UNITED STATES PATENT AND TRADEMARK OFFICE

BEFORE THE PATENT TRIAL AND APPEAL BOARD

R.J. REYNOLDS VAPOR COMPANY,

Petitioner

v.

FONTEM HOLDINGS 1 B.V.,

Patent Owner

Case IPR2016-01268
Patent 8,365,742

SUPPLEMENTAL EVIDENCE DECLARATION OF DR. ROBERT H. STURGES

1. I have been retained by the law firm of Brinks Gilson & Lione on behalf of R.J. Reynolds Vapor Company ("Petitioner") in connection with IPR2016-01268. I previously provided three declarations ("Petition Declaration," Ex. 1015; "Supplemental Declaration," Ex. 1020; "Reply Declaration," Ex. 1027) concerning the technical subject matter relevant to the petition in IPR2016-01268.

2. My background and qualification are contained in my Petition Declaration.

3. My list of prior testimony and updated lists are attached to my Petition and Reply declarations.

4. The information I considered is identified in my prior declarations and deposition testimony.

5. I submit this supplemental evidence declaration in response to Patent Owner's Objections To Petitioner's Evidence Under 37 C.F.R. § 42.64(b)(1) dated July 12, 2017 (Paper 32). Specifically, Patent Owner provided the following objection to my Reply Declaration (Ex. 1027):

> Paragraphs 33-67 offer opinions regarding airflow, aerosols, pore sizes, "standard drawing practices," "aerodynamic forces," compression, tensile strength, "providing the needed holes," electrical resistance, atomization, a "slipstream," thermal efficiency, and "heating wire in lightbulbs, heaters and other things" that are not

> based on sufficient facts or data and are not the product of reliable principles and methods.

While no explanation is provided as to why the opinions in paragraphs 33-67 "are not based on sufficient facts or data and are not the product of reliable principles and methods" and no specific examples are provided, I nonetheless have attempted to provide some additional "facts and data" for my opinions.

6. Patent Owner identifies the term "standard drawing practices" in its objection. "Standard drawing practices" as used in paragraph 35 of my reply declaration refers to common techniques used in mechanical section drawings that require enough views to ensure that the internal features of the object depicted are shown. Ex. A (Engineering Design Graphics), p. 235. Typically, two or more sectional drawings are necessary to adequately show an object's internal features. *See Id.* Drawings that do not make clear what is being shown would normally be rejected and not used by designers and engineers. With respect to Patent Owner's argument that an exit hole purportedly exists in the bulge section of Hon's porous body 27 as illustrated in Figs. 6 and 8, as I mentioned in my reply declaration, no such hole is described or illustrated in Hon 043. If an exit hole existed, the person having ordinary skill in the art ("PHOSTA") would have expected that, consistent with standard drawing practices, the inventor would have clearly illustrated such a

hole using sectional views and/or perspective drawings. However, there are no such views or drawings in Hon 043 illustrating the purported exit hole.

7. Patent Owner identifies the term "aerodynamic forces" in its objection. I used the term "aerodynamic forces" in paragraph 43 of my reply declaration when explaining that the PHOSITA would have understood that air-entrained atomized liquid droplets (*i.e.*, aerosol) are pulled through the downstream portion of Hon 043's porous body 27 via aerodynamic forces caused by the user drawing on the device but that those forces were insufficient to pull unatomized liquid out of the porous body at the downstream end of Hon 043's porous body 27. Thus, as Hon 043 explains:

> After the atomization, the large diameter droplets stick to the wall under the action of eddy flow and are reabsorbed by the porous body 27 via the overflow hole 29, whereas the small diameter droplets float in stream and forms aerosols, which are sucked out via the aerosol passage 12, gas vent 17 and mouthpiece 15.

Ex. 1003-00011. As the PHOSITA would have understood, the "aerodynamic forces" that influence the manner in which the "small diameter droplets" forming an aerosol travel through Hon 043's porous body are created by pressure differences that result from the user drawing on the Hon 043 device. The aerosol will flow towards lower pressure areas in accordance with Bernoulli's law. As the

PHOSITA would have understood, the aerosol pulled through the pores of Hon 043's porous body when the user draws upon the device will have a velocity distribution profile where the airflow through the center of the pores is faster than at the edges of the pores. Ex. B (DOE Fundamentals Handbook, Thermodynamics, Heat Transfer, And Fluid Flow, Volume 3 of 3), p. 18 ("The velocity of the fluid in contact with the pipe wall is essentially zero and increases the further away from the wall."). Although the DOE Handbook refers to fluid flow through a pipe, the PHOSITA would have expected that the same general principles apply to fluid flow through the pores of a porous body. The DOE Handbook identifies that the term "fluid" includes both liquids and gases. Ex. B, p. 2. As can be seen in Fig. 5 from the DOE Handbook (copied below), similar velocity profiles exist both for laminar flow and for turbulent flow. *Id.*

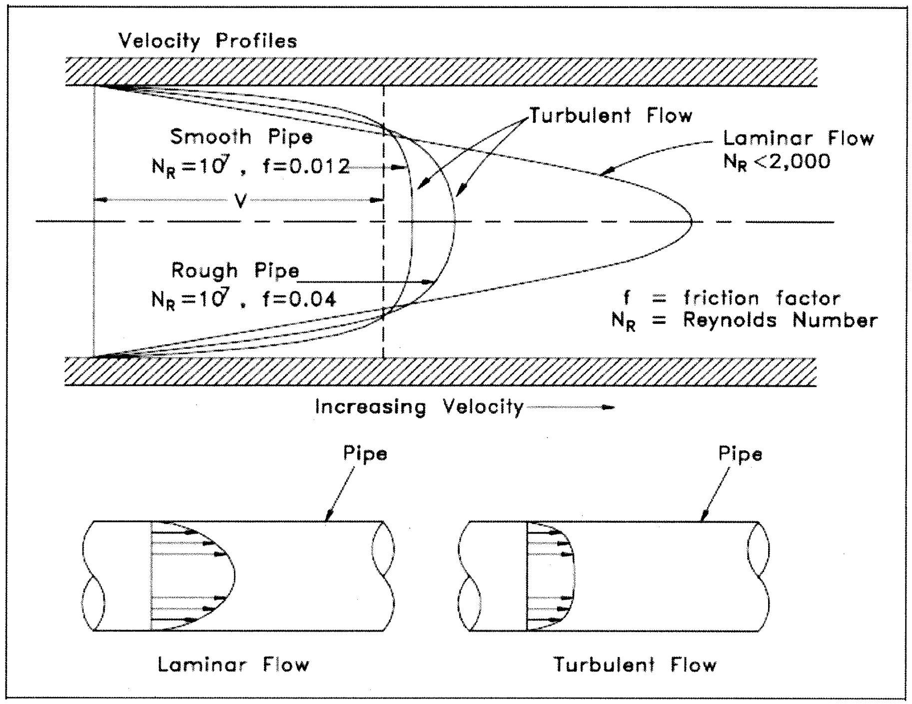


Figure 5 Laminar and Turbulent Flow Velocity Profiles

8. Thus, the PHOSITA would have understood that the velocity of the aerosol pulled through Hon 043's porous body when the user draws on the device is higher in the center of Hon 043's pores than at the edges, and that the small liquid droplets entrained in Hon 043's aerosol stream, particularly those located in the center of the air stream, would pass through the pores of Hon 043's porous body without being "picked off" or otherwise reabsorbed by the porous body. In other words, as the PHOSITA would have understood, because Hon 043's porous body is permeable to airflow, then it is also permeable to aerosol flow.

9. However, although the aerodynamic forces of the user drawing on Hon 043's device are sufficient to pull aerosol from Hon 043's porous body, the

PHOSITA would have understood that these forces would not be sufficient to pull unatomized liquid out of Hon 043's porous body at the downstream end. As the PHOSITA would have understood, as liquid is spent at the upstream end of Hon 043's porous body, unatomized liquid would move by capillary action from the downstream end towards the upstream end of the porous body. The PHOSITA would have expected that the aerodynamic forces associated with the user drawing on the Hon 043 device would not be sufficient to overcome the capillary action that moves the unatomized liquid from the downstream end to the upstream end of Hon 043's porous body.

10. Also, the PHOSITA would have understood that the pressure in Hon 043's ejection holes 24 is much lower than that in the porous body in the immediate vicinity of the ejection holes, because the air stream moves faster in the ejection holes than in the bulk porous body. Under Bernoulli's law, an increase in flow velocity will result in a decrease in pressure. Ex. B, p. 23. As a result, the PHOSTA would have also understood that Hon 043 is designed such that there is a significant pressure differential between the porous body in the area of ejection holes 24 and the inside of ejection holes 24, which causes unatomized liquid droplets to enter the ejection holes 24 as described in Hon 043. In contrast, the pressure differential between the downstream end and the area immediately outside of the downstream end of the porous body is not as great. The PHOSITA would

have understood that this pressure difference is not sufficient to overcome the forces associated with the general movement of unatomized liquid from the downstream end to the upstream end via capillary action, which explains why the PHOSITA would have understood that unatomized liquid droplets exit the porous body into ejection holes 24 but would not exit the porous body at the downstream end of the porous body.

11. Patent Owner refers to the words "compression" and "tensile strength" in its objection but does not identify where those words are used. Both those terms are used in paragraph 44 of my reply declaration, where I explain why the stress-strain curves and the purported tensile strengths of certain materials on which Mr. Meyst relies is not relevant to the extent to which Hon 043's porous body would bend or sag but for the presence of cavity wall 27. The term compression is used in conjunction with my description of Fig. 6a of Ex. 2019 and Fig. 5 from Ex. 2018, which are both compressive strain curves. Ex. 2019 at 4 (p. 187), Ex. 2018 at 6 (p. 821). As I mentioned in my reply declaration, a compression strain curve measures a material's ability to withstand compression. Compression tests are conducted by applying a load at both ends of a specimen as shown in Fig. 2.1(b) of Manufacturing Processes for Engineering Materials. Ex. C (Manufacturing Processes for Engineering Materials), p. 27. Tensile strength is the opposite of compression. It measures a material's ability to withstand stretching.

Tensile strength tests are conducted by applying a "pulling" force on both ends of the sample as shown in Fig. 2.1(a), which is copied below. Ex. C, p. 27.

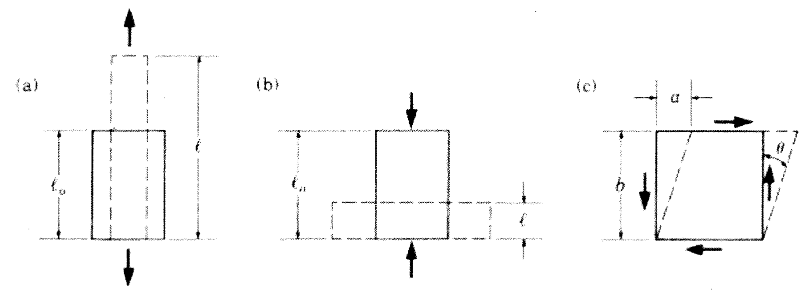
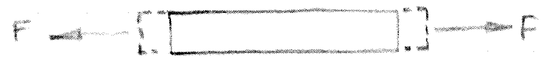


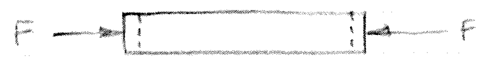
FIGURE 2.1 Types of strain. (a) Tensile. (b) Compressive. (c) Shear. All deformation processes in manufacturing involve strains of these types. Tensile strains are involved in stretching sheet metal to make car bodies, compressive strains in forging metals to make turbine disks, and shear strains in making holes by punching.

Ex. C, p. 27

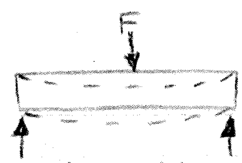
Neither compression nor tensile strength are a measurement of a material's ability to withstand bending or sagging. I note that Mr. Meyst agrees. Ex. D, 61:4-63:1. Bending or sagging is the type of deformation that the porous body in Hon 043 would exhibit when a user coughs or intentionally blows into the device (or when the porous body is loaded with liquid) but for the support provided by the cavity wall. Below I have shown arrows representing the forces used for material compression testing ("compressive stress strain curves"), material tensile strength testing and material bending testing.



TENSION



COMPRESSION



BENDING

12. Patent Owner refers to "electrical resistance" in its objection. In paragraph 56 of my reply declaration, I discuss that resistance of a wire is directly proportional to the length of the wire and inversely proportional to the cross-sectional area of the wire. I further explain that a PHOSITA would have known that if the diameter of a wire is doubled, the cross sectional area grows by four times and the resistance drops by four times. This relationship between wire diameter and resistance is well-known. Ex. E (Marks' Standard Handbook for Mechanical Engineers), p. 15-4 ($R = \rho l/a$); *see* http://www.rapidtables.com/calc/wire/wire-gauge-chart.htm. For example, as can be calculated on the website, 20 gauge wire has a diameter of 0.032 inches and a resistance of 10.182 kΩ/1000 ft. By comparison, 14 gauge wires has a diameter of 0.0641 inches (twice the diameter of 20 gauge) and a resistance of 2.5194 kΩ

/1000 ft (about four times less than 20 gauge). Marks Handbook provides nearly identical data. Ex. E, p. 15-6. A PHOSITA also would have understood that if the voltage and current are the same for two wires, these two wires will have the same resistance according to Ohm's law (resistance equals voltage divided by current, R = E/I, where R is resistance, E is voltage, and I is current). Ex. E, p. 15-7. For two wires with the same length – *e.g.*, two wires having the length of the wire illustrated in Fig. 6 of Hon 043's device – a thicker wire (*i.e.*, a wire with a larger diameter) will have lower resistance than the thinner or smaller diameter wire. Under the same voltage, the currents in the two wires will not be the same – the current in the thicker wire will be greater than in the thinner wire to obtain the same voltage (*i.e.*, E = IR, where E is voltage, I is current and R is resistance). Thus, to generate the same amount of energy per puff (*i.e.*, J = IET, where J is energy, I is current, E is voltage, and T=time)(Ex. E, p. 15-3, Table 15.1.1), the thicker wire (*i.e.*, having the higher current, I) would be on (*i.e.*, conducting current) for a shorter amount of time (T) than the thinner wire (having lower current, I). This means that the thicker wire uses more current per unit time than the thinner wire. This higher use of current per unit time means that the discharge rate of the battery will be higher for a thick wire than a thin wire. In other words, the higher discharge rate in the thicker wire means that the battery will be draining faster than if a thin wire were used. A PHOSITA would have known that higher discharge rates may result in a significant reduction of available battery capacity, thus decreased efficiency and shortened life of a battery. Ex. E, p. 15-16 ("The ampere-hour capacity of batteries falls off rapidly with increase in discharge rate."). Therefore, as I explained in paragraphs 56 and 57 of my reply declaration, a PHOSITA would have understood that increasing the diameter of Hon's heating wire 26 would lower the resistance of the wire, resulting in faster discharge of the current in a shorter time in order to obtain the same heating power, and thus would likely decrease, not increase the heating efficiency of the Hon device. In other words, a user will not get as many puffs out of a Hon device with a thicker wire compared to a Hon device with a thinner wire.

13. Ex. A is a compilation of pages from the book James H. Earle, Engineering Design Graphics, AutoCAD® Release 12, 8th Ed., 1994 ("Earle") from my personal library. Ex. A includes the cover page, title page, copyright page and pages 235-254 of the Earle book.

14. I have examined Ex. A and it is a true and correct copy of the pages from the Earle book. The Earle book is available for inspection and copying at the offices of Brinks Gilson & Lione, Suite 3600 NBC Tower, 455 Cityfront Plaza Drive, Chicago IL 60611-5599.

15. On July 17, 2017, I downloaded a copy of the book DOE Fundamentals Handbook, Thermodynamics, Heat Transfer, And Fluid Flow,

Volume 3 of 3, 1992 from the Department of Energy website: https://energy.gov/sites/prod/files/2013/06/f2/h1012v3_0.pdf. I have used this website in the past to access information.

16. Ex. B is a true and correct copy of the book DOE Fundamentals Handbook, Thermodynamics, Heat Transfer, And Fluid Flow, Volume 3 of 3, 1992 that I obtained from the Department of Energy website.

17. Ex. C is a compilation of pages from Serope Kalpakjian, Manufacturing Processes for Engineering Materials, 1984 ("Kalpakjian") from my personal library. Ex. C includes the cover page, title page, copyright page and pages 25-96 of the Kalpakjian book.

18. I have examined Ex. C and it is a true and correct copy of the pages from the Kalpakjian book. The Kalpakjian book is available for inspection and copying at the offices of Brinks Gilson & Lione, Suite 3600 NBC Tower, 455 Cityfront Plaza Drive, Chicago IL 60611-5599.

19. Ex. E is a compilation of pages from Eugene A. Avallone and Theodore Baumeister III, Marks' Standard Handbook for Mechanical Engineers, 9th Ed., 1978, ("Marks' Handbook") from my personal library. Ex. E includes the cover page, title page, copyright page and pages 15-1 to 15-100 of the Marks' Handbook.

20. I have examined Ex. E and it is a true and correct copy of the pages from the Marks' Handbook. The Marks' Handbook is available for inspection and copying at the offices of Brinks Gilson & Lione, Suite 3600 NBC Tower, 455 Cityfront Plaza Drive, Chicago IL 60611-5599.

