glass-transition temperature (Section 7.2.3), which, for the type of rubber tires are made of, is −115°C. The steel and fiber reinforcements in the tire are not significantly affected by this operation. At −115°C, the rubber becomes very brittle, and when processed in a hammer mill (shown in Fig. 17.6c), it shatters into millions of small pieces.
- The rubber, in the form of small particles, is then easily separated from the fiber or fiber mesh and the steel reinforcements, which are then recycled separately from the rubber.
- Colors can be added to the granulated material.

The resulting artificial mulch, referred to by its trade name of Permalife Softstuff™, has successfully been used as a playground material, for landscaping, and as a support material beneath artificial turf widely used in sports stadiums around the world. From a functional standpoint, this material can be tailored to the particular application: (a) For playgrounds, a soft material can be developed with a large-size granule to prevent injury in the case of a child falling, while also making its ingestion unlikely. (b) For professional athletic competitions, the material can be made stiffer, so that the likelihood of injury is low while athletic performance is optimized.

The major advantages of this material are as follows:

- The material will not cause splinters or attract insects (as opposed to natural wood mulch).
- Maintenance is less costly and more environmentally friendly, as opposed to natural grass surfaces, which require fertilizer and maintenance for optimum performance.
- The material does not stain, as opposed to wood or grass surfaces.

Source: Courtesy of M. Sergia and N. Menonna, Permalife, Inc.

Sustainable Manufacturing. As is now universally acknowledged, the natural resources on this Earth are limited, thus necessitating conservation of both materials and energy. The concept of *sustainable manufacturing* emphasizes the need for conserving resources, particularly through proper maintenance and reuse. While profitability is important to an organization, sustainable manufacturing is meant to meet purposes such as (a) increasing the life cycle of products, (b) eliminating harm to the environment and the ecosystem, and (c) ensuring our collective well-being, especially that of future generations.

EXAMPLE 40.2 Sustainable Manufacturing in the Production of Nike Athletic Shoes

Among numerous examples from industry, the production of Nike shoes indicates the benefits of sustainable manufacturing. These athletic shoes are assembled with the use of adhesives. 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.

Regarding another component of the shoe, the rubber outsoles 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 Fig. 14.5d). With about 40 factories using thousands of molds and producing over a million outsoles a day, flashing constitutes the largest chunk of waste in manufacturing the shoes. In order to reduce this 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 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 will discuss 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 susceptibility to oxidation and corrosion and to the various surface-treatment processes described in Chapter 34.

The following considerations are significant in the selection of materials for various products:

- Do the materials selected have the appropriate manufacturing characteristics?
- Can some of the materials be replaced by others that are less expensive?
- Do the materials under consideration have properties that meet minimum requirements and specifications?
- Are the raw materials (also called *stock*) specified available in standard shapes, dimensions, tolerances, and surface characteristics?
- Is the supplier of the materials reliable? Can the materials be delivered in the required quantities within the required time frame? Are there likely to be significant price increases or fluctuations?
- Does the material present any environmental hazards or concerns?

Material selection has become easier and faster because of the increasing availability of extensive computer databases that provide greater accessibility and accuracy. In order to facilitate the selection of materials, **expert-system software** (called **smart databases**, Section 39.8) has been developed. With the proper input of product design and functional requirements, these systems are capable of identifying appropriate materials for a specific application, just as an expert or a team of experts would.

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 40.2). Depending on the type of material (metal, polymer, ceramic, etc.) materials generally are available as castings, extrusions, forgings, powder metals, drawn rod and wire, and rolled bars, plates, sheets, and foil.

Purchasing materials in shapes that require the least amount of additional processing obviously is an important economic consideration. Also relevant are such characteristics as surface finish and quality, dimensional tolerances, straightness, and flatness. (See, e.g., Figs. 27.4, 23.13, and 23.14, and Table 11.2.) The better and the

TABLE 40.2

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; W = wire. Lowercase letters indicate limited availability.

more consistent these characteristics are, the less additional processing will be required. Note, for example, that if we want to produce simple shafts with good dimensional accuracy, roundness, straightness, and surface finish, then we could purchase round bars that are first turned or drawn and then centerless-ground (Fig. 26.22) to the dimensions specified. Unless the facilities in a plant have the capability of producing round bars economically, it generally is cheaper to purchase them. If we need to make a stepped shaft (i.e., a shaft having different diameters along its length, as shown in Fig. IV.3), we could purchase a round bar with a diameter at least equal to the largest diameter of the stepped shaft and then turn it on a lathe or process it by some other means in order to reduce its diameter.

Each manufacturing operation produces parts that have specific shapes, surface finishes, and dimensional accuracies. 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 cross-sectional 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.
- The wall thickness of welded tubing is generally more uniform than that of seamless tubing, which is typically produced by the Mannesmann process (Fig. 13.18).

Manufacturing Characteristics of Materials. Manufacturing characteristics of materials generally include castability, workability, formability, machinability, weldability, and hardenability by heat treatment. Raw materials have to be formed, shaped, machined, ground, fabricated, or heat treated into individual components having specific shapes and dimensions; consequently, a knowledge of their manufacturing characteristics is essential.

Recall that the quality of the raw material can greatly influence its manufacturing properties. The following are typical examples (see also individual processes):

- A bar with a longitudinal seam, or lap, will develop cracks during simple upsetting and heading operations.
- Round rods with internal defects such as hard inclusions will crack during further processing.
- Porous castings will develop a poor surface finish when subsequently machined.
- Parts that are nonuniformly heat treated and cold-drawn bars that are not properly stress relieved will distort during subsequent processing.
- Incoming stock that has variations in composition and microstructure cannot be heat treated or machined consistently and uniformly.
- Sheet-metal stock having variations in its cold-worked conditions will exhibit different degrees of springback during bending and other forming operations because of differences in yield stress.
- 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. There are several factors that influence the reliability of material supplies: shortages, strikes, geopolitical factors, and the reluctance of suppliers to produce materials in a particular shape or quality. 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 scrap metal, plastic bottles, etc.; it often requires that individual components of a product be taken apart and separated. Also, obviously, if much effort and time has to be expended in doing so, recycling may become prohibitively expensive. Some general guidelines to facilitate the process during the life cycle of a product are as follows:

- Reduce the number of parts and types of materials in products.
- Reduce the variety of product models.
- Use a modular design to facilitate disassembly.
- For plastic parts, use single types of polymers as much as possible.
- Mark plastic parts for ease of identification, as is done with plastic food containers and bottles (See Section 7.8).
- Avoid using coatings, paints, and plating; instead, use molded-in colors in plastic parts.
- Avoid using adhesives, rivets, and other permanent joining methods in assembly; instead, use fasteners, especially snap-in fasteners.