21. I declare under penalty of perjury under the laws of the United States that the foregoing is true and correct.

July 26, 2017

Robert H. Sturges

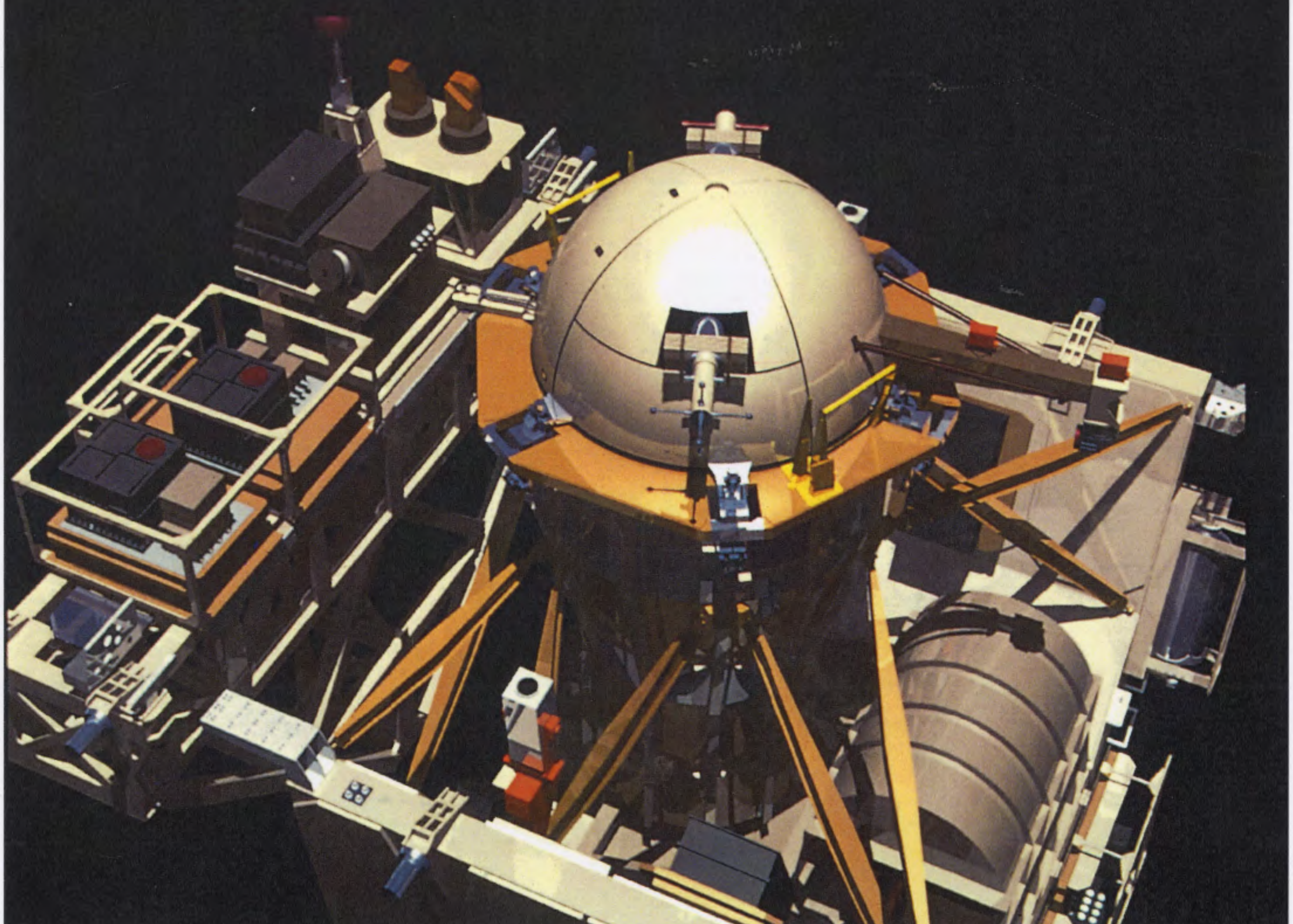
Dr. Robert H. Sturges

Exhibit A

ENGINEERING DESIGN GRAPHICS

AutoCAD® Release 12 • Eighth Edition

James H. Earle



ENGINEERING DESIGN GRAPHICS

AutoCAD® Release 12
Eighth Edition

JAMES H. EARLE

Texas A & M University

Addison-Wesley Publishing Company
Reading, Massachusetts • Menlo Park, California • New York
Don Mills, Ontario • Wokingham, England • Amsterdam • Bonn
Sydney • Singapore • Tokyo • Madrid • San Juan • Milan • Paris

Executive Editor: Michael Payne
Sponsoring Editor: Denise Descoteaux
Senior Production Supervisor: David Dwyer
Senior Production Coordinator: Genevra A. Hanke
Cover Design Supervisor: Peter M. Blaiwas
Senior Manufacturing Manager: Roy Logan
Manufacturing Coordinator: Judy Sullivan
Text Designer: Jean Hammond
Copyeditor: Jerrold C. Moore
Proofreader: Phyllis Coyne
Layout Artist: Julia M. Fair

Cover credits: Courtesy of Trilby Wallace, McDonnell Douglas Space Systems, Kennedy Space Center, Florida, and Intergraph Corporation, Huntsville, Alabama.

Library of Congress Cataloging-in-Publication Data

Earle, James H.
Engineering design graphics : AutoCAD release 12 / James H. Earle.
-- 8th ed.
p. cm.
Includes index.
ISBN 0-201-51982-8 (alk. paper)
1. Engineering design. 2. Engineering graphics. I. Title.
TA174.E23 1994
620'.0042'028566869--dc20 93-15665
CIP

1 2 3 4 5 6 7 8 9 10 – DOC – 9796959493

16

Sections

16.1 Introduction

Correctly drawn orthographic views that show all hidden lines may not clearly describe an object's internal details. This shortcoming can be overcome by imagining that part of the object has been cut away and shown in a cross-sectional view, called a **section**.

16.2 Basics of Sectioning

Figure 16.1 shows pictorially a section created by passing an imaginary cutting plane through the object to reveal its internal features. Think of the cutting plane as a knife-edge cutting through the object. **Figure 16.1A** shows the standard top and front views, and **Fig. 16.1B** shows the method of drawing a section. The front view is full section, with the portion cut by the imaginary plane cross-hatched. Hidden lines usually are omitted because they are not needed.

Figure 16.2 shows two types of cutting planes. Either is acceptable although the one with pairs of

SECTIONS VS. VIEWS

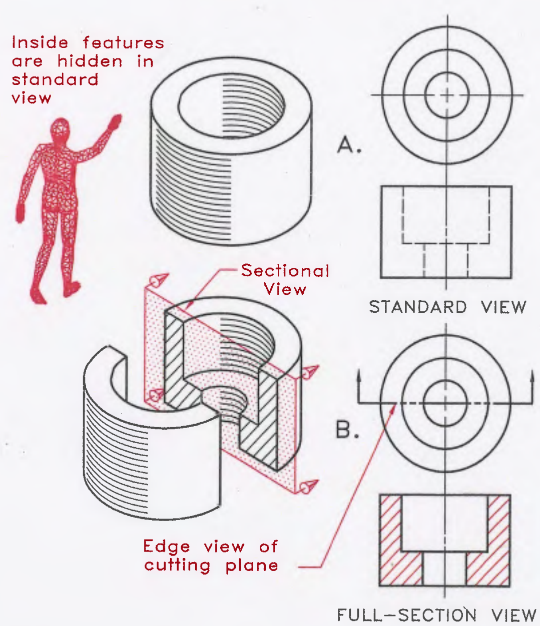


Figure 16.1 This drawing compares a standard orthographic view with a full-section view that shows the internal features of the same object.

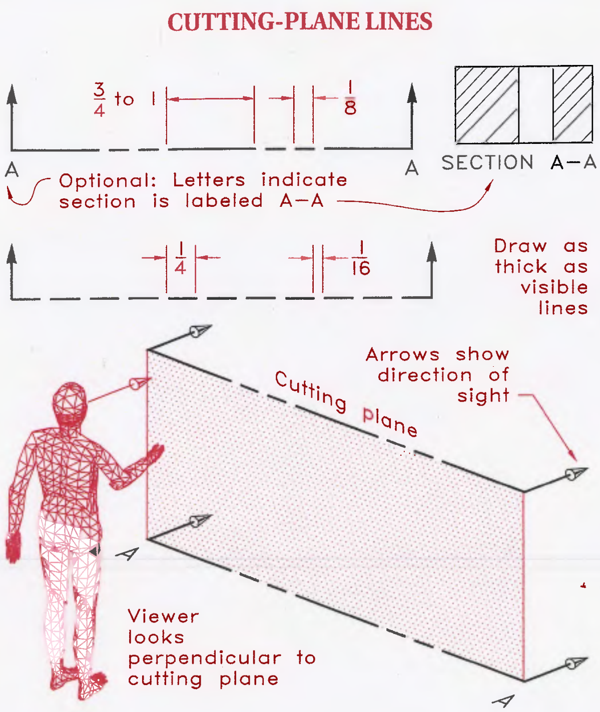


Figure 16.2 Use cutting-plane lines to represent sections (the cutting edge). The cutting plane marked A–A produces a section labeled A–A.

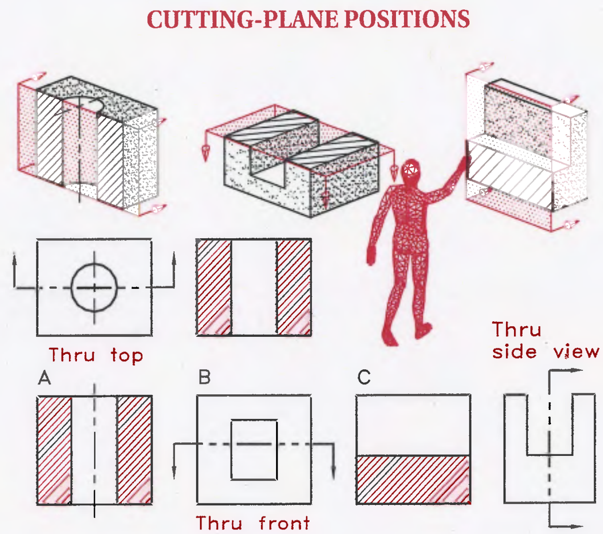


Figure 16.3 The three standard positions of cutting planes through orthographic (A) top, (B) front, and (C) side sectional views as sections. The arrows point in the direction of your line of sight for each section.

short dashes is most often used. The spacing and proportions of the dashes depend on the size of the drawing. The line thickness of the cutting plane is the same as the visible object line. Letters placed at each end of the cutting plane are used to label the sectional view, such as section A–A.

The sight arrows at the ends of the cutting plane are always perpendicular to the cutting plane. In the sectional view, the observer is looking in the direction of the sight arrows, perpendicular to the surface of the cutting plane.

Figure 16.3 shows the three basic positions of sections and their respective cutting planes. In each case perpendicular arrows point in the direction of the line of sight. For example, the cutting plane in **Fig. 16.3A** passes through and removes the front of the top view and the line of sight is perpendicular to the remainder of the top view.

The top view appears as a section when the cutting plane passes through the front view and the line of sight is downward **(Fig. 16.3B)**. When the cutting plane passes vertically through the side view **(Fig. 16.3C)**, the front view becomes a section.

16.3 Sectioning Symbols

Figure 16.4 shows the hatching symbols used to distinguish between different materials in sections. Although these symbols may be used to indicate the materials in a section, you should provide supplementary notes specifying the materials to ensure clarity.

The cast-iron symbol (evenly spaced section lines) may be used to represent any material and is the symbol used most often. Draw cast-iron symbols with a 2H pencil, slant the lines upward at 30°, 45°, or 60° angles, and space the lines about 1/16 inch apart (close together in small areas and farther apart in larger areas).

STANDARD HATCHING SYMBOLS

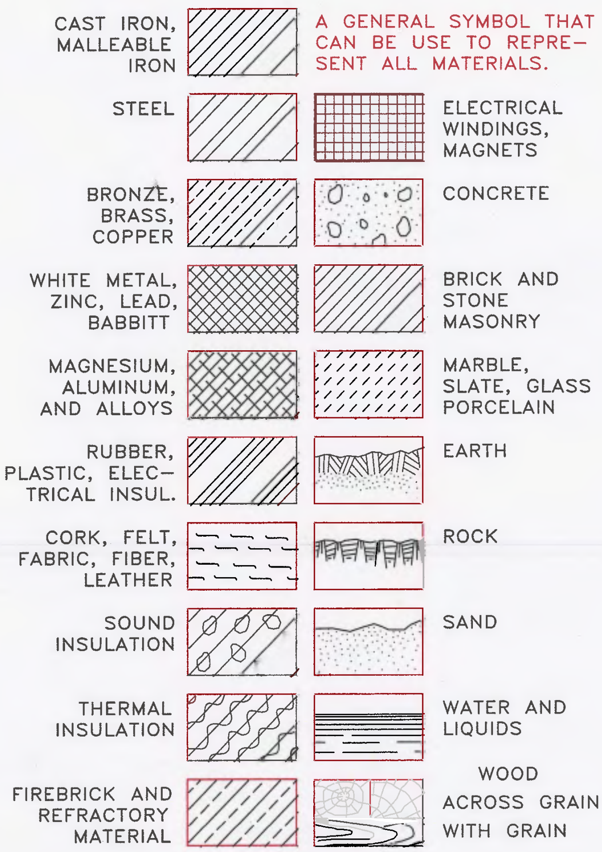


Figure 16.4 Use these symbols for hatching parts in section. The cast-iron symbol may be used for any material.

AUTOCAD HATCHING SYMBOLS

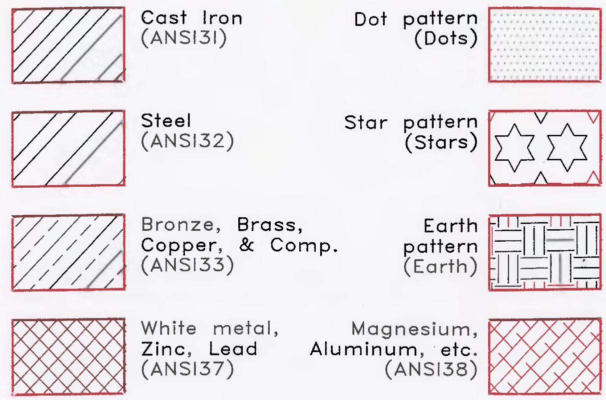


Figure 16.5 These are a few of the hatching symbols provided by AutoCAD. Use the pattern scale factor to vary the spacing between lines and dashes.

SECTION-LINE SPACING

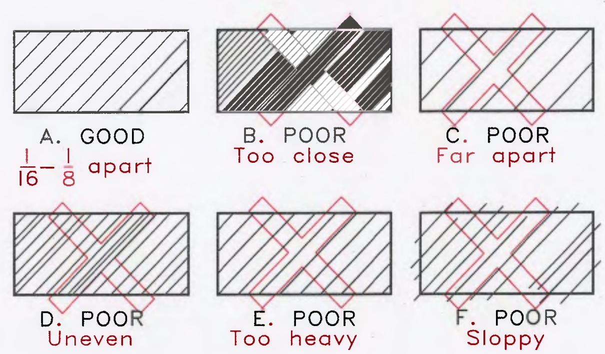


Figure 16.6 Techniques:

A Section lines are thin lines drawn 1/16 to 1/8 in. apart.

B–F Avoid these typical section lining errors.

Computer Method **Figure 16.5** shows a few of the many cross-sectional symbols available with AutoCAD. You may vary the spacing between the lines and the dash lengths by changing the pattern scale factor.

Figure 16.6A shows properly drawn section lines: thin and evenly spaced. **Figure 16.6B–F** show common errors of section lining.

Section thin parts such as sheet metal, washers, and gaskets by completely blacking in the areas **(Fig. 16.7)**, because space does not permit the drawing of section lines. Show large parts with an outline section to save time and effort.

You should hatch sectioned areas with symbols that are neither parallel nor perpendicular to the outlines of the parts lest they be confused with serrations or other machining treatments of the surface. **(Fig. 16.8)**.

LARGE AND THIN PARTS IN SECTION

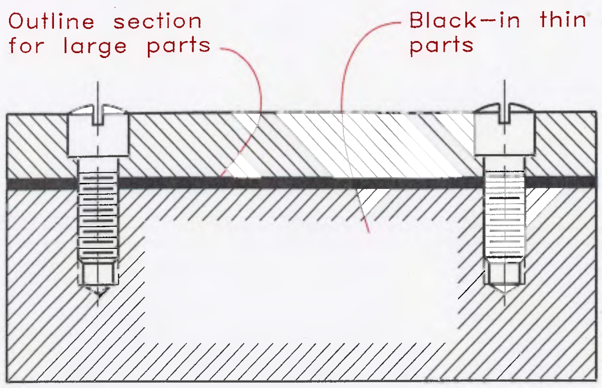


Figure 16.7 Black–in thin parts and hatch large areas around their outlines (outline sectioning) to save time and effort.

HATCH-LINE ANGLES

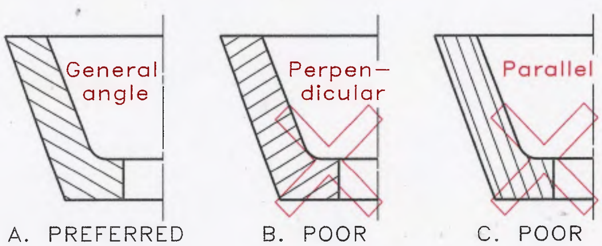


Figure 16.8 Draw section lines at angles that are neither parallel nor perpendicular to the outline of a part, so that they are not misunderstood as machining features.

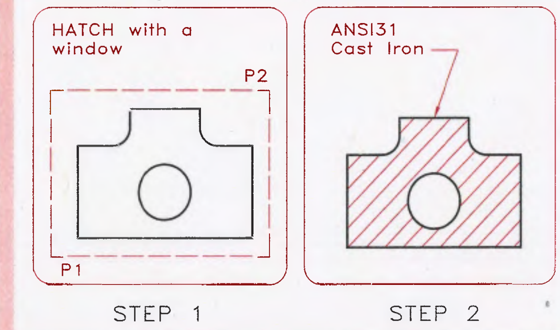


Figure 16.9 Sectioning by computer:
Step 1 `Command:` <u>`HATCH`</u> (CR)
`Pattern (? or name/U, style):` <u>`ANSI31`</u> (CR)
`Scale for pattern <default>:` <u>`.50`</u> (CR)
`Angle for pattern <default>:` <u>`0`</u> (CR)
`Select objects:` <u>`W`</u> (window option)
Step 2 Press (CR) and the hatch is completed. Press Ctrl C to terminate hatching if desired.

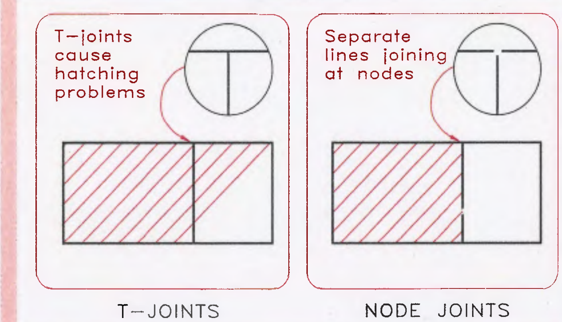


Figure 16.10 For the `HATCH` command to work properly, the outlines of areas to be section lined must be drawn with perfect outlines, that is, without T-joints, gaps, or overlaps.

Computer Method **Figure 16.9** shows the method of applying section symbols to an area with AutoCAD. After assigning the proper hatch symbol with the `HATCH` command, select the area to be sectioned with a window, and the section lines are drawn. To vary the spacing of the section lines, change the scale factor of the `HATCH` command. `BHATCH`, another hatching command, is covered in Chapter 36.

The lines used to depict sectioned areas must intersect perfectly at each corner point; no T-joints are permitted **(Fig. 16.10)**. Poor intersections may cause hatching symbols to fill the desired area improperly.

16.4 Sectioning Assemblies of Parts

When sectioning an assembly of several parts, draw section lines at varying angles to distinguish the parts from each other **(Fig. 16.11A)**. Using dif-

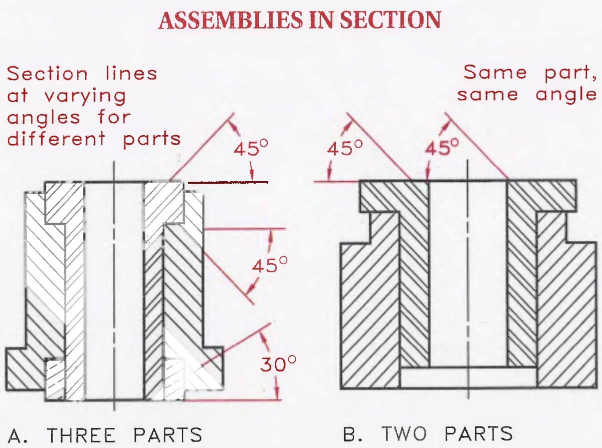


Figure 16.11 Hatching assemblies:

A Draw section lines of different parts in an assembly at varying angles to distinguish the parts.

B Draw section lines on separated portions of the same part (both sides of a hole here) in the same direction.

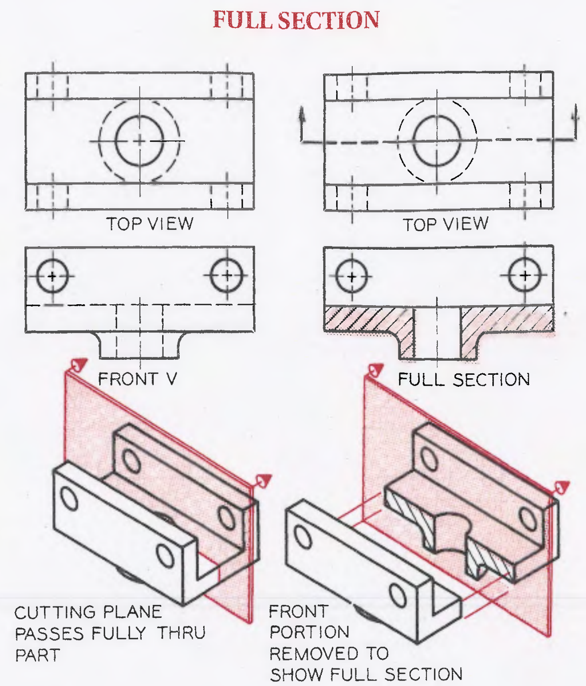


Figure 16.12 A full section is found by passing a cutting plane fully through the top view of this part, removing half of it. The arrows at each end of the cutting plane indicate the direction of your sight. The sectional view shows the part's internal features clearly.

ferent material symbols in an assembly also helps distinguish between the parts and their materials. Cross-hatch the same part at the same angle and with the same symbol even though portions of the part may be separated **(Fig. 16.11B)**.

16.5 Full Sections

A cutting plane passed fully through an object and removing half of it forms a full section view. **Figure 16.12** shows two orthographic views of an object with all its hidden lines. We can describe the part better by passing a cutting plane through the top view to remove half of it. The arrows on the cutting plane indicate the direction of sight. The front view becomes a full section, showing the surfaces cut by the cutting plane.

Figure 16.13 shows a full section through a cylindrical part, with half the object removed. **Figure 16.13A** shows the correctly drawn sectional view. A common mistake in constructing sections is omitting the visible lines behind the cutting plane **(Fig. 16.13B)**.

Omit hidden lines in sectional views unless you consider them necessary for a clear understanding of the view. Also, omit cutting planes if you consider them unnecessary. **Figure 16.14** shows a full section of a part from which the cutting plane was omitted because its path is obvious.

Parts Not Requiring Section Lining

Many standard parts, such as nuts and bolts, rivets, shafts, and set screws, do not require section lining even though the cutting plane passes through them **(Fig. 16.15)**. These parts have no internal features, so sections through them would be of no value. Other parts not requiring section lining are roller bearings, ball bearings, gear teeth, dowels, pins, and washers.

FULL SECTION: CYLINDRICAL PART

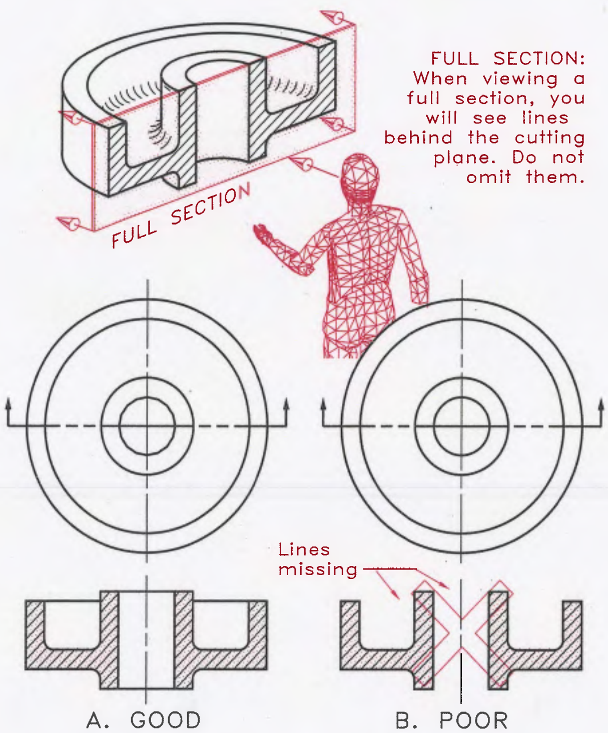


Figure 16.13
A When a cutting plane is passed through a cylinder to obtain a full section, you will see lines behind the plane, not just the cut surface.

B Showing only the lines at the cutting plane's surface yields an incomplete view.

FULL SECTION: PULLEY ARM

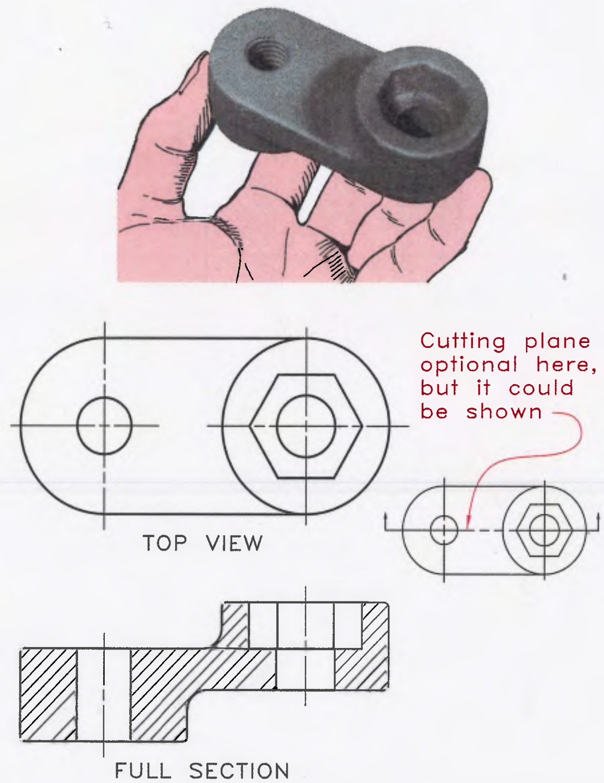


Figure 16.14 The cutting plane of a section can be omitted if its location is obvious.

Ribs

Ribs are not section lined when the cutting plane passes flatwise through them **(Fig. 16.16A)**, because to do so would give a misleading impression of the rib. But ribs do require section lining when the cutting plane passes perpendicularly through them and shows their true thickness **(Fig. 16.16B)**.

Figure 16.17 shows an alternative method of section lining webs and ribs. The outside ribs in **Fig. 16.17A** do not require section lining because the cutting plane passes flatwise through them

PARTS NOT HATCHED

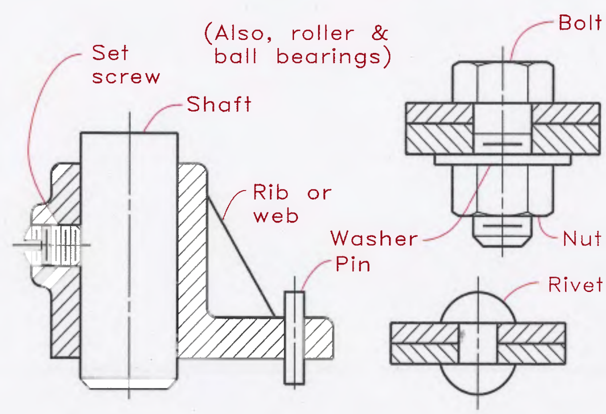


Figure 16.15 By conventional practice these parts are not section lined even though cutting planes pass through them.

RIBS IN SECTION

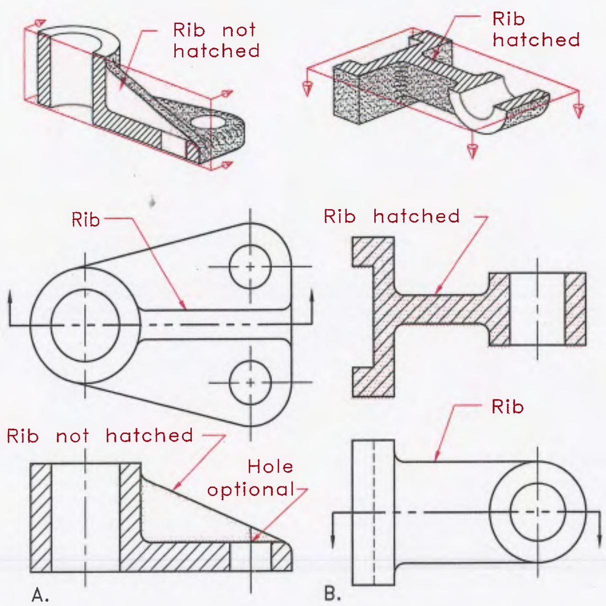


Figure 16.16
A Do not hatch a rib cut in a flatwise direction.

B Hatch ribs when cutting planes pass through them, showing their true thickness.

RIBS AND WEBS IN SECTION

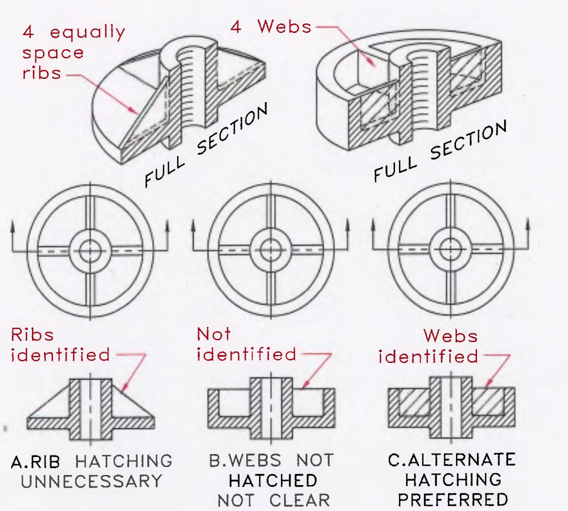


Figure 16.17
A You need not hatch well-defined outside ribs in section.

B Define poorly identified webs by (C) using alternative hatching to call attention to them.

RIBS IN SECTION

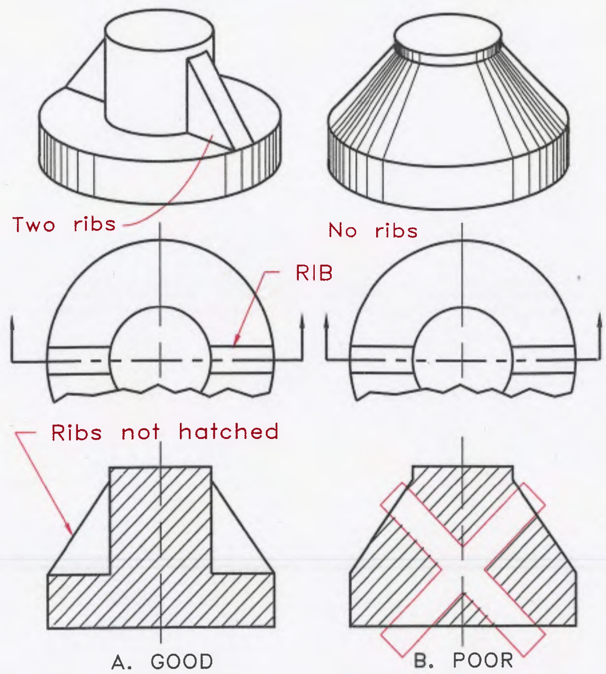


Figure 16.18 In this case (A) the ribs are not hatched to better describe the part than by (B) hatching them. When you use partial views to save space for drawing sections, remove the portion from the side adjacent to the section.

and they are well identified. As a rule, webs do not require cross-hatching, but the webs shown in **Fig. 16.17B** are not well identified in the front section and could go unnoticed. Therefore using alternate section lines as shown in **Fig. 16.17C** is better. Here, extending every-other section line through the webs ensures that they can be identified easily.

By not section lining the ribs in **Fig. 16.18A** we provided an effective section view of the part. If we had section lined the ribs, the section would give the impression that the part is solid and conical (**Fig. 16.18B**).

16.6 Half Sections

A half section is a view obtained by passing a cutting plane halfway through an object and removing a quarter of it to show both external and

HALF SECTION

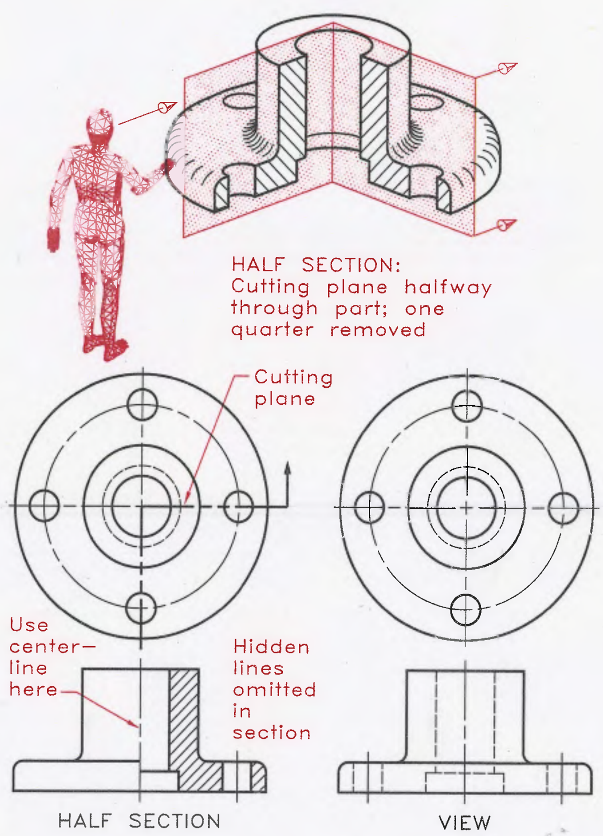


Figure 16.19 In a half section the cutting plane passes halfway through the object, removing a quarter of it, to show half the outside and half the inside of the object. Omit hidden lines unless you need them to clarify the view.

HALF SECTION: PULLEY

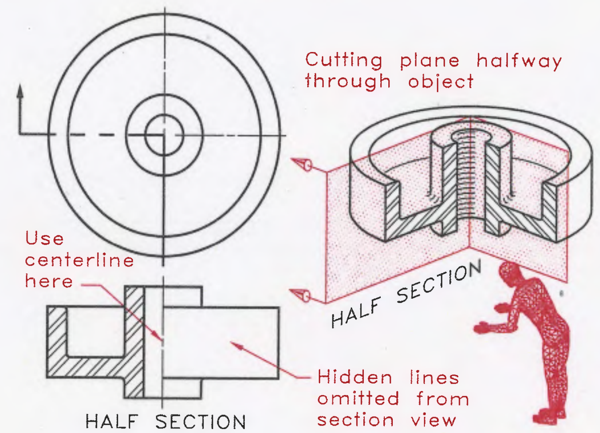


Figure 16.20 This half section describes the part that is shown orthographically and pictorially.

HALF SECTION

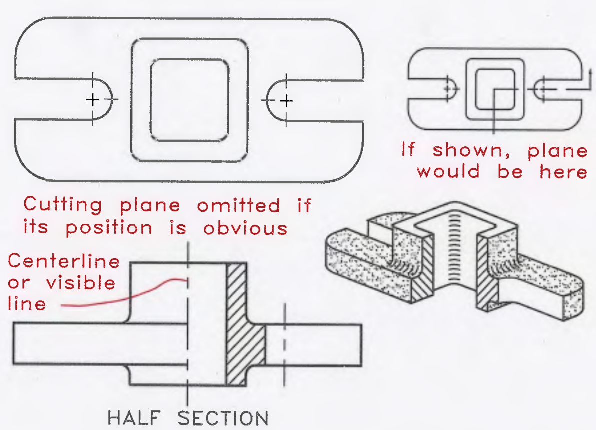


Figure 16.21 The cutting plane can be omitted when its location is obvious. The parting line between the section and the view may be a visible line or a centerline if the part is not cylindrical.

internal features. Half sections are used with symmetrical parts and with cylinders, in particular, as shown in **Fig. 16.19**. By comparing the half section with the standard front view, you can see that both internal and external features show more clearly in a half section than in a view. Hidden lines are unnecessary, and we've omitted them to simplify the section. **Figure 16.20** shows a half section of a pulley.