As an example of this type of 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 by 55%.

Cost of Materials and Processing. Because of its processing history, the unit cost of a raw material (typically, cost per unit weight) depends not only on the material itself, but also on its shape, size, and condition. For example, because more operations are involved in the production of thin wire than in that of round rod, the unit cost of the thin wire is much higher. Similarly, powder metals generally are more expensive than bulk metals. Furthermore, the cost of materials typically decreases as the quantity purchased increases. Likewise, certain segments of industry (such as automotive companies) purchase materials in very large quantities; the larger the quantity, the lower is the cost per unit weight (bulk discount).

Table 6.1 shows the cost per unit volume relative to that of carbon steel. The benefit of cost per volume can be seen by the following simple example: In the design of a steel cantilevered rectangular beam supporting a certain load at its end, a maximum deflection is specified. Using equations from handbooks, and assuming that the weight of the beam can be neglected, we can determine an appropriate cross section of the beam. Since all dimensions are now known, the volume of the beam can be calculated; then the cost of the beam can be determined by multiplying the volume by the cost of the material per unit volume. Note, on the other hand, if the cost is given per unit weight, we first have to calculate the weight of the beam and then determine the cost.

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 product is no longer cost competitive, alternative and less costly materials may 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, (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 at the junctions, because aluminum has a higher contact resistance than copper.

Scrap. When scrap is produced during manufacturing, as in sheet-metal fabricating, forging, and machining (Table 40.3), the value of the scrap is deducted from the material's cost in order to obtain the net material cost. As expected, the value of the scrap depends on the type of metal and on the demand for it; typically, it is between

TABLE 40.3

Approximate Percentages of Scrap Produced in Various Manufacturing Processes

Process	Scrap (%)	Process	Scrap (%)
Machining	10–60	Permanent-mold casting	10
Hot forging	20–25	Powder metallurgy	<5
Sheet-metal forming	10–25	Rolling	<1
Hot extrusion	15		

10 and 40% of the original cost of the material. Note that, in machining, scrap can be very high, whereas operations such as rolling, ring rolling, and powder metallurgy (all of which are net- or near-net-shape processes) produce the least scrap.

EXAMPLE 40.3 Effect of Workpiece Hardness on Cost in Drilling

Gear blanks forged from 8617 alloy steel and having a hardness range from 149 to 156 HB required the drilling of a hole 75 mm in diameter in the hub. The blanks were drilled with a standard helix drill. After only 10 pieces, however, the drill became dull, temperatures increased excessively, and the drilled holes had developed a rough internal surface finish. In order to improve machinability and reduce galling, the hardness of the gear blanks was increased to range from 217 to 241 HB by heating them to 840°C and then quenching them in oil. When blanks at this hardness level were drilled, galling was reduced, surface finish was improved, drill life increased to 50 pieces, and the cost of drilling was reduced by 80%.

Source: ASM International.

40.6 Material Substitution

There is hardly a product on the global market today for which the *substitution of materials* has not played a major role in helping companies maintain their competitive positions. 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.

Although new products continually appear on the market, the majority of the design and manufacturing activities is concerned with improving existing products. There are several reasons for substituting materials in existing products:

1. Reduce the costs of materials and processing.
2. Improve manufacturing, assembly, and installation, and allow conversion to automated assembly.
3. Improve the performance of the product, such as by reducing its weight and by improving resistance to wear, fatigue, and corrosion.
4. Increase stiffness-to-weight and strength-to-weight ratios.
5. Reduce the need for 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.

8. Improve robustness to reduce variations in performance or environmental sensitivity of the product.
9. Increase the ease of recycling for environmental reasons.

Substitution of Materials in the Automobile Industry. The automobile is a good example of the effective substitution of materials in order to achieve one or more of the objectives outlined previously. Some examples of material substitution in automobiles are as follows:

- 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 (see Section 5.5.6).
- Metallic engine components replaced with ceramic and composites parts.
- All-metal driveshafts replaced with composite-material driveshafts.
- 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.
- Leather seats in automobiles in some luxury cars (including Mercedes) can now be replaced (offered as an option) with synthetic materials in response to concerns raised by advocacy groups.

Because the automobile industry is a major consumer of both metallic and nonmetallic materials, there is constant competition among suppliers, particularly in steel, aluminum, and plastics. Industry engineers and management continually are investigating the relative advantages and limitations of these principal materials in their applications, recycling and other environmental considerations, and relative costs and benefits (in particular).

Substitution of Materials in the Aircraft and Aerospace Industries

- Conventional aluminum alloys (particularly 2000 and 7000 series) are being 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. (See Example 9.4.)
- Forged parts are being 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 are replacing traditional aluminum airframe components (Fig. 40.1), and metal-matrix composites are replacing some of the aluminum and titanium in structural components.

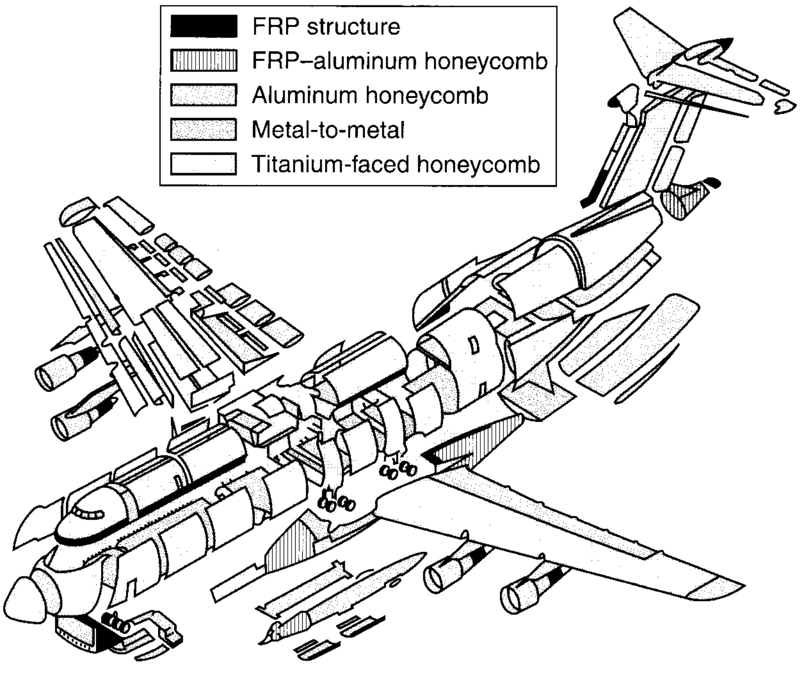


FIGURE 40.1 Advanced materials in the Lockheed C-5A transport aircraft. *Note*: FRP = fiber-reinforced plastic.