Note omission of the cutting plane from the half section shown in **Fig. 16.21** because the cutting plane's location is obvious. **Because the parting line of the half section is not at a centerline, you may use a solid line or a centerline to separate the sectional half from the half that appears as an external view.**

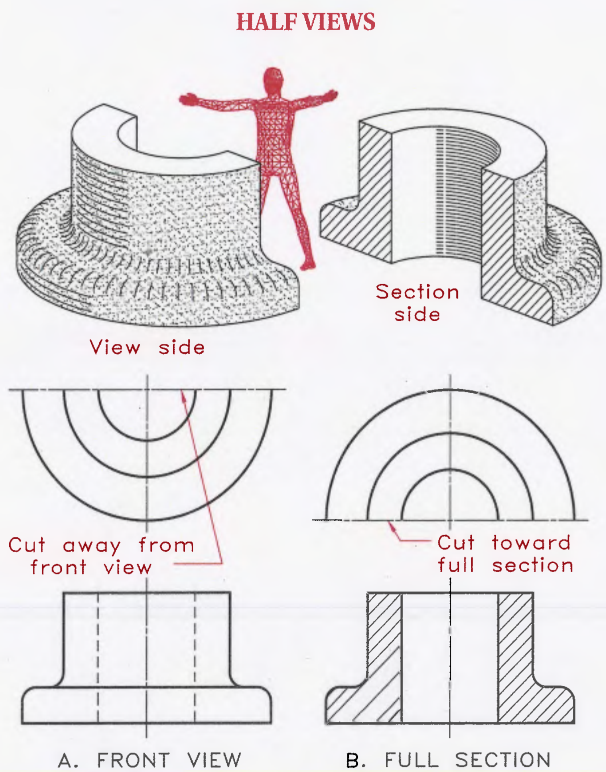


Figure 16.22 Half views of symmetrical objects can be used to conserve space and drawing time. (A) The omitted portion of the view is away from the front view. (B) The omitted portion of the top view is adjacent to the section. In half sections, the omitted half view may be either adjacent to or away from the section.

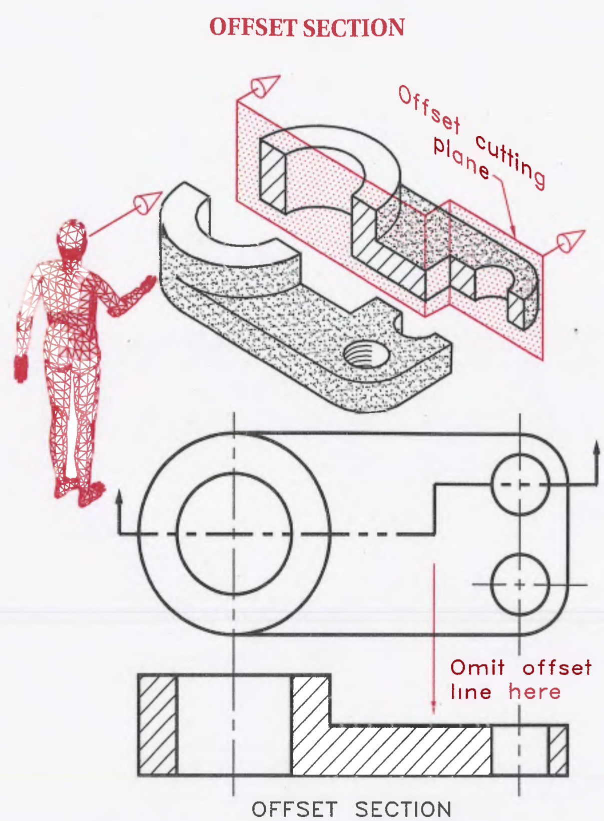


Figure 16.23 An offset section is formed by a cutting plane that must be offset to pass through features not in a single plane. Here, the offset cutting plane is drawn in the top view and the front view is drawn as an offset section.

16.7 Half Views

Figure 16.22 shows **half views**, or conventional methods of representing symmetrical views that require less space and less time to draw than full views. A half top view is sufficient when drawn adjacent to the section view or front view. For half views (not sections), the removed half is the half away from the adjacent view **(Fig. 16.22A)**. For full sections, the removed half is the half nearest the section **(Fig. 16.22B)**. When drawing partial views with half sections, you may omit either the near or the far halves of the partial views.

16.8 Offset Sections

An **offset section** is a full section in which the cutting plane is offset to pass through important features that do not lie in a single plane. **Figure 16.23** shows an offset section in which the plane is offset to pass through the large hole and one of the small holes. The cut formed by the offset is not shown in the section because it is imaginary.

16.9 Broken-Out Sections

A **broken-out section** shows a partial view of a part's interior features. The broken-out section of the part shown in **Fig. 16.24** reveals details of the

BROKEN-OUT SECTION

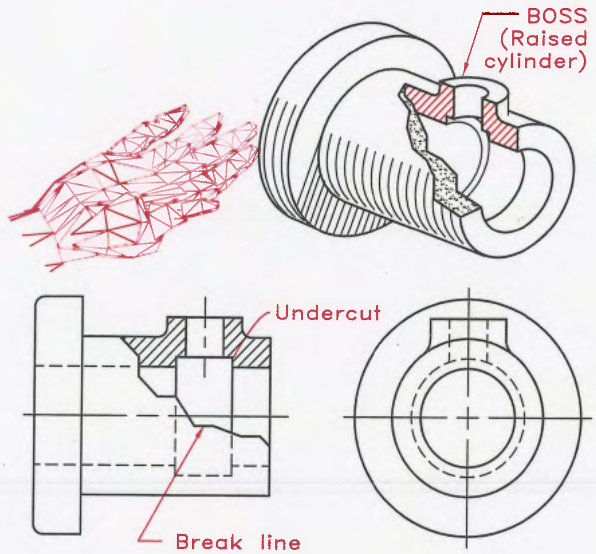


Figure 16.24 To find a broken-out section, imagine that part of the object has been broken away to reveal interior features.

BROKEN-OUT SECTION: PULLEY

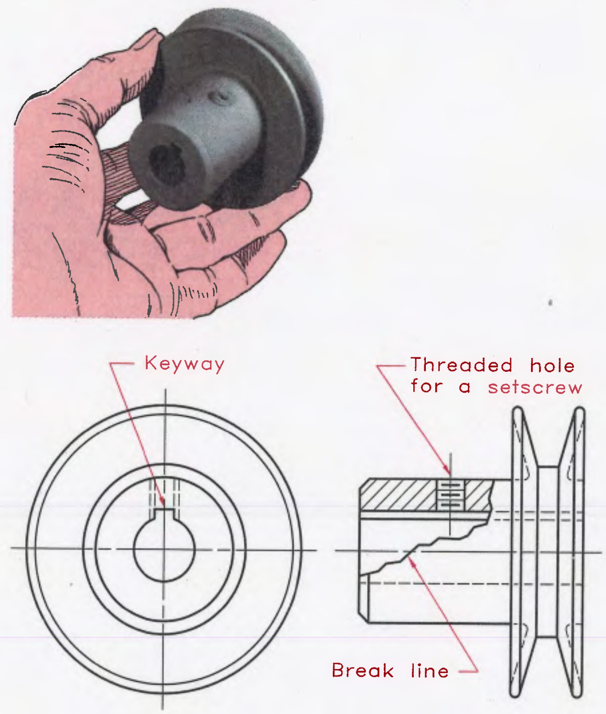


Figure 16.25 This broken-out section effectively shows the keyway and threaded hole for a setscrew in the pulley.

wall thickness to describe the part better. The irregular lines representing the break are conventional breaks (discussed later in this chapter).

The broken-out section of the pulley in **Fig. 16.25** clearly depicts the keyway and threaded hole for a setscrew. This method shows the part efficiently, with the minimum of views.

16.10 Revolved Sections

A **revolved section** describes a part when you revolve its cross section about an axis of revolution and place it on the view where the revolution occurred. Note the use of revolved sections to explain two cross sections of the shaft shown in **Fig. 16.26** (with and without conventional breaks). Conventional breaks are optional; you may draw a revolved section on the view without them.

REVOLVED SECTIONS

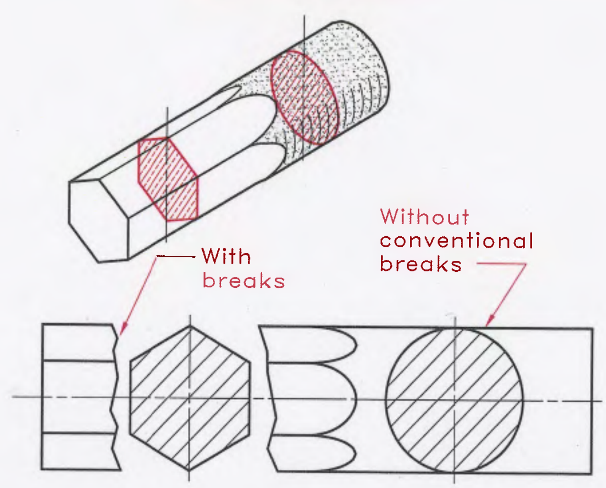


Figure 16.26 Revolved sections show cross-sectional features of a part to eliminate the need for supplementary orthographic views. You may superimpose revolved sections on the given views or use conventional breaks to separate them from the given views.

REVOLVED SECTION

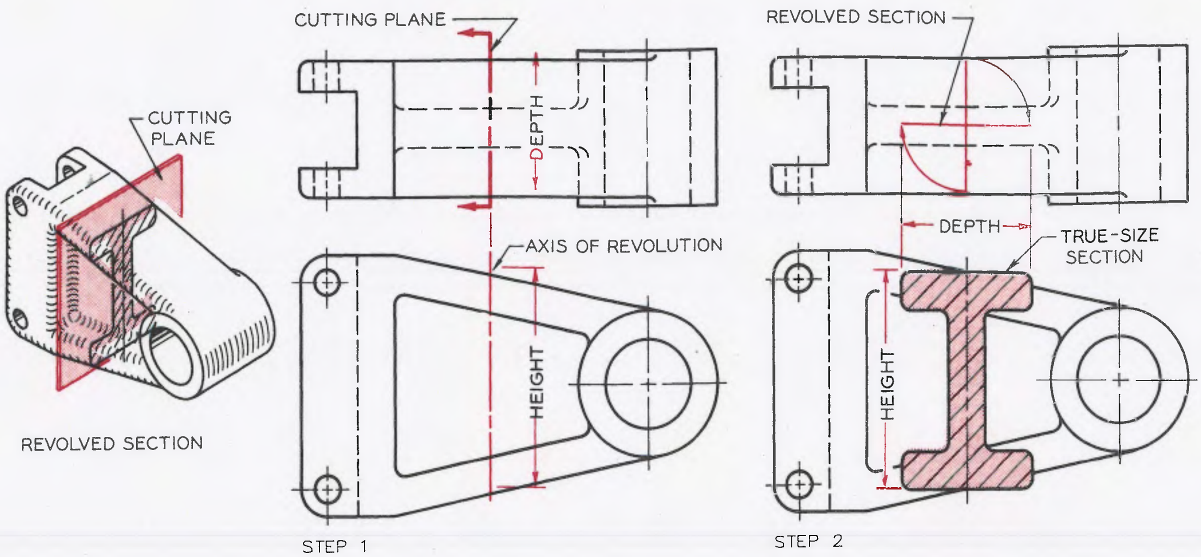


Figure 16.27 Drawing revolved sections:
Step 1 Show an axis of revolution in the front view. The cutting plane would appear as an edge in the top view if you were to show it.

Step 2 Revolve the vertical section in the top view to show the section at true size in the front view. Do not draw object lines through the revolved section.

REVOLVED SECTIONS

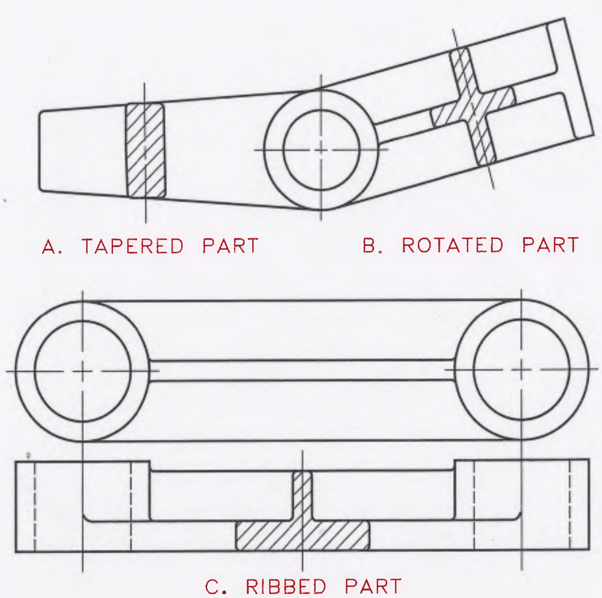


Figure 16.28 These revolved sections help describe the cross sections of the two parts and make complex orthographic views unnecessary.

A revolved section helps to describe the part shown in **Fig. 16.27**. Imagine passing a cutting plane through the top view of the part (step 1). Then imagine revolving the cutting plane in the top view to obtain a true-size revolved section in the front view (step 2). Conventional breaks could be used on each side of the revolved section.

Figure 16.28 demonstrates how to use typical revolved sections to show cross sections through parts without having to draw additional orthographic views.

16.11 Conventional Revolutions

In **Fig. 16.29A**, the middle hole is omitted because it does not pass through the center of the circular plate. However, in **Fig. 16.29B**, the hole does pass through the plate's center and is shown in the section. Although the cutting plane does not pass through one of the symmetrically spaced holes in

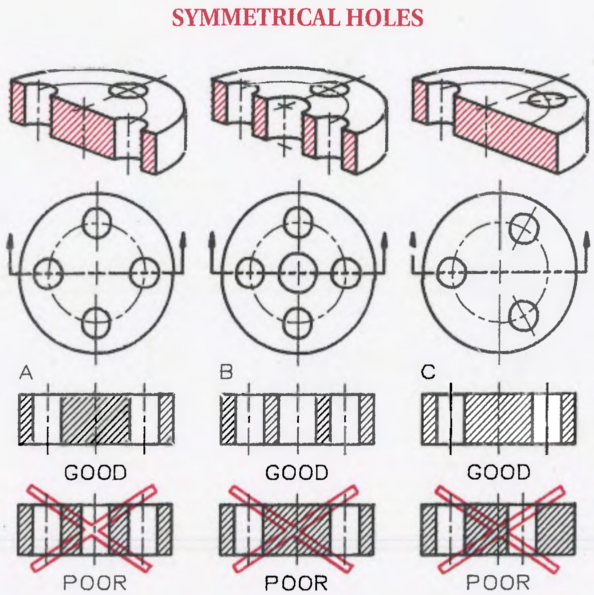


Figure 16.29 Revolve symmetrically spaced holes to show their true radial distances from their centers in sectional views. (A) Do not show the middle hole because it is not at the center of the plate. (B) Show the center hole because it is at the center of the plate. (C) Rotate one of the holes to the cutting plane to make the sectional view symmetrical and more descriptive.

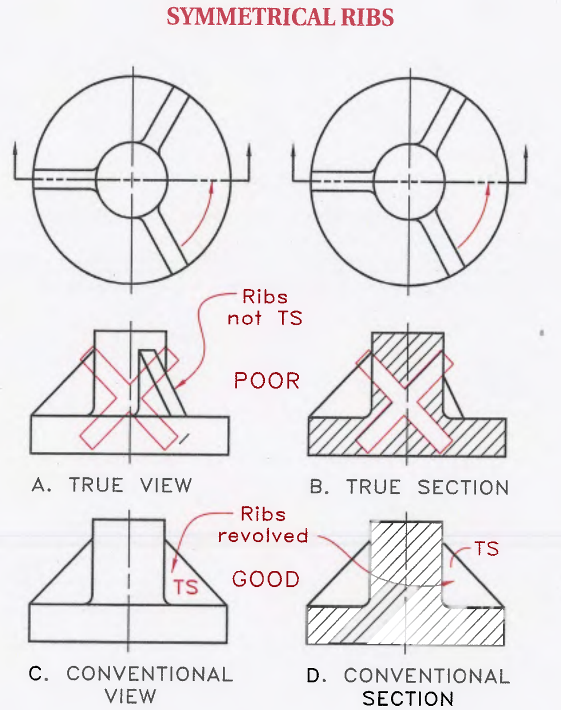


Figure 16.30 Show symmetrically spaced ribs revolved in both orthographic and sectional views to show them true size as a conventional practice.

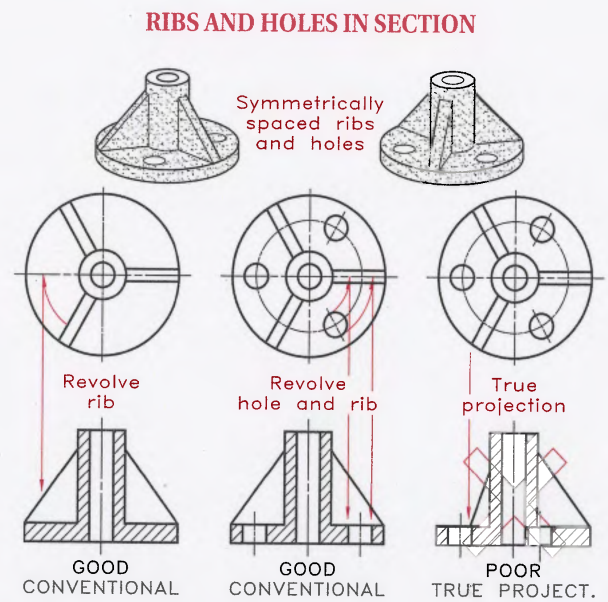


Figure 16.31 Show parts having symmetrically spaced ribs and holes in section with ribs rotated to show their true size and holes rotated to show them at their true radial distance from the center.

the top view (**Fig. 16.29C**), the hole is revolved to the cutting plane to show the full section.

When ribs are symmetrically spaced about a hub (**Fig. 16.30**), it is conventional practice to revolve them so that they appear true size in both views and sections. **Figure 16.31** illustrates the conventional practice of revolving both holes and ribs (or webs) of symmetrical parts. Revolution gives a better description of the parts in a manner that is easier to draw.

A cutting plane may be positioned in either of two ways shown in **Fig. 16.32**. Even though the cutting plane does not pass through the ribs and holes in **Fig. 16.32A**, they may appear in section as if the cutting plane passed through them. The path of the cutting plane also may be revolved, as shown in **Fig. 16.32B**. In this case the ribs are revolved to their true-size position in the section view, although the plane does not cut them.

CUTTING-PLANE POSITIONS

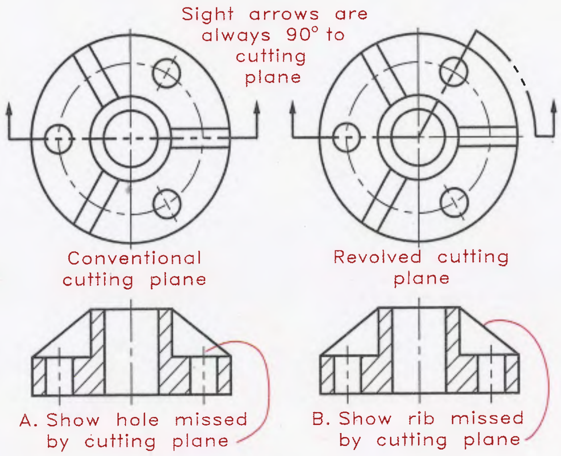


Figure 16.32 Show symmetrically located ribs true size in section whether the cutting plane passes through them or not. You may revolve the path of the cutting plane through certain features if you want, but the sight arrows are always perpendicular to the cutting plane.

SPOKES IN SECTION

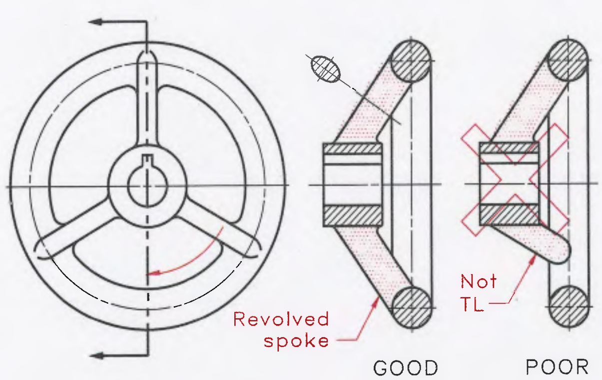


Figure 16.33 Revolve symmetrically spaced spokes to show them at true size in section. Do not section line spokes.

The same principles apply to symmetrically spaced spokes (**Fig. 16.33**). Draw only the revolved, true-size spokes and do not section line them. If the spokes shown in **Fig. 16.34A** were hatched, they could be misunderstood as a solid web, as shown in **Fig. 16.34B**.

WEBS AND SPOKES IN SECTION

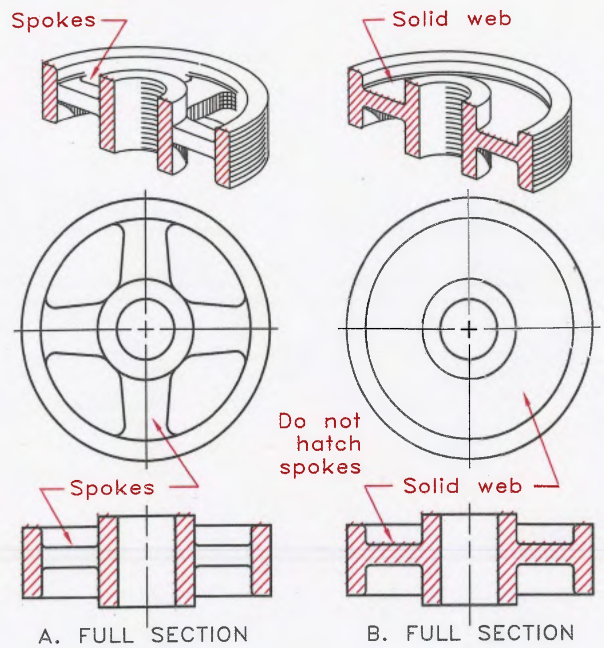


Figure 16.34

A Do not hatch spokes when the cutting plane passes through them.

B Hatch solid webs in sections of this type.

Revolving the symmetrically positioned lugs shown in **Fig. 16.35** yields their true size in both the front view and section. The same principles of rotation apply to the part shown in **Fig. 16.36**, where the inclined arm appears in the section as if it had been revolved to the centerline in the top view and then projected to the sectional view.

16.12 Removed Sections

A **removed section** is a revolved section that is shown outside the view in which it was revolved (**Fig. 16.37**). Centerlines are used as axes of rotation to show the locations from which the sections are taken. Where space does not permit revolution on

LUGS IN SECTION

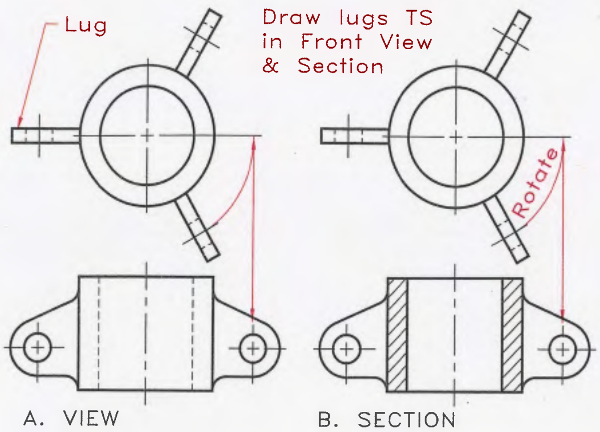


Figure 16.35 Revolve symmetrically spaced lugs (flanges) to show their true size in the (A) front view and (B) in sections.

REVOLVED FEATURES

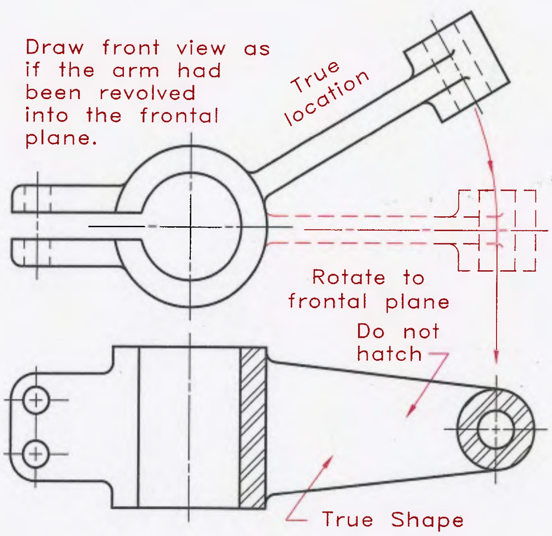


Figure 16.36 It is conventional practice to revolve a part with an inclined arm extending from a circular hub as if it were true shape in the sectional view.

the given view **(Fig. 16.38A)**, removed sections must be used instead of revolved sections **(Fig. 16.38B)**.

Removed sections do not have to position directly along an axis of revolution adjacent to the view from which they were revolved. Instead,

REMOVED SECTIONS

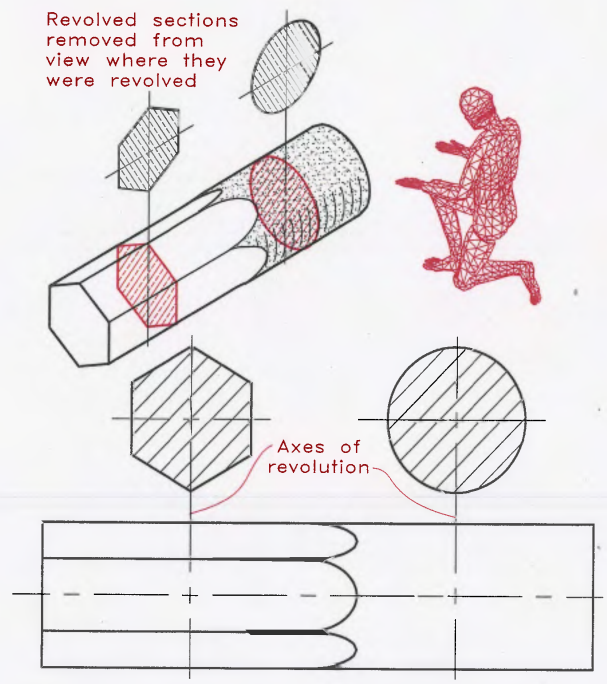


Figure 16.37 Removed sections are revolved sections that are drawn outside the object along their axes of revolution.

REVOLVED AND REMOVED SECTIONS

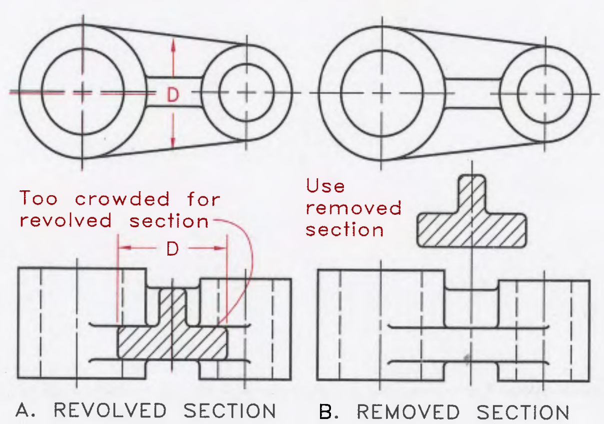


Figure 16.38 Removed sections are necessary when space does not permit the use of revolved sections.

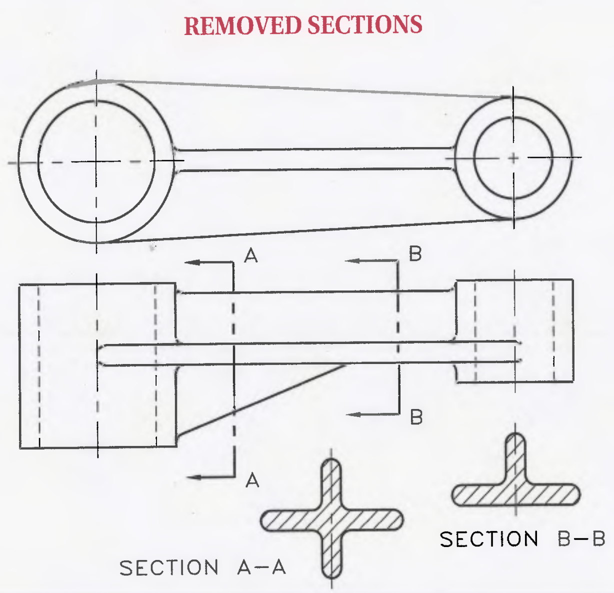


Figure 16.39 Lettering each end of a cutting plane (such as A–A) identifies the removed section labeled Section A–A shown elsewhere on the drawing.

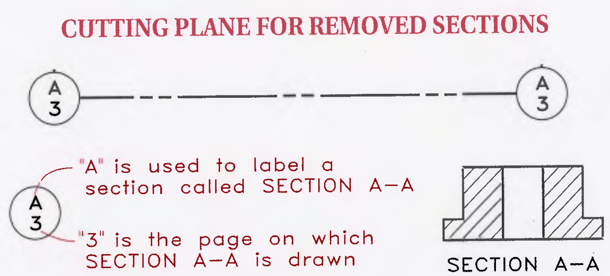


Figure 16.40 If placing a removed section on another page in a set of drawings is necessary, label each end of the cutting plane with a letter and a number. Here, the letters refer to Section A–A, and the numbers mean that Section A–A appears on page 3.

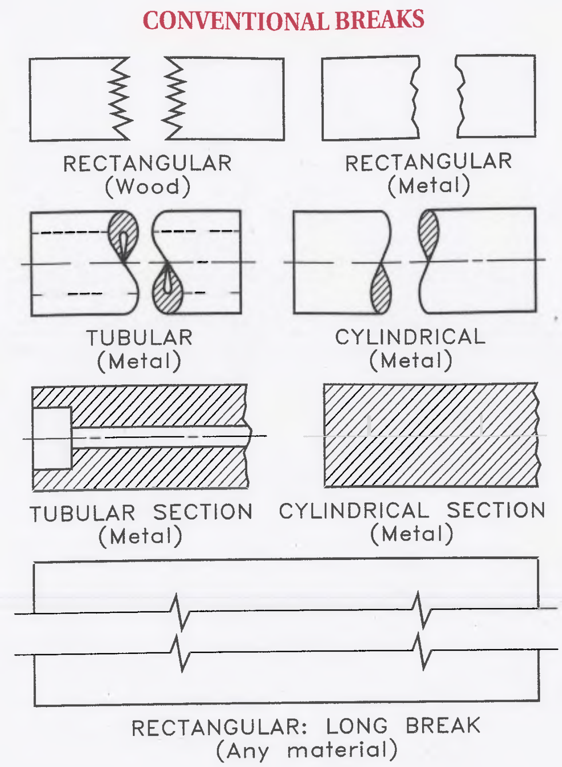


Figure 16.41 These conventional breaks indicate that a portion of an object is not shown.

removed sections can be located elsewhere on a drawing if they are properly labeled **(Fig. 16.39)**. For example, the plane labeled with an A at each end identifies the location of section A–A; the same applies to section B–B.

When a set of drawings consists of multiple sheets, removed sections and the views from which they are taken may appear on different sheets. When this method of layout is necessary, label the cutting plane in the view from which the section was taken and the sheet on which the section appears **(Fig. 16.40)**.

16.13 Conventional Breaks

Figure 16.41 shows types of conventional breaks to use when you remove portions of an object. You may draw the "figure-eight" breaks used for cylindrical and tubular parts freehand **(Fig. 16.42)** or with a compass when they are larger **(Fig. 16.43)**.

One use of conventional breaks is to shorten a long piece by removing the portion between the breaks so that it may be drawn at a larger size **(Fig. 16.44)**. The dimension specifies the true length of

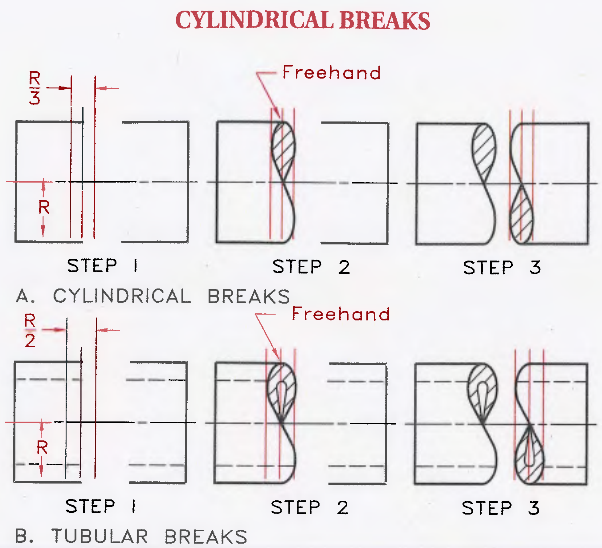


Figure 16.42 Guidelines aid in drawing conventional breaks in (A) cylindrical and (B) tubular sections freehand, as shown here. The radius, R, establishes the width of both "figure-eight" break symbols.

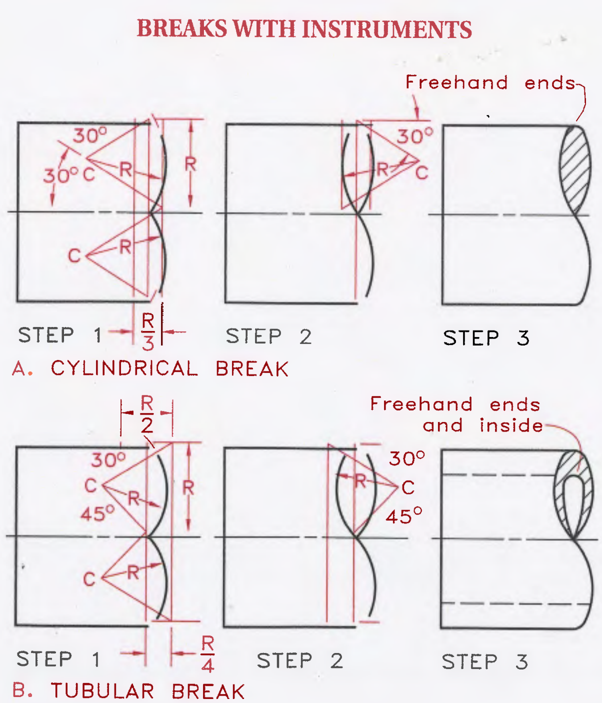


Figure 16.43 Instruments help in drawing conventional breaks in larger cylindrical and tubular parts.

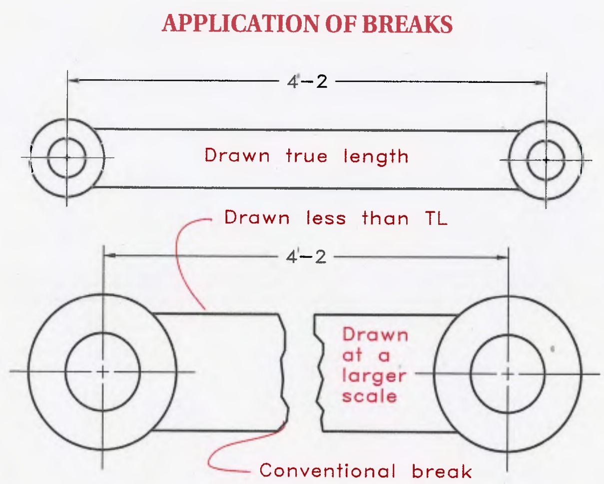


Figure 16.44 The use of conventional breaks allows this part to be drawn effectively at a larger scale. It is permissible to insert a revolved section between the breaks.

the part, and the breaks indicate that a portion of the length has been removed.

16.14 Phantom (Ghost) Sections

A phantom or ghost section depicts parts as if they were being X-rayed. In **Fig. 16.45**, the cutting plane is used in the normal manner, but the section lines are drawn as dashed lines. If the object were shown as a regular full section, the circular hole through the front surface could not be shown in the same view.