EXAMPLE 40.4 **Material Substitution in Common Products**

In the following list, the commonly available products can be made of either set of materials mentioned:

- **a.** Metal vs. wooden baseball bat.
- **b.** Metal vs. reinforced-plastic or wood handle for a hammer.
- **c.** Plastic vs. metal intake manifold.
- **d.** Cast-iron vs. aluminum lawn chair.
- **e.** Plastic vs. sheet-metal light-switch plate.

These products are given as typical examples, and on the basis of the topics covered in various chapters throughout this book, the choice of materials can be reviewed with regard to their respective advantages and limitations.

EXAMPLE 40.5 **Material Changes between C-5A and C-5B Military Cargo Aircraft**

Table 40.4 shows the changes made in materials for various components of the two aircraft listed and the reasons for the changes.

Source: After H.B. Allison, Lockheed-Georgia.

TABLE 40.4

Material Changes from C-5A to C-5B Military Cargo Aircraft

Item	C-5A Material	C-5B Material	Reason for change
Wing panels	7075-T6511	7175-T73511	Durability
Main frame:			
Forgings	7075-F	7049-01	Stress-corrosion resistance
Machined frames	7075-T6	7049-T73	Stress-corrosion resistance
Frame straps	7075-T6 plate	7050-T7651 plate	Stress-corrosion resistance
Fuselage skin	7079-T6	7475-T61	Material availability
Fuselage underfloor	7075-T6 forging	7049-T73 forging	Stress-corrosion resistance end fittings
Wing–pylon attach fitting	4340 alloy steel	PH13-8Mo	Corrosion prevention
Aft ramp lock hooks	D6-AC	PH13-8Mo	Corrosion prevention
Hydraulic lines	AM350 stainless steel	21-6-9 stainless steel	Improved field repair
Fuselage failsafe straps	6Al-4V titanium	7475-T61 aluminum	Titanium strap debonding

40.7 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. 40.2. Note, for instance, that sand casting (Section 11.2.1) cannot produce thin parts, whereas cold rolling (Section 13.3) is a process capable of producing very thin materials, as evidenced by a product such as aluminum foil.

Equally important as to overall dimensions are the capabilities of various processes to meet stringent dimensional tolerance and surface-finish requirements, as shown in Fig. 40.3. Note, for example, how sand casting is at the extreme opposite corner of microfabrication (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 boundaries that span three orders of magnitude. Thus, if a turning

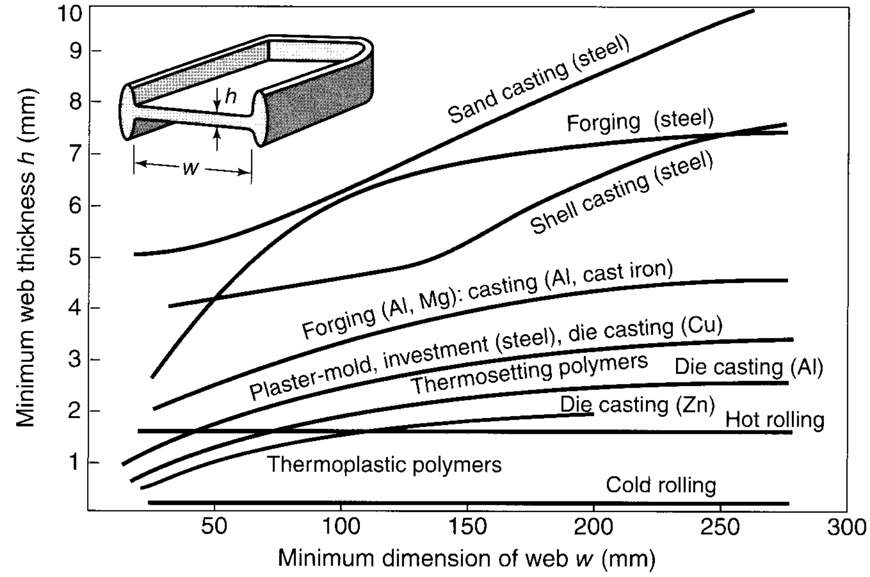


FIGURE 40.2 Manufacturing process capabilities for minimum part dimensions. *Source*: After J.A. Schey.

operation is carried out on an old lathe using inappropriate tools and processing parameters, then the tolerances and surface finish will, of course, be poor.

In the sections that follow, we describe important aspects of process capabilities as they relate to manufacturing processes and production operations.

Dimensional Tolerances and Surface Finish. The dimensional tolerances and surface finish produced are particularly important in subsequent assembly operations (because of possible difficulties in fitting the parts together for assembly) and in the proper operation of machines and instruments (because their performance can affect tolerances and finish). The dimensional tolerance and surface finish typically obtained by various manufacturing processes are illustrated qualitatively in Fig. 40.3.

Closer tolerances and better surface finish can be achieved by subsequent additional finishing operations (Section 26.7), but at higher cost, as shown in Fig. 40.4. Also, the finer the surface finish required, the longer is the manufacturing time (Fig. 40.5). In the machining of aircraft structural members made of titanium alloys, it has been observed that as much as 60% of the cost of machining may be expended in the final machining pass in order to maintain proper tolerances and surface finish. Thus, unless otherwise required, and with appropriate technical and economic justification, parts should be made 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. For example, bearings, bolts, spark plugs, plastic containers, tires, automobiles, and lawn mowers are produced in very large quantities, whereas jet engines, diesel engines, locomotives, and medical equipment are produced in 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.