16.15 Auxiliary Sections

You may use auxiliary sections to supplement the principal views of orthographic projections **(Fig. 16.46)**. Pass auxiliary cutting plane A–A through the front view and project the auxiliary view from the cutting plane as indicated by the sight arrows. Section A–A gives a cross-sectional description of the part that would be difficult to depict by other principal orthographic views.

PHANTOM SECTION

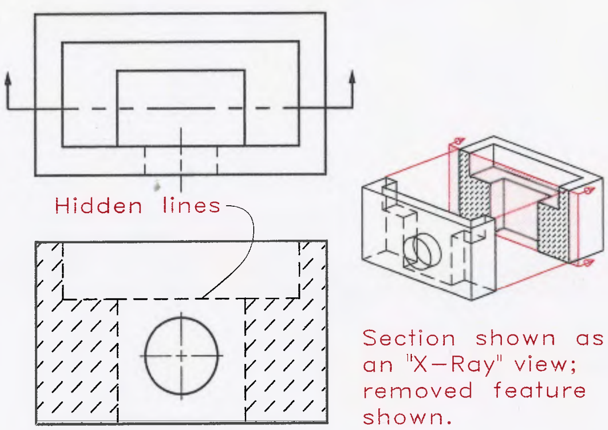


Figure 16.45 Phantom sections give an "X-ray" view of an object, allowing you to show the hole in the front of the cutting plane. Draw section lines as dashed lines.

AUXILIARY SECTION

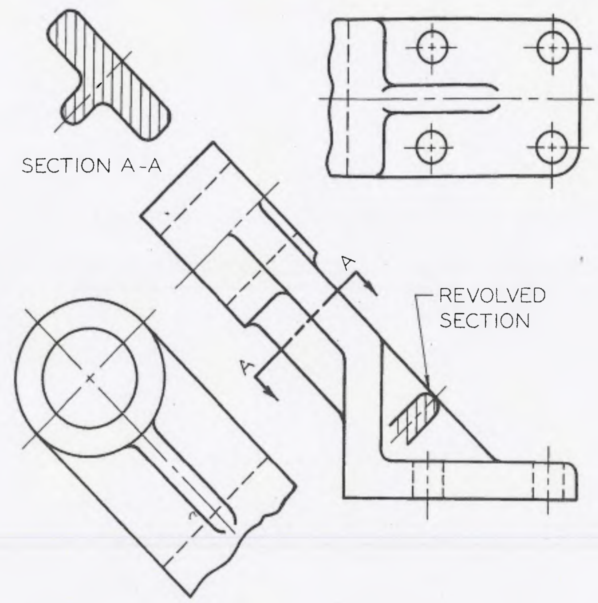


Figure 16.46 Auxiliary sections can be used as supplementary views to add clarity to a drawing.

Problems

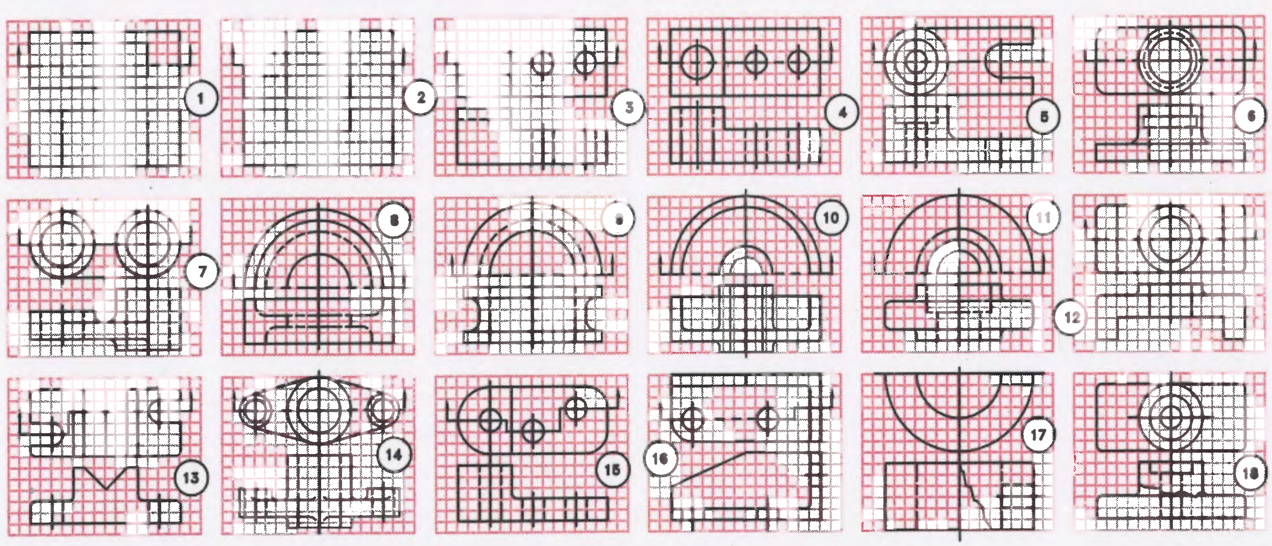


Figure 16.47 Problems 1–24. Introductory sections.

(continued)

Solve the problems shown in **Fig. 16.47** on size A sheets by drawing two solutions per sheet. Each grid space equals 0.20 in., or 5 mm.

1–16. Draw the sections indicated by the cutting planes.

17–20. Draw broken-out sections.

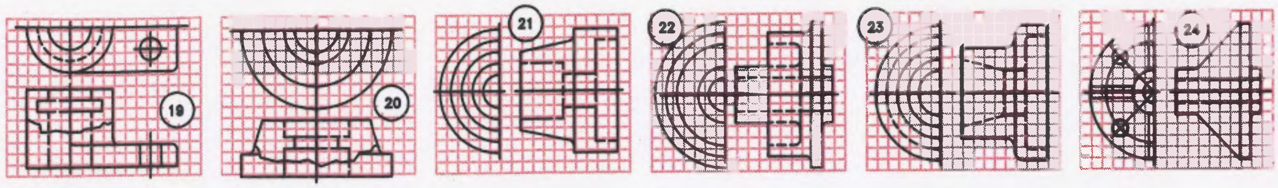


Fig. 16.47 continued

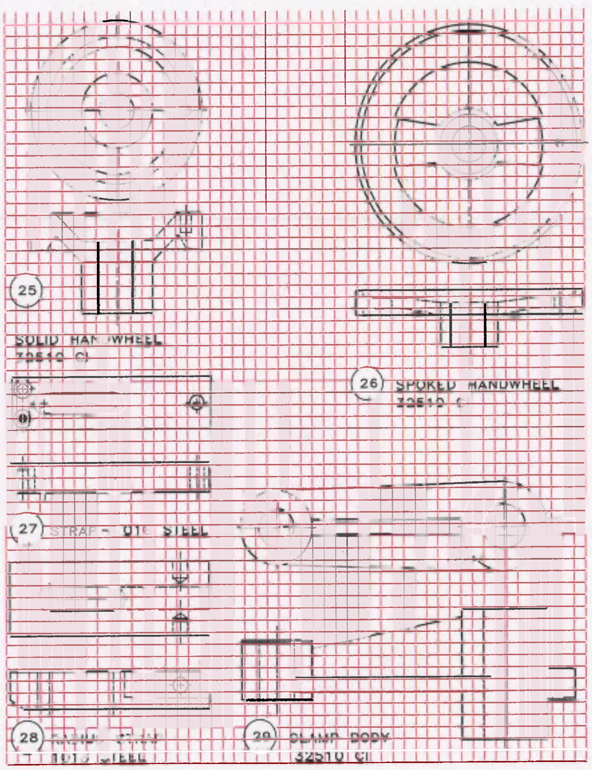


Figure 16.48 Problems 25–29.

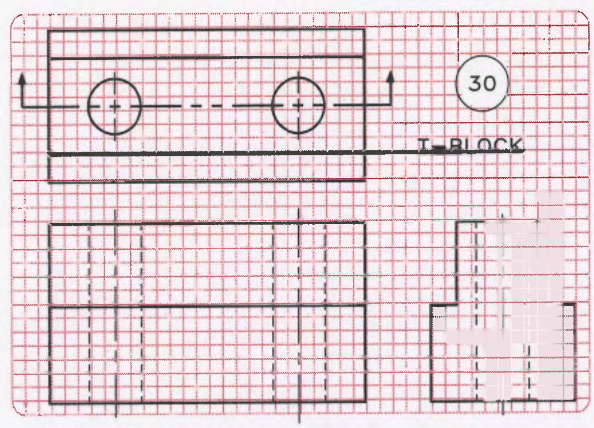


Figure 16.49 Problem 30. Full section.

21–24. Draw half sections.

25–29. (Fig. 16.48) Construct views of these fixtures on size A sheets and draw appropriate sections to describe them clearly. Exercise your design ability and add finish marks, fillets, and rounds where you believe they are needed but not shown. Use the grid—0.20 in., or 5 mm—to determine the full-size dimensions of the parts. Select the best scale and sheet size for each solution. (Courtesy of Jergens, Incorporated, Cleveland, Ohio.)

30–34. (Figs. 16.49–16.53) Complete these drawings as full sections. Draw one solution per size AH sheet. Each grid space equals 0.20 in., or 5 mm. Show the cutting planes in each solution.

35–36. (Figs. 16.54–16.55) Complete the drawings as half sections with one solution per size AH sheet. Each grid space equals 0.20 in., or 5 mm. Show the cutting planes in each solution.

37–38. (Figs. 16.56–16.57) Complete the drawings as offset sections. Draw one solution per size AH sheet. Each grid space equals 0.20 in., or 5 mm. Show the cutting planes in each solution.

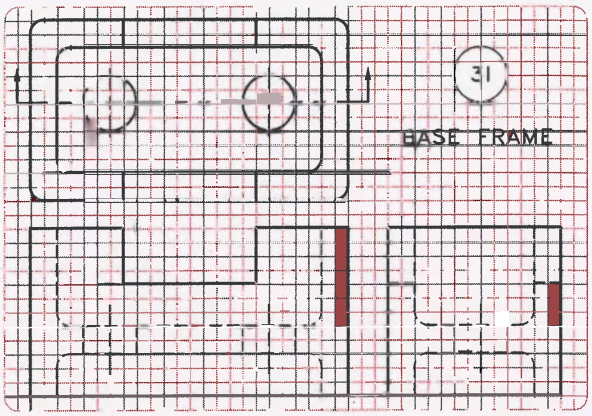


Figure 16.50 Problem 31. Full section.

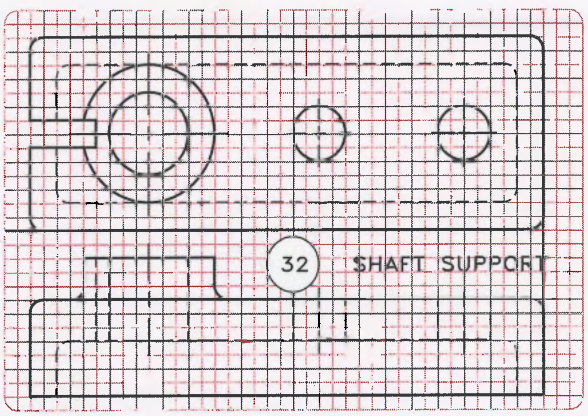


Figure 16.51 Problem 32. Full section.

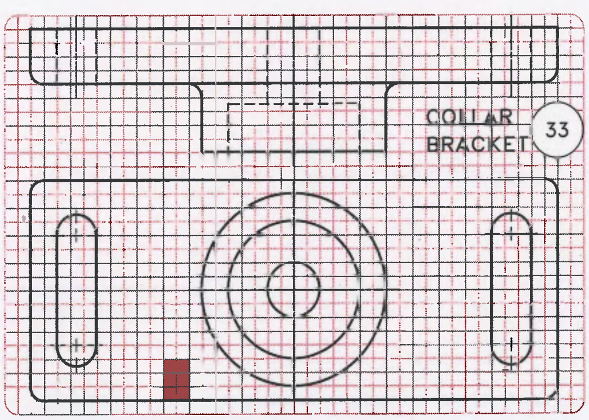


Figure 16.52 Problem 33. Full section.

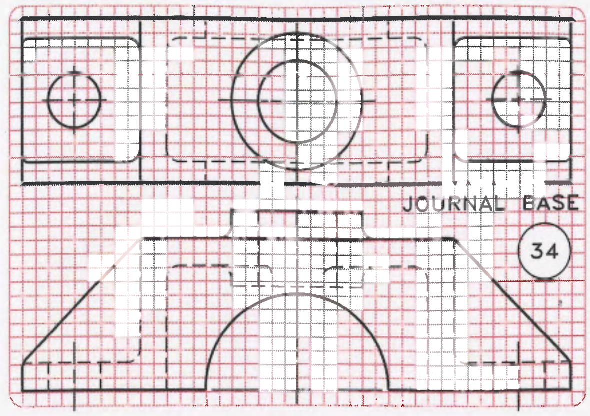


Figure 16.53 Problem 34. Full section.

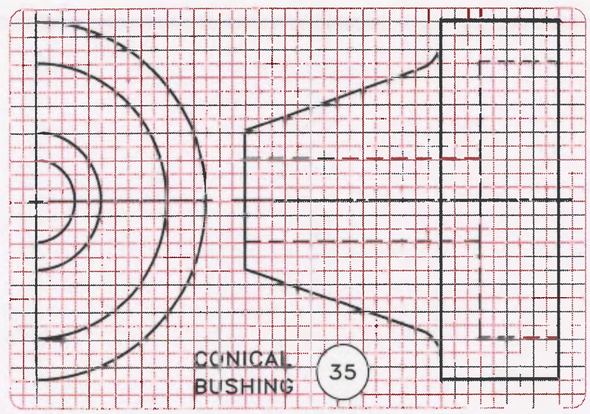


Figure 16.54 Problem 35. Half section.

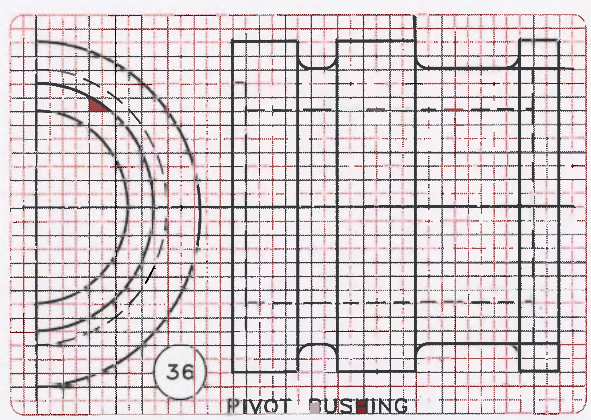


Figure 16.55 Problem 36. Half section.

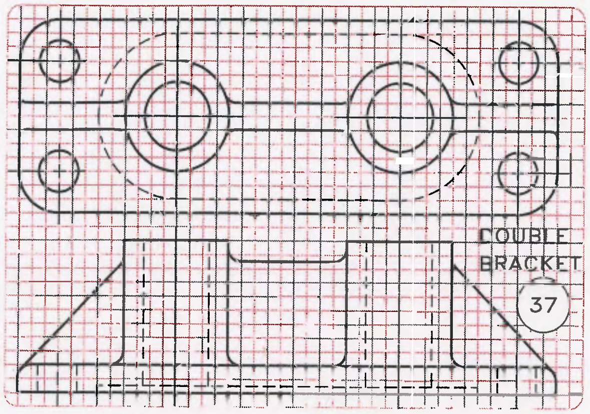


Figure 16.56 Problem 37. Offset section.

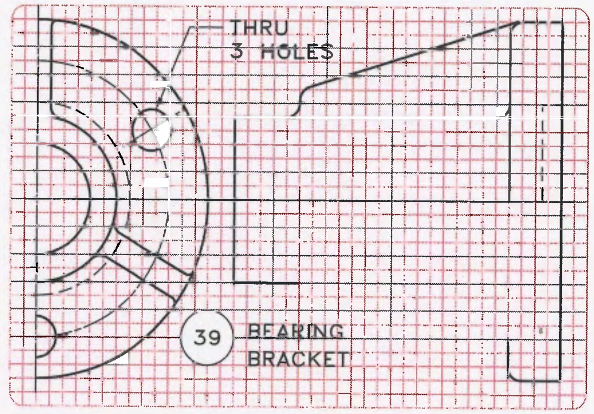


Figure 16.58 Problem 39. Full section.

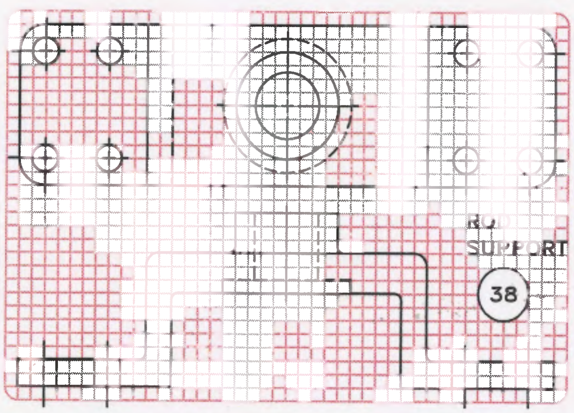


Figure 16.57 Problem 38. Offset section.

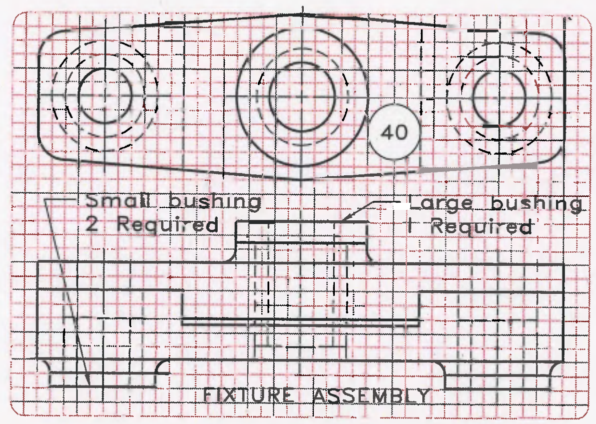


Figure 16.59 Problem 40. Full section in assembly.