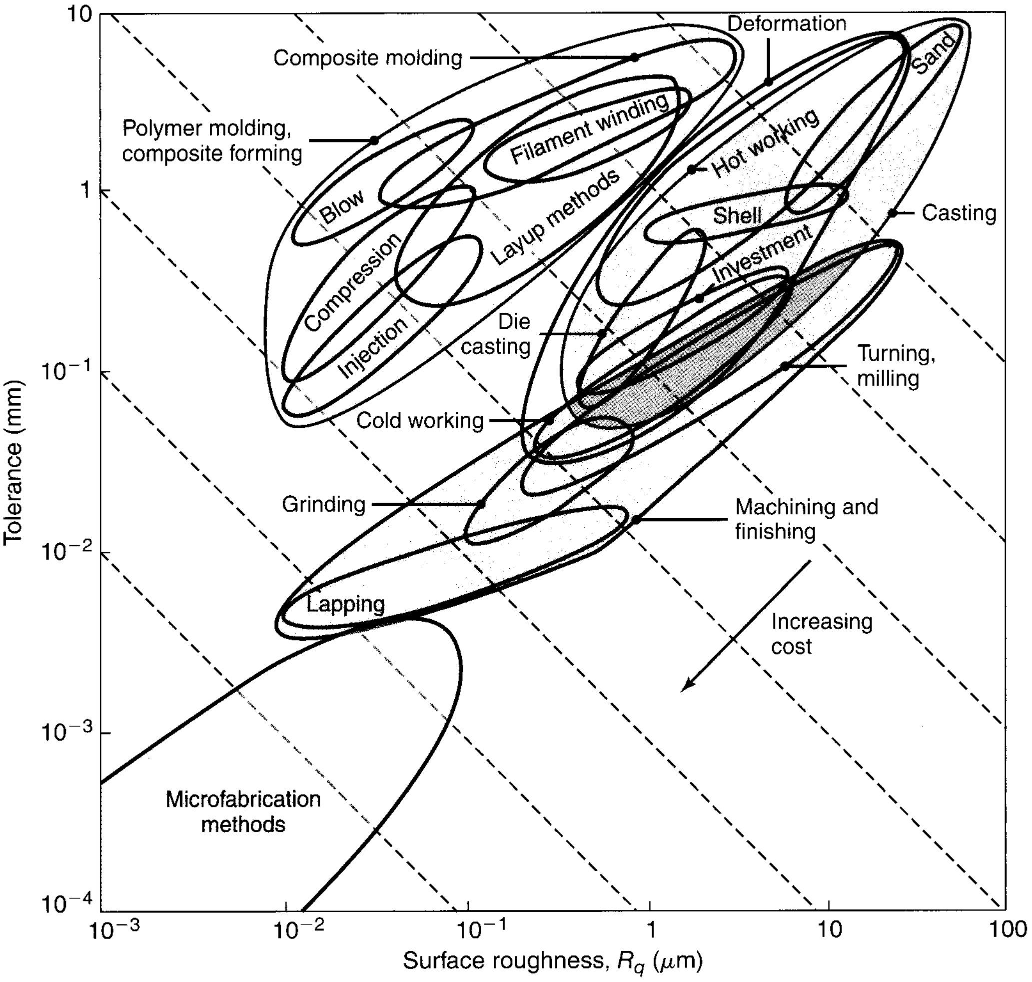


FIGURE 40.3 A plot of achievable tolerance versus surface roughness for assorted manufacturing operations. The dashed lines indicate cost factors, where an increase in precision corresponding to the separation of two neighboring lines gives an increase in cost for a given process (within a factor of two). *Source*: M.F. Ashby, *Materials Selection in Design*, Butterworth-Heineman, 1999.

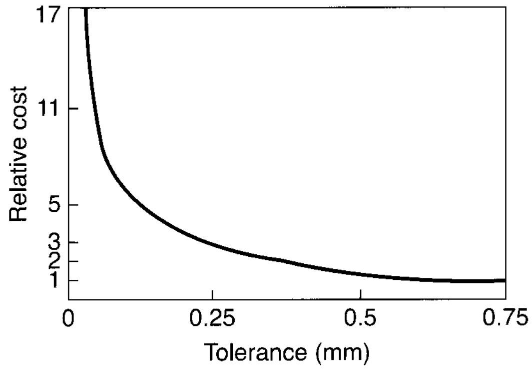


FIGURE 40.4 Dependence of manufacturing cost on dimensional tolerances.

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, such as per hour, per day, or per year. The production rate obviously can be increased by using multiple equipment and highly automated machines. Recall that processes such as die casting, powder metallurgy, deep drawing, wire drawing, and roll forming are high-production-rate operations. By contrast, sand casting, conventional and electrochemical machining, metal spinning, superplastic forming, adhesive and diffusion bonding, and the processing of reinforced plastics generally are relatively slow operations.

Lead Time. Lead time generally is defined as the length of time between the receipt of an order for a product and its delivery to the customer at a specified time. The selection of a manufacturing process and operation is greatly influenced by the time required to start production. Depending on the die's shape complexity, size, and material, the lead time for such processes as forging, extrusion, die casting, roll forming, and sheet-metal forming can range from weeks to months. Lead time can be a critical factor in an increasingly competitive marketplace, as well as under adverse conditions such as emergencies and wartime.

By contrast to the processes listed in the previous paragraph, processes such as machining, grinding, and advanced material-removal processes have significant built-in flexibility, due to the fact that they utilize machinery and tooling that can readily be adapted to most production requirements in a very short time. 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 *rapid prototyping*, Chapter 20.)

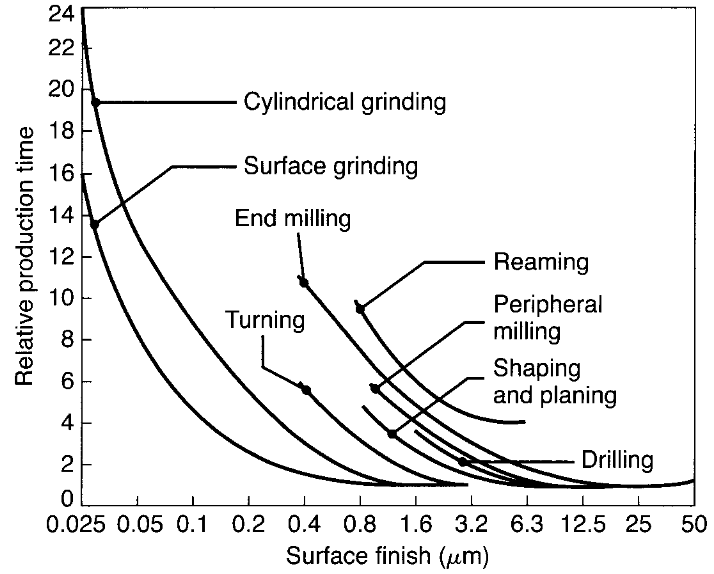


FIGURE 40.5 Relative production time as a function of surface finish produced by various manufacturing processes. (See also Fig. 26.35.)