39. (Fig. 16.58) Complete the partial view as a full section. Draw the views on a size AH sheet. Each grid space equals 0.20 in., or 5 mm. Show the cutting plane in your solution.

40. (Fig. 16.59) Complete the front view as a full section of the assembly. Draw the views on a size AH sheet. Each grid space equals 0.20 in., or 5 mm. Show the cutting plane in your solution.

Exhibit B

DOE-HDBK-1012/3-92
JUNE 1992

DOE FUNDAMENTALS HANDBOOK
THERMODYNAMICS, HEAT TRANSFER, AND FLUID FLOW
Volume 3 of 3

U.S. Department of Energy
Washington, D.C. 20585

FSC-6910

This document has been reproduced directly from the best available copy.

Available to DOE and DOE contractors from the Office of Scientific and Technical Information. P. O. Box 62, Oak Ridge, TN 37831; prices available from (615) 576-8401. FTS 626-8401.

Available to the public from the National Technical Information Service, U.S. Department of Commerce, 5285 Port Royal Rd., Springfield, VA 22161.

Order No. DE92019791

ABSTRACT

The *Thermodynamics, Heat Transfer, and Fluid Flow Fundamentals Handbook* was developed to assist nuclear facility operating contractors provide operators, maintenance personnel, and the technical staff with the necessary fundamentals training to ensure a basic understanding of the thermal sciences. The handbook includes information on thermodynamics and the properties of fluids; the three modes of heat transfer - conduction, convection, and radiation; and fluid flow, and the energy relationships in fluid systems. This information will provide personnel with a foundation for understanding the basic operation of various types of DOE nuclear facility fluid systems.

Key Words: Training Material, Thermodynamics, Heat Transfer, Fluid Flow, Bernoulli's Equation

FOREWORD

The *Department of Energy (DOE) Fundamentals Handbooks* consist of ten academic subjects, which include Mathematics; Classical Physics; Thermodynamics, Heat Transfer, and Fluid Flow; Instrumentation and Control; Electrical Science; Material Science; Mechanical Science; Chemistry; Engineering Symbology, Prints, and Drawings; and Nuclear Physics and Reactor Theory. The handbooks are provided as an aid to DOE nuclear facility contractors.

These handbooks were first published as Reactor Operator Fundamentals Manuals in 1985 for use by DOE Category A reactors. The subject areas, subject matter content, and level of detail of the Reactor Operator Fundamentals Manuals was determined from several sources. DOE Category A reactor training managers determined which materials should be included, and served as a primary reference in the initial development phase. Training guidelines from the commercial nuclear power industry, results of job and task analyses, and independent input from contractors and operations-oriented personnel were all considered and included to some degree in developing the text material and learning objectives.

The *DOE Fundamentals Handbooks* represent the needs of various DOE nuclear facilities' fundamentals training requirements. To increase their applicability to nonreactor nuclear facilities, the Reactor Operator Fundamentals Manual learning objectives were distributed to the Nuclear Facility Training Coordination Program Steering Committee for review and comment. To update their reactor-specific content, DOE Category A reactor training managers also reviewed and commented on the content. On the basis of feedback from these sources, information that applied to two or more DOE nuclear facilities was considered generic and was included. The final draft of each of these handbooks was then reviewed by these two groups. This approach has resulted in revised modular handbooks that contain sufficient detail such that each facility may adjust the content to fit their specific needs.

Each handbook contains an abstract, a foreword, an overview, learning objectives, and text material, and is divided into modules so that content and order may be modified by individual DOE contractors to suit their specific training needs. Each subject area is supported by a separate examination bank with an answer key.

The *DOE Fundamentals Handbooks* have been prepared for the Assistant Secretary for Nuclear Energy, Office of Nuclear Safety Policy and Standards, by the DOE Training Coordination Program. This program is managed by EG&G Idaho, Inc.

OVERVIEW

The *Department of Energy Fundamentals Handbook* entitled *Thermodynamics, Heat Transfer, and Fluid Flow* was prepared as an information resource for personnel who are responsible for the operation of the Department's nuclear facilities. A basic understanding of the thermal sciences is necessary for DOE nuclear facility operators, maintenance personnel, and the technical staff to safely operate and maintain the facility and facility support systems. The information in the handbook is presented to provide a foundation for applying engineering concepts to the job. This knowledge will help personnel more fully understand the impact that their actions may have on the safe and reliable operation of facility components and systems.

The *Thermodynamics, Heat Transfer, and Fluid Flow* handbook consists of three modules that are contained in three volumes. The following is a brief description of the information presented in each module of the handbook.

Volume 1 of 3

Module 1 - Thermodynamics

This module explains the properties of fluids and how those properties are affected by various processes. The module also explains how energy balances can be performed on facility systems or components and how efficiency can be calculated.

Volume 2 of 3

Module 2 - Heat Transfer

This module describes conduction, convection, and radiation heat transfer. The module also explains how specific parameters can affect the rate of heat transfer.

Volume 3 of 3

Module 3 - Fluid Flow

This module describes the relationship between the different types of energy in a fluid stream through the use of Bernoulli's equation. The module also discusses the causes of head loss in fluid systems and what factors affect head loss.

The information contained in this handbook is by no means all encompassing. An attempt to present the entire subject of thermodynamics, heat transfer, and fluid flow would be impractical. However, the *Thermodynamics, Heat Transfer, and Fluid Flow* handbook does present enough information to provide the reader with a fundamental knowledge level sufficient to understand the advanced theoretical concepts presented in other subject areas, and to better understand basic system and equipment operations.

Department of Energy
Fundamentals Handbook

THERMODYNAMICS, HEAT TRANSFER, AND FLUID FLOW, Module 3 Fluid Flow

TABLE OF CONTENTS

TABLE OF CONTENTS (Cont.)

TABLE OF CONTENTS (Cont.)

LIST OF FIGURES

LIST OF TABLES

REFERENCES

- Streeter, Victor L., Fluid Mechanics, 5th Edition, McGraw-Hill, New York, ISBN 07-062191-9.

- Knudsen, J. G. and Katz, D. L., Fluid Dynamics and Heat Transfer, McGraw-Hill, New York.

- McDonald, A. T. and Fox, R. W., Introduction to Fluid Mechanics, 2nd Edition, John Wiley and Sons, New York, ISBN 0-471-98440-X.

- Crane Company, Flow of Fluids Through Valves, Fittings, and Pipe, Crane Co. Technical Paper No. 410, Chicago, Illinois, 1957.

- Esposito, Anthony, Fluid Power with Applications, Prentice-Hall, Inc., New Jersey, ISBN 0-13-322701-4.

- Wallis, Graham, One-Dimensional Two-Phase Flow, McGraw-Hill, New York, 1969.

- Academic Program for Nuclear Power Plant Personnel, Volume III and IV, General Physics Corporation, Library of Congress Card #A 397747, June 1982 and April 1982.

TERMINAL OBJECTIVE

1.0 Given conditions affecting the fluid flow in a system, **EVALUATE** the effects on the operation of the system.

ENABLING OBJECTIVES

1.1 **DESCRIBE** how the density of a fluid varies with temperature.

1.2 **DEFINE** the term buoyancy.

1.3 **DESCRIBE** the relationship between the pressure in a fluid column and the density and depth of the fluid.

1.4 **STATE** Pascal's Law.

1.5 **DEFINE** the terms mass flow rate and volumetric flow rate.

1.6 **CALCULATE** either the mass flow rate or the volumetric flow rate for a fluid system.

1.7 **STATE** the principle of conservation of mass.

1.8 **CALCULATE** the fluid velocity or flow rate in a specified fluid system using the continuity equation.

1.9 **DESCRIBE** the characteristics and flow velocity profiles of laminar flow and turbulent flow.

1.10 **DEFINE** the property of viscosity.

1.11 **DESCRIBE** how the viscosity of a fluid varies with temperature.

1.12 **DESCRIBE** the characteristics of an ideal fluid.

1.13 **DESCRIBE** the relationship between the Reynolds number and the degree of turbulence of the flow.

1.14 **DESCRIBE** the relationship between Bernoulli's equation and the First Law of Thermodynamics.

ENABLING OBJECTIVES (Cont.)

1.15 **DEFINE** the term head with respect to its use in fluid flow.

1.16 **EXPLAIN** the energy conversions that take place in a fluid system between the velocity, elevation, and pressure heads as flow continues through a piping system.

1.17 Given the initial and final conditions of the system, **CALCULATE** the unknown fluid properties using the simplified Bernoulli equation.

1.18 **DESCRIBE** the restrictions applied to Bernoulli's equation when presented in its simplest form.

1.19 **EXPLAIN** how to extend the Bernoulli equation to more general applications.

1.20 **RELATE** Bernoulli's principle to the operation of a venturi.

1.21 **DEFINE** the terms head loss, frictional loss, and minor losses.

1.22 **DETERMINE** friction factors for various flow situations using the Moody chart.

1.23 **CALCULATE** the head loss in a fluid system due to frictional losses using Darcy's equation.

1.24 **CALCULATE** the equivalent length of pipe that would cause the same head loss as the minor losses that occur in individual components.

1.25 **DEFINE** natural circulation and forced circulation.

1.26 **DEFINE** thermal driving head.

1.27 **DESCRIBE** the conditions necessary for natural circulation to exist.

1.28 **EXPLAIN** the relationship between flow rate and temperature difference in natural circulation flow.

1.29 **DESCRIBE** how the operator can determine whether natural circulation exists in the reactor coolant system and other heat removal systems.

1.30 **DESCRIBE** how to enhance natural circulation flow.

1.31 **DEFINE** two-phase flow.

ENABLING OBJECTIVES (Cont.)

1.32 **DESCRIBE** two-phase flow including such phenomena as bubbly, slug, and annular flow.

1.33 **DESCRIBE** the problems associated with core flow oscillations and flow instability.

1.34 **DESCRIBE** the conditions that could lead to core flow oscillation and instability.

1.35 **DESCRIBE** the phenomenon of pipe whip.

1.36 **DESCRIBE** the phenomenon of water hammer.

1.37 **DEFINE** the terms net positive suction head and cavitation.

1.38 **CALCULATE** the new volumetric flow rate, head, or power for a variable speed centrifugal pump using the pump laws.

1.39 **DESCRIBE** the effect on system flow and pump head for the following changes:
   a. Changing pump speeds
   b. Adding pumps in parallel
   c. Adding pumps in series

Intentionally Left Blank

CONTINUITY EQUATION

Understanding the quantities measured by the volumetric flow rate and mass flow rate is crucial to understanding other fluid flow topics. The continuity equation expresses the relationship between mass flow rates at different points in a fluid system under steady-state flow conditions.

EO 1.1 **DESCRIBE how the density of a fluid varies with temperature.**

EO 1.2 **DEFINE the term buoyancy.**

EO 1.3 **DESCRIBE the relationship between the pressure in a fluid column and the density and depth of the fluid.**

EO 1.4 **STATE Pascal's Law.**

EO 1.5 **DEFINE the terms mass flow rate and volumetric flow rate.**

EO 1.6 **CALCULATE either the mass flow rate or the volumetric flow rate for a fluid system.**

EO 1.7 **STATE the principle of conservation of mass.**

EO 1.8 **CALCULATE the fluid velocity or flow rate in a specified fluid system using the continuity equation.**

Introduction

Fluid flow is an important part of most industrial processes; especially those involving the transfer of heat. Frequently, when it is desired to remove heat from the point at which it is generated, some type of fluid is involved in the heat transfer process. Examples of this are the cooling water circulated through a gasoline or diesel engine, the air flow past the windings of a motor, and the flow of water through the core of a nuclear reactor. Fluid flow systems are also commonly used to provide lubrication.

Fluid flow in the nuclear field can be complex and is not always subject to rigorous mathematical analysis. Unlike solids, the particles of fluids move through piping and components at different velocities and are often subjected to different accelerations.

Even though a detailed analysis of fluid flow can be extremely difficult, the basic concepts involved in fluid flow problems are fairly straightforward. These basic concepts can be applied in solving fluid flow problems through the use of simplifying assumptions and average values, where appropriate. Even though this type of analysis would not be sufficient in the engineering design of systems, it is very useful in understanding the operation of systems and predicting the approximate response of fluid systems to changes in operating parameters.

The basic principles of fluid flow include three concepts or principles; the first two of which the student has been exposed to in previous manuals. The first is the principle of momentum (leading to equations of fluid forces) which was covered in the manual on Classical Physics. The second is the conservation of energy (leading to the First Law of Thermodynamics) which was studied in thermodynamics. The third is the conservation of mass (leading to the continuity equation) which will be explained in this module.

Properties of Fluids

A *fluid* is any substance which flows because its particles are not rigidly attached to one another. This includes liquids, gases and even some materials which are normally considered solids, such as glass. Essentially, fluids are materials which have no repeating crystalline structure.

Several properties of fluids were discussed in the Thermodynamics section of this text. These included temperature, pressure, mass, specific volume and density. *Temperature* was defined as the relative measure of how hot or cold a material is. It can be used to predict the direction that heat will be transferred. *Pressure* was defined as the force per unit area. Common units for pressure are pounds force per square inch (psi). *Mass* was defined as the quantity of matter contained in a body and is to be distinguished from weight, which is measured by the pull of gravity on a body. The *specific volume* of a substance is the volume per unit mass of the substance. Typical units are ft^3/lbm. *Density*, on the other hand, is the mass of a substance per unit volume. Typical units are lbm/ft^3. Density and specific volume are the inverse of one another. Both density and specific volume are dependant on the temperature and somewhat on the pressure of the fluid. As the temperature of the fluid increases, the density decreases and the specific volume increases. Since liquids are considered incompressible, an increase in pressure will result in no change in density or specific volume of the liquid. In actuality, liquids can be slightly compressed at high pressures, resulting in a slight increase in density and a slight decrease in specific volume of the liquid.

Buoyancy

Buoyancy is defined as the tendency of a body to float or rise when submerged in a fluid. We all have had numerous opportunities of observing the buoyant effects of a liquid. When we go swimming, our bodies are held up almost entirely by the water. Wood, ice, and cork float on water. When we lift a rock from a stream bed, it suddenly seems heavier on emerging from the water. Boats rely on this buoyant force to stay afloat. The amount of this buoyant effect was first computed and stated by the Greek philosopher Archimedes. When a body is placed in a fluid, it is buoyed up by a force equal to the weight of the water that it displaces.

If a body weighs more than the liquid it displaces, it sinks but will appear to lose an amount of weight equal to that of the displaced liquid, as our rock. If the body weighs less than that of the displaced liquid, the body will rise to the surface eventually floating at such a depth that will displace a volume of liquid whose weight will just equal its own weight. A floating body displaces its own weight of the fluid in which it floats.

Compressibility

Compressibility is the measure of the change in volume a substance undergoes when a pressure is exerted on the substance. Liquids are generally considered to be incompressible. For instance, a pressure of 16,400 psig will cause a given volume of water to decrease by only 5% from its volume at atmospheric pressure. Gases on the other hand, are very compressible. The volume of a gas can be readily changed by exerting an external pressure on the gas

Relationship Between Depth and Pressure

Anyone who dives under the surface of the water notices that the pressure on his eardrums at a depth of even a few feet is noticeably greater than atmospheric pressure. Careful measurements show that the pressure of a liquid is directly proportional to the depth, and for a given depth the liquid exerts the same pressure in all directions.

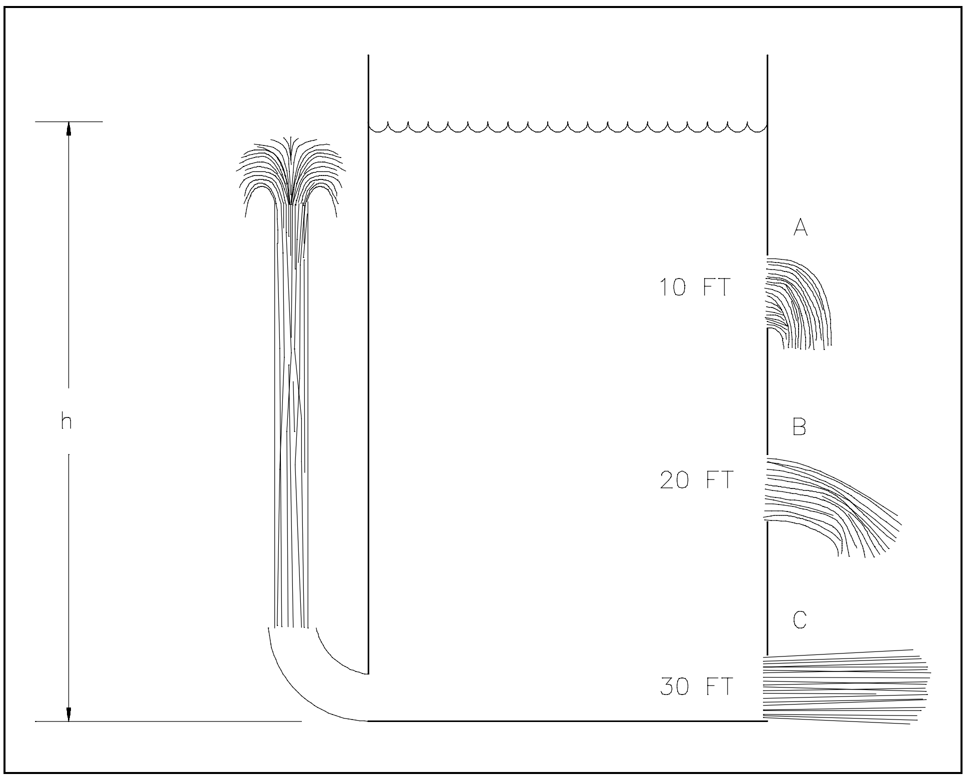


Figure 1 Pressure Versus Depth

As shown in Figure 1 the pressure at different levels in the tank varies and this causes the fluid to leave the tank at varying velocities. Pressure was defined to be force per unit area. In the case of this tank, the force is due to the weight of the water above the point where the pressure is being determined.