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. In order to appreciate the importance of robustness in manufacturing processes, let's briefly consider a situation in which a simple plastic gear is being produced by injection molding (Section 19.3), but significant and unpredictable variations in quality arise as the gears are being produced. There are several well-understood variables and parameters in the injection molding of plastics, including the effects of raw-material (such as pellets) quality, speed, and temperatures within the system; all these are independent variables; hence, they can be controlled.

However, there are certain other variables, called *noise*, that are largely beyond the control of the operator. Among these are ambient-temperature and humidity variations in the plant throughout the day, dust in the air entering the plant from an open door (and thus possibly contaminating the pellets being fed into the hoppers of the injection-molding machine), and variability in the performance of individual operators during different shifts. Obviously, these variables are difficult or impossible to control precisely.

In order to obtain or sustain good product quality, it is necessary to understand the effects, if any, of each element of noise in the operation. 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 affect its behavior in the molding machine? (c) How different are the performances of different operators during different shifts, and why are they different? and (d) Are there inherent variations in machine performance during the day and, if so, how and why?

Such an investigation will make it possible to establish new operating parameters so that variations in, say, ambient temperature and the plant environment do not affect gear quality adversely. Note that these considerations are equally valid for any manufacturing operation, although some (such as bulk-deformation processes) are less sensitive to noise than others (especially microelectronics manufacturing).

40.8 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 the Workpiece Materials. Recall that some materials can be processed at room temperature, whereas others require elevated temperatures—and hence furnaces, appropriate tooling, and various controls. Some materials are easy to work with because they are soft and ductile. Other materials, such as those which are hard, brittle, and abrasive, require special processing technologies and equipment.

Materials have different manufacturing characteristics, such as castability, forgeability, workability, machinability, and weldability. Note from Table 40.5 that few materials have favorable characteristics in all of these relevant categories. For example, a material that is castable or forgeable may later present difficulties in subsequent processes, such as machining, grinding, and finishing, that may be required for an acceptable surface finish and dimensional accuracy.

Materials have different responses to the rate of deformation (strain-rate sensitivity, Sections 2.2.7 and 7.3) to which they are subjected. Thus, the speed at which a particular machine is operated can affect product quality, including the development of external and internal defects. Impact extrusion or drop forging, for example, may not be appropriate for materials with high strain-rate sensitivity, whereas such materials will perform well in a hydraulic press or in direct extrusion.

Geometric Features of the Part. Features such as part shape, size, and thickness, dimensional tolerances, and surface-finish requirements greatly influence the selection of a process or processes, as described throughout this chapter 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 40.7.)

Process Selection Considerations. The factors involved in process selection are summarized by the following questions:

1. Are some or all of the parts or components that are needed 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 with 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 final dimensions 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 properly?

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	A	A	A	B	A	A	A	A	B	A
Plaster	—	—	—	—	A	A	A	—	—	—
Ceramic	A	A	A	A	B	B	A	A	B	A
Investment	A	A	A	—	A	B	A	A	A	A
Permanent	B	B	—	—	A	A	A	—	—	—
Die	—	—	—	—	A	A	A	—	—	—
Forging										
Hot	A	A	A	A	A	A	A	A	A	A
Cold	A	A	A	—	A	B	A	—	—	—
Extrusion										
Hot	A	A	A	B	A	A	A	A	A	A
Cold	A	B	A	—	A	—	A	B	—	—
Impact	—	—	—	—	A	A	A	—	—	—
Rolling	A	A	A	—	A	A	A	A	A	B
Powder metals	A	A	A	A	A	A	A	A	A	A
Sheet-metal forming	A	A	A	—	A	A	A	A	A	A
Machining	A	A	A	—	A	A	A	B	A	B
Chemical	A	B	A	B	A	A	A	B	B	B
ECM	—	A	B	A	—	—	B	A	A	A
EDM	—	B	B	A	B	—	B	B	B	A
Grinding	A	A	A	A	A	A	A	A	A	A
Welding	A	A	A	—	A	A	A	A	A	A

Note: A = Generally processed by this method; B = Can be processed by this method, but may present some difficulties; — = Usually not processed by this method. Product quality and productivity depend greatly on the techniques and equipment used, operator skill, and proper control of processing variables.

EXAMPLE 40.6 Process Substitution in Making Common Products

The following list gives some typical choices that can be made in process selection for the products listed:

- **a.** Forged vs. cast crankshaft.
- **b.** Forged vs. powder-metallurgy connecting rod.
- **c.** Sheet metal vs. cast hubcap.
- **d.** Machining vs. precision forming of a large gear.
- **e.** Forging vs. powder-metallurgy production of a spur gear.
- **f.** Thread rolling vs. machining a threaded fastener.
- **g.** Casting vs. stamping a metal frying pan.
- **h.** Formed aluminum tubing vs. cast iron for outdoor furniture.
- **i.** Welding vs. mechanical fastening of machine-tool structures.

EXAMPLE 40.7 Process Selection in Making a Simple Part

You are asked to produce the simple axisymmetric part shown in Fig. 40.6a; 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, wear resistance, and resistance to elevated temperatures.

Which manufacturing process would you choose, and how would you organize the production facilities to manufacture a cost-competitive, high-quality product? 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 reduces the total manufacturing time and cost. Because it is relatively simple, this part can be manufactured by (a) casting or powder metallurgy, (b) forging, or upsetting, (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 (heading, Fig. 14.11) a 25-mm round bar in a suitable die to form the larger end. Another possibility is partial direct *extrusion* of a 38-mm diameter bar 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.

The part also can be made by *machining* a 38-mm-diameter bar stock to reduce the lower section to 25 mm. Machining this part will require much more time than forming it, and a considerable amount of material inevitably will be wasted as metal chips (Table 40.3). However, unlike net-shape processes, which generally require special dies, machining

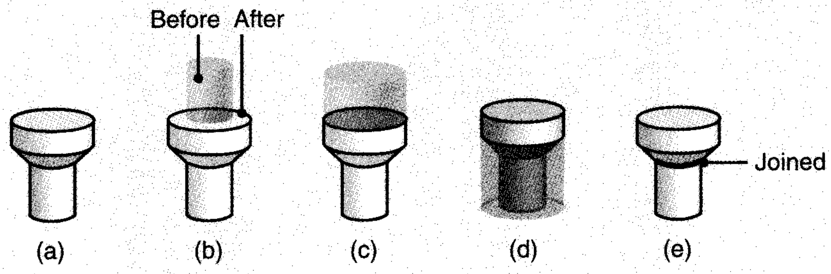


FIGURE 40.6 Various methods of making a simple part: (a) casting or powder metallurgy, (b) forging or upsetting, (c) extrusion, (d) machining, and (e) joining two pieces.

involves no special tooling, and this operation can be carried out easily on a CNC lathe at high rates. Note 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. For a high production quantity and rate, producing the part by a heading operation or by cold extrusion (a variation of closed-die forging, 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, the part can be made in two pieces, and joining them would be the most appropriate choice.

EXAMPLE 40.8 Manufacturing a Sheet-metal Part by Different Methods

A simple, dish-shaped part can be formed from sheet metal by placing a round, flat piece of sheet metal between a pair of male and female dies in a press and then closing the dies by applying a vertical force (Fig. 40.7a). Parts like this typically are formed in such manner at high production rates; the method is generally known as *stamping* or *pressworking*.

Assume now that the size of the part is very large, say, 2 m in diameter and that the lot size is only 50 parts. We now have to reexamine the total operation. Is it economical to manufacture a set of dies 2 m in diameter (which would be very costly; see Section 14.7) when the total production quantity is very low? Are presses available with sufficient capacity to accommodate such large dies? Are there alternative methods of manufacturing this part? Does the part have to be made in one piece?

This part also can be made by *welding* smaller pieces of sheet metal, formed by other methods, as described in Chapter 16. (Note that large municipal water tanks and ships are made by this method.) Would a part manufactured by welding be acceptable for its intended purpose in the environment in which it will be used? Will it have the required properties and the desired shape after welding, or will it require additional processing?

The part also can be made by *explosive forming*, as shown in Fig. 40.7b. Because of the nature of the process, the deformation of the material in explosive forming takes place at a very high rate. Consequently, a series of questions has to be asked regarding this process (Section 16.11):

- **a.** Is the material capable of undergoing deformation at high rates without fracture or any detrimental effect on the final properties of the formed part?
- **b.** Can the dimensional tolerances and surface finish be held within acceptable limits?
- **c.** Is the life of the die sufficiently long, given that the die is subjected to the very high transient pressures generated in explosive forming?
- **d.** Can this operation be performed in a manufacturing plant within city limits, or should it be carried out in open country?
- **e.** Although explosive forming has the advantage of requiring only one die, is the operation economical?

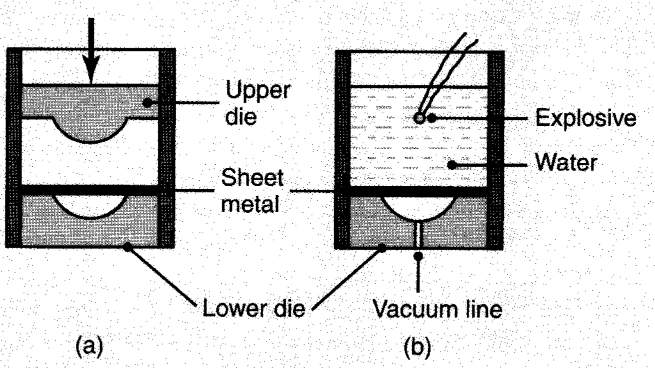


FIGURE 40.7 Two methods of making a dish-shaped sheet-metal part: (a) pressworking using a male and female die and (b) explosive forming using one die only.

40.9 Manufacturing Costs and Cost Reduction

The total cost of a product generally 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: Part II (casting); Part III (rolling, forging, extrusion, drawing, sheet-metal working, powder metallurgy, ceramics, polymer processing); Part IV (machining, abrasive processing, advanced machining); and Part VI (welding and various joining processes).

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 many technical and operational factors involved, calculating individual cost factors correctly 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) 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 now has a built-in added 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 20% of the price we pay for a ladder is attributed to potential product liability costs.

Materials Costs. Some cost data on materials are given in various tables throughout this book, as also listed in Table 40.1. Because of the different operations required in producing raw materials, their costs depend not only on the type of material (ferrous, nonferrous, nonmetallic, etc.), but also on its processing history (ingot, powder, drawn rod, extrusion), as well as its size, shape, and surface characteristics. For example, per unit weight, (a) drawn round bars are less expensive than bars that are ground to close tolerances and a fine surface finish, (b) square bars are more expensive than round bars, (c) cold-rolled plate is more expensive than hot-rolled plate, (d) thin wire is more expensive than thick wire, and (e) hot-rolled bars are much less expensive than metal powders of the same type.

Tooling Costs. Costs are involved in making the tools, dies, molds, patterns, and special jigs and fixtures required for manufacturing a product. Tooling costs can be very high, but they can be justified in high-volume production, such as automotive applications, where die costs can be on the order of $2 million. The expected life of tools and die, and their obsolescence because of product changes, also are important considerations.

Tooling costs are greatly influenced by the production process selected. For example, (a) the tooling cost of die casting is higher than that of sand casting; (b) the tooling cost of machining or grinding is much lower than that of powder metallurgy, forging, or extrusion; (c) carbide tools are more expensive than high-speed steel tools, but their life is longer; (d) if a part is to be manufactured by spinning, the tooling cost of conventional spinning is much lower than that of shear spinning; and (e) tooling for rubber-forming processes is less expensive than that of the die sets (male and female) used for the deep drawing and stamping of sheet metals.

Fixed Costs. These costs include electric power, fuel, taxes on real estate, rent, insurance, and capital (including depreciation and interest). The company has to meet fixed costs regardless of whether or not it has made a particular product; thus, fixed costs are not sensitive to production volume.

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

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.

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. 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 are for 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 by the worker 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. The cutting speeds for machining high-temperature alloys, for example, are lower than those for machining aluminum or plain-carbon steels. Consequently, the cost of machining aerospace materials is much higher than that of machining more common alloys, such as those of aluminum and steel.

Labor costs in manufacturing and assembly vary greatly from country to country (see Table I.4 in the General Introduction). It is not surprising that most of the products one purchases today are either made or assembled in countries where labor costs are low. On the other hand, firms located in countries with high labor rates tend to emphasize high value-added manufacturing tasks or high automation levels, so the labor component of the cost is significantly reduced.

For labor-intensive industries, such as machine building, 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 continues to increase because of important homeland security issues.
- Lengthy shipping times indicate that the benefits of just-in-time manufacturing approaches (Section 39.5) and their associated cost savings may not be realized. Also, 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 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, product defect rates can be significant.
- There are various other hidden costs, such as increased paperwork and documentation, lower productivity from existing employees because of lower morale, and difficulties in communication.

Indirect-labor Costs. These costs are generated in the servicing of the total manufacturing operation. They generally consist of the costs 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 finished parts, or they are not chargeable to a specific product, 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. Obviously, a large production quantity 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. As we have seen, 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. For example, an electric motor is made of relatively inexpensive materials, yet several different manufacturing operations are involved in the making of 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).

A typical breakdown of the costs in modern manufacturing is as follows:

Design	5%
Material	50%
Manufacturing	
Direct Labor	15%
Overhead	30%

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. In the foregoing breakdown, note the very small contribution of the *design phase*, yet the design phase generally has the largest influence on the *quality* and *success* of a product in the marketplace. The various opportunities for cost reduction have been discussed in a number of chapters throughout this book. Among these opportunities are the following:

- Simplifying both part design and the number of subassemblies required.
- Reducing the amount of materials used.
- Specifying broader dimensional tolerances and allowing rougher surface finish
- Using less expensive materials.
- Investigating alternative methods of manufacturing.
- Using more efficient machinery, systems, and equipment.

The introduction of more automated systems and the adoption of up-to-date technology in a manufacturing facility is an obvious means of reducing some costs. However, this approach must be undertaken with due care and only after a thorough **cost–benefit analysis**, which requires reliable input data and a consideration of the technical as well as the human factors involved. Advanced technologies, which can be very costly to implement, should be implemented only after a complete analysis of the more obvious cost factors, known as **return on investment** (ROI).

40.9.1 Value Analysis

Manufacturing adds *value* to materials as they become discrete products and are marketed. Because this 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 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.

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

$$\text{Value} = \frac{\text{Product function and performance}}{\text{Product cost}}. \qquad (40.1)$$

Thus, the goal of value analysis is to obtain maximum performance per unit cost. Value analysis generally consists of the following six phases:

1. *Information phase*: to gather data and determine costs.
2. *Analysis phase*: to define functions and identify problems as well as opportunities.
3. *Creativity phase*: to seek ideas in order to respond to problems and opportunities without judging the value of each idea.
4. *Evaluation phase*: to select the ideas to be developed and identify the costs involved.
5. *Implementation phase*: to present facts, costs, and values to the company management; to develop 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*: to reexamine the overall value-analysis process in order to make necessary adjustments.

Value analysis is an important and all-encompassing interdisciplinary activity, usually coordinated by a value engineer and conducted jointly by designers, manufacturing engineers, and quality-control, purchasing, and marketing personnel and managers. In order for value analysis to be effective, 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.

An example of product weight reduction is the development of the antilock braking system (ABS) for automotive applications. In 1989, the typical weight of a Bosch brand system was 6.2 kg. In 2001, its weight was 1.8 kg, a reduction of 70%, which also helped reduce the weight of the automobile. Note that, considering the function of the product and the fact that weight is related to the product's volume, reducing the size indicates that the ratio of surface area to volume increases.

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. Several guidelines have been established for designing products for economic production.
- 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 our environment and the ecosystem.
- Substitution of materials, modification of product design, and relaxing of dimensional tolerance and surface finish requirements are 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, but to counteract lower wages in developing countries, labor costs can be reduced further through highly automated and computer-controlled manufacturing 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
Return on investment
Scrap
Smart databases
Sustainable manufacturing
Value
Value analysis

BIBLIOGRAPHY

Anderson, D.M., **Design for Manufacturability & Concurrent Engineering**, CIM Press, 2003.

Ashby, M.F., **Materials Selection in Mechanical Design**, 3rd ed., Pergamon, 2005.

ASM Handbook, Vol. 20: **Materials Selection and Design**, ASM International, 1997.

Billatos, S., and Basaly, N., **Green Technology and Design for the Environment**, Taylor & Francis, 1997.

Boothroyd, G., Dewhurst, P., and Knight, W., **Product Design for Manufacture and Assembly**, 2nd ed., Dekker, 2001.

Bralla, J.G., **Design for Manufacturability Handbook**, 2nd ed., McGraw-Hill, 1999.

Cha, J., Jardim-Gonçalves, R., and Steiger-Garção, A., **Concurrent Engineering**, Taylor & Francis, 2003.

Dettmer, W.H., **Breaking the Constraints to World-Class Performance**, ASQ Quality Press, 1998.

Dorrf, R.C., and Kusiak, A. (eds.), **Handbook of Design, Manufacturing and Automation**, Wiley, 1995.

Giudice, F., La Rosa, G., and Risitano, A., **Product Design for the Environment**, CRC Press, 2006.

Harper, C.A. (ed.), **Handbook of Materials for Product Design**, McGraw-Hill, 2001.

Hartley, J.R., and Okamoto, S., **Concurrent Engineering: Shortening Lead Times, Raising Quality, and Lowering Costs**, Productivity Press, 1998.

Hundai, M. (ed.), **Mechanical Life Cycle Handbook**, CRC Press, 2001.

Madu, C. (ed.), **Handbook of Environmentally Conscious Manufacturing**, Springer, 2001.

Magrab, E.B., **Integrated Product and Process Design and Development: The Product Realization Process**, CRC Press, 1997.

Mangonon, P.C., **The Principles of Materials Selection for Design**, Prentice Hall, 1999.

McDonough, W., and Braungart, M., **Cradle to Cradle: Rethinking the Way We Make Things**, North Point Press, 2002.

Poli, C., **Design for Manufacturing: A Structured Approach**, Butterworth-Heinemann, 2001.

Shina, S.G. (ed.), **Successful Implementation of Concurrent Engineering Products and Processes**, Wiley, 1997.

Stoll, H.W., **Product Design Methods and Practices**, Dekker, 1999.

Swift, K.G., and Booker, J.D., **Process Selection: From Design to Manufacture**, 2nd ed., Butterworth-Heinemann, 2003.

Wenzel, H., and Hauschild, M., **Environmental Assessment of Products**, Vol. 2, Springer, 1997.

Wenzel, H., Hauschild, M., and Alting, L., **Environmental Assessment of Products**, Vol. 1, 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. Explain the difference between direct-labor cost and indirect-labor cost.

40.7. Describe your understanding of the following terms: (a) life expectancy, (b) life-cycle engineering, (c) sustainable manufacturing, and (d) green manufacturing.

40.8. Is there a significant difference bewteen cradle-to-grave and cradle-to-cradle production? Explain.

40.9. How would you define value? Explain.

40.10. What is the meaning and significance of the term "return on investment"? Explain.

QUALITATIVE PROBLEMS

40.11. Describe the major considerations involved in selecting materials for products.

40.12. What is meant by manufacturing process capabilities? Select four different manufacturing processes and describe their capabilities.

40.13. Comment on the magnitude and range of scrap shown in Table 40.3 and the reasons for the variations.

40.14. Explain why the value of the scrap produced in a manufacturing process depends on the type of material and processes involved.

40.15. Describe your observations concerning the information given in Table 6.1 and the reasons for those observations.

40.16. 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.

40.17. Explain how the high cost of some of the machinery listed in Table 40.6 can be justified.

40.18. On the basis of the topics covered in this book, explain the reasons for the relative positions of the curves shown in Fig. 40.2.

40.19. What factors are involved in the shape of the curve shown in Fig. 40.4? Explain.

40.20. Describe the problems that may have to be faced in reducing the quantity of materials in products. Give some examples.

40.21. Explain the reasons that there is a strong desire in industry to practice near-net-shape manufacturing.

40.22. State and explain your thoughts concerning cradle-to-cradle manufacturing.

40.23. List and explain the advantages and disadvantages of outsourcing manufacturing activities to countries with low labor costs.

SYNTHESIS, DESIGN, AND PROJECTS

40.24. 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.

40.25. 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.

40.26. What course of action would you take if the supply of a raw material selected for a product line becomes unreliable? Explain.

40.27. Estimate the position of the curves for the following processes in Fig. 40.5: (a) centerless grinding, (b) electrochemical machining, (c) chemical milling, and (d) extrusion.

40.28. 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?

40.29. 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.

40.30. 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*).

40.31. 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.

40.32. Describe your own thoughts concerning the replacement of aluminum beverage cans with cans made of steel.

40.33. 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.

40.34. 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.

40.35. 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.

40.36. 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.

40.37. What products do you know of that would be very difficult to disassemble for recycling purposes? Explain.

Index

LIST OF TABLES

General Introduction

Part I Fundamentals of Materials: Behavior and Manufacturing Properties

Part II Metal-Casting Processes and Equipment

Part III Forming and Shaping Processes and Equipment

REFERENCES TO VARIOUS TOPICS
(Page Numbers are in Parentheses)

Material Properties

Tables 2.1 (57), 2.2 (59), 2.3 (62), and Figs. 2.4, 2.6, 2.7, 2.8, 2.15, 2.16, 2.17, 2.29
Tables 3.1 (89), 3.2 (90), and Figs. 3.1, 3.2, 3.3
Tables 5.2 (139), 5.4 (141), and 5.5 (144)
Tables 6.3 through 6.10 (153–162)
Tables 7.1 (172), 7.2 (180), 7.3 (186)
Tables 8.1 (199), 8.2 (202), 8.3 (206)
Tables 9.1 (218), 9.2 (220), 9.3 (228), and Figs. 9.3, 9.5, 9.7
Table 10.1 (248)
Table 11.3 (281)
Tables 12.3 (304), 12.4 (305), 12.5 (305), and Fig. 12.4
Tables 16.2 (392), 16.3 (398), 16.4 (409), and Fig. 16.14
Tables 17.3 (455), 17.4 (455), 17.5 (456), and Fig. 17.10
Table 20.2 (529)
Tables 22.1 (593), 22.2 (594), 22.3 (594), 22.5 (600), and Figs. 22.1, 22.9
Table 26.1 (721)
Table 32.3 (931)
Section 17.7 and Fig. 17.14
Table 18.1 (466)
Tables 19.1 (485), 19.2 (521)
Table 20.1 (528)
Tables 23.1 (617), 23.6 (627), 23.8 (635), 23.10 (644)
Tables 26.3 (735), 26.4 (736)
Table 27.1 (761)
Tables 28.1 (801), 28.2 (809), 28.3 (810), and Fig. 28.19
Tables 29.1 (847), 29.3 (856)
Table VI.1 (864)
Table 30.1 (866)
Table 32.4 (932)
Table 34.1 (989)
Table 37.2 (1056), and Fig. 37.3
Table 39.1 (1121)
Tables 40.3 (1146), 40.5 (1153), and Figs. 40.2, 40.3, 40.5

Manufacturing Characteristics of Materials

Table I.3 (16)
Table 4.1 (120)
Table 5.8 (147)
Table 6.2 (152)
Tables 12.1 (297), 12.6 (308)
Table 14.3 (348)
Table 16.3 (398), and Fig. 16.33
Tables 17.1 (447), 17.2 (453)
Tables 21.1 (559), 21.2 (571)
Fig. 22.2

Dimensional Tolerances and Surface Finish

Table 11.2 (261)
Table 23.1 (617), and Figs. 23.13, 23.14
Fig. 25.16
Fig. 27.4
Figs. 35.19, 35.20
Figs. 40.4, 40.5

Capabilities of Manufacturing Processes

Tables 11.1 (259), 11.2 (261)
Table III.1 (315)
Table 14.1 (337), 14.4 (353)
Table 16.1 (383)

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 16.13

General Costs and Economics

Tables I.5 (30), I.6 (33), I.7 (34), and Section I.10
Table 6.1 (152)
Section 12.4
Section 14.9
Section 16.15
Table 17.6 (461) and Section 17.8
Table 19.2 (521) and Section 19.16
Section 25.8
Section 26.9 and Fig. 26.35
Section 27.11
Section 31.8
Section 32.7
Section 37.11
Section 39.9
Table 40.6 (1157) and Section 40.9

CONVERSION FACTORS

1 in. = 25.4 mm = 0.0254 m
1 mm = 0.0394 in.
1 ft = 0.3048 m
1 lb (force) = 4.448 N
1 lb (mass) = 0.4536 kg
1 ton = 2240 lb

1 tonne = 1000 kg
1 psi = 6.895 kPa
1 ksi = 6.895 MPa
1 MPa = 145 psi
1 ft·lb = 1.356 J
1 BTU = 1055 J = 778 ft·lb

1 hp = 746 W = 550 ft·lb/s
1 kW = 1.34 hp = 3413 BTU/hr
°F = 9/5 °C + 32
°C = 5/9 (°F − 32)
K = °C + 273.15

LIST OF EXAMPLES