Example:

$$\text{Pressure} = \frac{\text{Force}}{\text{Area}}$$

$$= \frac{\text{Weight}}{\text{Area}}$$

$$P = \frac{m\ g}{A\ g_c}$$

$$= \frac{\rho\ V\ g}{A\ g_c}$$

where:

m = mass in lbm

g = acceleration due to earth's gravity $32.17 \frac{\text{ft}}{\text{sec}^2}$

g_c = $32.17\ \frac{\text{lbm-ft}}{\text{lbf-sec}^2}$

A = area in ft^2

V = volume in ft^3

ρ = density of fluid in $\frac{\text{lbm}}{\text{ft}^3}$

The volume is equal to the cross-sectional area times the height (h) of liquid. Substituting this in to the above equation yields:

$$P = \frac{\rho\ A\ h\ g}{A\ g_c}$$

$$P = \frac{\rho\ h\ g}{g_c}$$

This equation tells us that the pressure exerted by a column of water is directly proportional to the height of the column and the density of the water and is independent of the cross-sectional area of the column. The pressure thirty feet below the surface of a one inch diameter standpipe is the same as the pressure thirty feet below the surface of a large lake.

Example 1:

If the tank in Figure 1 is filled with water that has a density of 62.4 lbm/ft^3, calculate the pressures at depths of 10, 20, and 30 feet.

Solution:

$$P = \frac{\rho\ h\ g}{g_c}$$

$$P_{10\ \text{feet}} = \left(62.4\ \frac{\text{lbm}}{\text{ft}^3}\right)(10\ \text{ft})\left(\frac{32.17\ \frac{\text{ft}}{\text{sec}^2}}{32.17\ \frac{\text{lbm-ft}}{\text{lbf-sec}^2}}\right)$$

$$= 624\ \frac{\text{lbf}}{\text{ft}^2}\left(\frac{1\ \text{ft}^2}{144\ \text{in}^2}\right)$$

$$= 4.33\ \frac{\text{lbf}}{\text{in}^2}$$

$$P_{20\ \text{feet}} = \left(62.4\ \frac{\text{lbm}}{\text{ft}^3}\right)(20\ \text{ft})\left(\frac{32.17\ \frac{\text{ft}}{\text{sec}^2}}{32.17\ \frac{\text{lbm-ft}}{\text{lbf-sec}^2}}\right)$$

$$= 1248\ \frac{\text{lbf}}{\text{ft}^2}\left(\frac{1\ \text{ft}^2}{144\ \text{in}^2}\right)$$

$$= 8.67\ \frac{\text{lbf}}{\text{in}^2}$$

$$P_{30\ feet} = \left(62.4\ \frac{lbm}{ft^3}\right)(30\ ft)\left(\frac{32.17\ \frac{ft}{sec^2}}{32.17\ \frac{lbm-ft}{lbf-sec^2}}\right)$$

$$= 1872\ \frac{lbf}{ft^2}\left(\frac{1\ ft^2}{144\ in^2}\right)$$

$$= 13.00\ \frac{lbf}{in^2}$$

Example 2:

A cylindrical water tank 40 ft high and 20 ft in diameter is filled with water that has a density of 61.9 lbm/ft^3.

(a) What is the water pressure on the bottom of the tank?

(b) What is the average force on the bottom?

Solution:

(a)

$$P = \frac{\rho\ h\ g}{g_c}$$

$$P = \left(61.9\ \frac{lbm}{ft^3}\right)(40\ ft)\left(\frac{32.17\ \frac{ft}{sec^2}}{32.17\ \frac{lbm-ft}{lbf-sec^2}}\right)$$

$$= 2476\ \frac{lbf}{ft^2}\left(\frac{1\ ft^2}{144\ in^2}\right)$$

$$= 17.2\ \frac{lbf}{in^2}$$

(b)

$$\text{Pressure} = \frac{\text{Force}}{\text{Area}}$$

$$\text{Force} = (\text{Pressure})\ (\text{Area})$$

$$\text{Area} = \pi r^2$$

$$F = \left(17.2\ \frac{\text{lbf}}{\text{in}^2}\right) \pi\ (10\ \text{ft})^2 \left(\frac{144\ \text{in}^2}{1\ \text{ft}^2}\right)$$

$$= 7.78\ \text{x}\ 10^5\ \text{lbf}$$

Pascal's Law

The pressure of the liquids in each of the previously cited cases has been due to the weight of the liquid. Liquid pressures may also result from application of external forces on the liquid. Consider the following examples. Figure 2 represents a container completely filled with liquid. A, B, C, D, and E represent pistons of equal cross-sectional areas fitted into the walls of the vessel. There will be forces acting on the pistons C, D, and E due to the pressures caused by the different depths of the liquid. Assume that the forces on the pistons due to the pressure caused by the weight of the liquid are as follows: A = 0 lbf, B = 0 lbf, C = 10 lbf, D = 30 lbf, and E = 25 lbf. Now let an external force of 50 lbf be applied to piston A. This external force will cause the pressure at all points in the container to increase by the same amount. Since the pistons all have the same cross-sectional area, the increase in pressure will result in the forces on the pistons all increasing by 50 lbf. So if an external force of 50 lbf is applied to piston A, the force exerted by the fluid on the other pistons will now be as follows: B = 50 lbf, C = 60 lbf, D = 80 lbf, and E = 75 lbf.

This effect of an external force on a confined fluid was first stated by Pascal in 1653.

> *Pressure applied to a confined fluid is transmitted undiminished throughout the confining vessel of the system.*

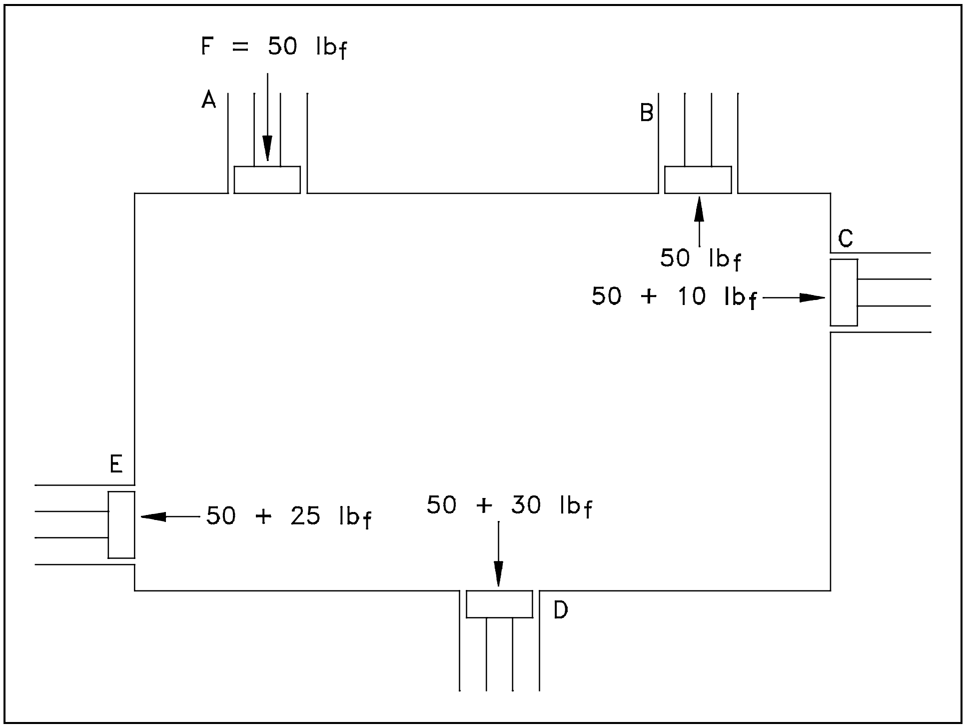


Figure 2 Pascal's Law

Control Volume

In thermodynamics, a *control volume* was defined as a fixed region in space where one studies the masses and energies crossing the boundaries of the region. This concept of a control volume is also very useful in analyzing fluid flow problems. The boundary of a control volume for fluid flow is usually taken as the physical boundary of the part through which the flow is occurring. The control volume concept is used in fluid dynamics applications, utilizing the continuity, momentum, and energy principles mentioned at the beginning of this chapter. Once the control volume and its boundary are established, the various forms of energy crossing the boundary with the fluid can be dealt with in equation form to solve the fluid problem. Since fluid flow problems usually treat a fluid crossing the boundaries of a control volume, the control volume approach is referred to as an "open" system analysis, which is similar to the concepts studied in thermodynamics. There are special cases in the nuclear field where fluid does not cross the control boundary. Such cases are studied utilizing the "closed" system approach.

Regardless of the nature of the flow, all flow situations are found to be subject to the established basic laws of nature that engineers have expressed in equation form. Conservation of mass and conservation of energy are always satisfied in fluid problems, along with Newton's laws of motion. In addition, each problem will have physical constraints, referred to mathematically as boundary conditions, that must be satisfied before a solution to the problem will be consistent with the physical results.

Volumetric Flow Rate

The *volumetric flow rate* ($\dot{V}$) of a system is a measure of the volume of fluid passing a point in the system per unit time. The volumetric flow rate can be calculated as the product of the cross-sectional area (A) for flow and the average flow velocity (v).

$$\dot{V} = A\ v \tag{3-1}$$

If area is measured in square feet and velocity in feet per second, Equation 3-1 results in volumetric flow rate measured in cubic feet per second. Other common units for volumetric flow rate include gallons per minute, cubic centimeters per second, liters per minute, and gallons per hour.

Example:

A pipe with an inner diameter of 4 inches contains water that flows at an average velocity of 14 feet per second. Calculate the volumetric flow rate of water in the pipe.

Solution:

Use Equation 3-1 and substitute for the area.

$$\dot{V} = (\pi\ r^2)\ v$$

$$\dot{V} = (3.14)\ \left(\frac{2}{12}\ \text{ft}\right)^2\ \left(14\ \frac{\text{ft}}{\text{sec}}\right)$$

$$\dot{V} = 1.22\ \frac{\text{ft}^3}{\text{sec}}$$

Mass Flow Rate

The *mass flow rate* ($\dot{m}$) of a system is a measure of the mass of fluid passing a point in the system per unit time. The mass flow rate is related to the volumetric flow rate as shown in Equation 3-2 where ρ is the density of the fluid.

$$\dot{m} = \rho\dot{V} \tag{3-2}$$

If the volumetric flow rate is in cubic feet per second and the density is in pounds-mass per cubic foot, Equation 3-2 results in mass flow rate measured in pounds-mass per second. Other common units for measurement of mass flow rate include kilograms per second and pounds-mass per hour.

Replacing $\dot{V}$ in Equation 3-2 with the appropriate terms from Equation 3-1 allows the direct calculation of the mass flow rate.

$$\dot{m} = \rho \ A \ v \tag{3-3}$$

Example:

The water in the pipe of the previous example had a density of 62.44 lbm/ft^3. Calculate the mass flow rate.

Solution:

$$\dot{m} = \rho \ \dot{V}$$

$$\dot{m} = (62.44 \ \frac{\text{lbm}}{\text{ft}^3}) \ (1.22 \ \frac{\text{ft}^3}{\text{sec}})$$

$$\dot{m} = 76.2 \ \frac{\text{lbm}}{\text{sec}}$$

Conservation of Mass

In thermodynamics, you learned that energy can neither be created nor destroyed, only changed in form. The same is true for mass. Conservation of mass is a principle of engineering that states that all mass flow rates into a control volume are equal to all mass flow rates out of the control volume plus the rate of change of mass within the control volume. This principle is expressed mathematically by Equation 3-4.

$$\dot{m}_{in} = \dot{m}_{out} + \frac{\Delta m}{\Delta t} \tag{3-4}$$

where:

$\frac{\Delta m}{\Delta t}$ = the increase or decrease of the mass within the control volume over a (specified time period)

Steady-State Flow

Steady-state flow refers to the condition where the fluid properties at any single point in the system do not change over time. These fluid properties include temperature, pressure, and velocity. One of the most significant properties that is constant in a steady-state flow system is the system mass flow rate. This means that there is no accumulation of mass within any component in the system.

Continuity Equation

The continuity equation is simply a mathematical expression of the principle of conservation of mass. For a control volume that has a single inlet and a single outlet, the principle of conservation of mass states that, for steady-state flow, the mass flow rate into the volume must equal the mass flow rate out. The continuity equation for this situation is expressed by Equation 3-5.

$$\dot{m}_{inlet} = \dot{m}_{outlet} \qquad (3\text{-}5)$$

$$(\rho A v)_{inlet} = (\rho A v)_{outlet}$$

For a control volume with multiple inlets and outlets, the principle of conservation of mass requires that the sum of the mass flow rates into the control volume equal the sum of the mass flow rates out of the control volume. The continuity equation for this more general situation is expressed by Equation 3-6.

$$\Sigma\ \dot{m}_{inlets} = \Sigma\ \dot{m}_{outlets} \qquad (3\text{-}6)$$

One of the simplest applications of the continuity equation is determining the change in fluid velocity due to an expansion or contraction in the diameter of a pipe.

Example: Continuity Equation - Piping Expansion

Steady-state flow exists in a pipe that undergoes a gradual expansion from a diameter of 6 in. to a diameter of 8 in. The density of the fluid in the pipe is constant at 60.8 lbm/ft^3. If the flow velocity is 22.4 ft/sec in the 6 in. section, what is the flow velocity in the 8 in. section?

Solution:

From the continuity equation we know that the mass flow rate in the 6 in. section must equal the mass flow rate in the 8 in. section. Letting the subscript 1 represent the 6 in. section and 2 represent the 8 in. section we have the following.

$$\dot{m}_1 = \dot{m}_2$$

$$\rho_1 A_1 v_1 = \rho_2 A_2 v_2$$

$$v_2 = v_1 \frac{\rho_1}{\rho_2} \frac{A_1}{A_2}$$

$$= v_1 \frac{\pi r_1^2}{\pi r_2^2}$$

$$= \left(22.4 \frac{ft}{sec}\right) \frac{(3 \text{ in})^2}{(4 \text{ in})^2}$$

$$v_2 = 12.6 \frac{ft}{sec}$$

So by using the continuity equation, we find that the increase in pipe diameter from 6 to 8 inches caused a decrease in flow velocity from 22.4 to 12.6 ft/sec.

The continuity equation can also be used to show that a decrease in pipe diameter will cause an increase in flow velocity.

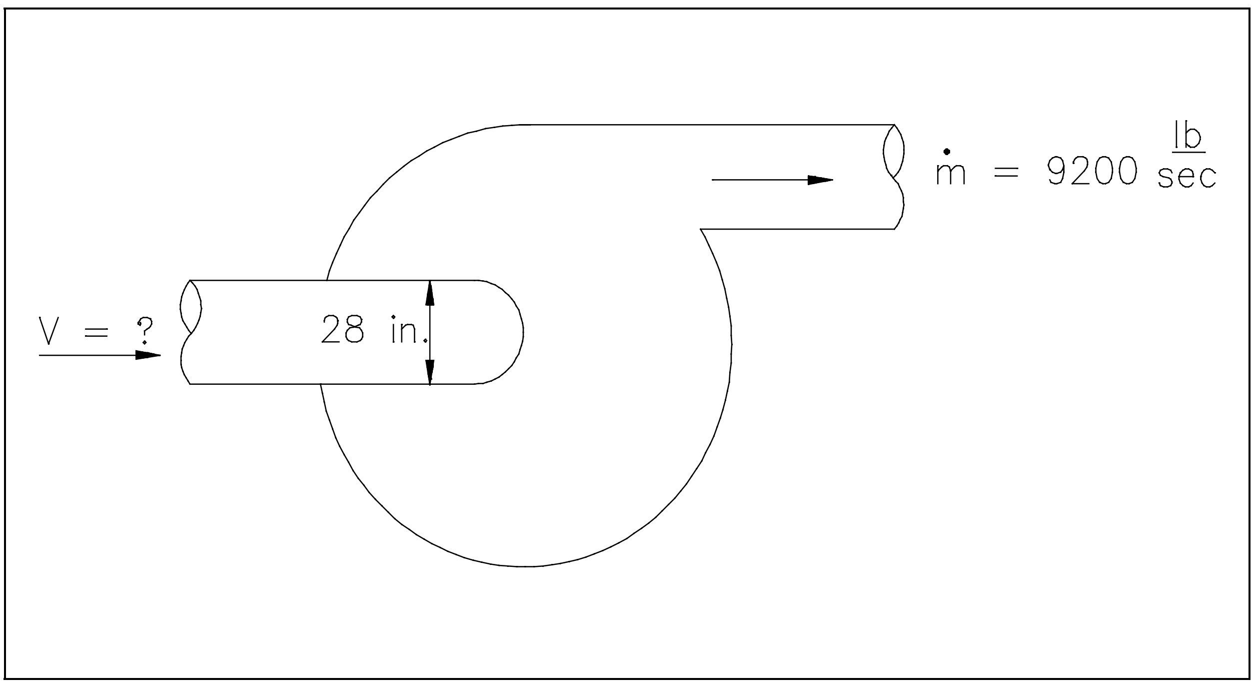


Figure 3 Continuity Equation

Example: Continuity Equation - Centrifugal Pump

The inlet diameter of the reactor coolant pump shown in Figure 3 is 28 in. while the outlet flow through the pump is 9200 lbm/sec. The density of the water is 49 lbm/ft^3. What is the velocity at the pump inlet?

Solution:

$$A_{inlet} = \pi r^2 = (3.14)\left(14\ \text{in}\left(\frac{1\ \text{ft}}{12\ \text{in}}\right)\right)^2$$

$$= 4.28\ \text{ft}^2$$

$$\dot{m}_{inlet} = \dot{m}_{outlet} = 9200\ \frac{\text{lbm}}{\text{sec}}$$

$$(\rho A v)_{inlet} = 9200\ \frac{\text{lbm}}{\text{sec}}$$

$$v_{inlet} = \frac{9200\ \frac{\text{lbm}}{\text{sec}}}{A\rho}$$

$$= \frac{9200\ \frac{\text{lbm}}{\text{sec}}}{(4.28\ \text{ft}^2)\left(49\ \frac{\text{lbm}}{\text{ft}^3}\right)}$$

$$v_{inlet} = 43.9\ \frac{\text{ft}}{\text{sec}}$$

The above example indicates that the flow rate into the system is the same as that out of the system. The same concept is true even though more than one flow path may enter or leave the system at the same time. The mass balance is simply adjusted to state that the sum of all flows entering the system is equal to the sum of all the flows leaving the system if steady-state conditions exist. An example of this physical case is included in the following example.

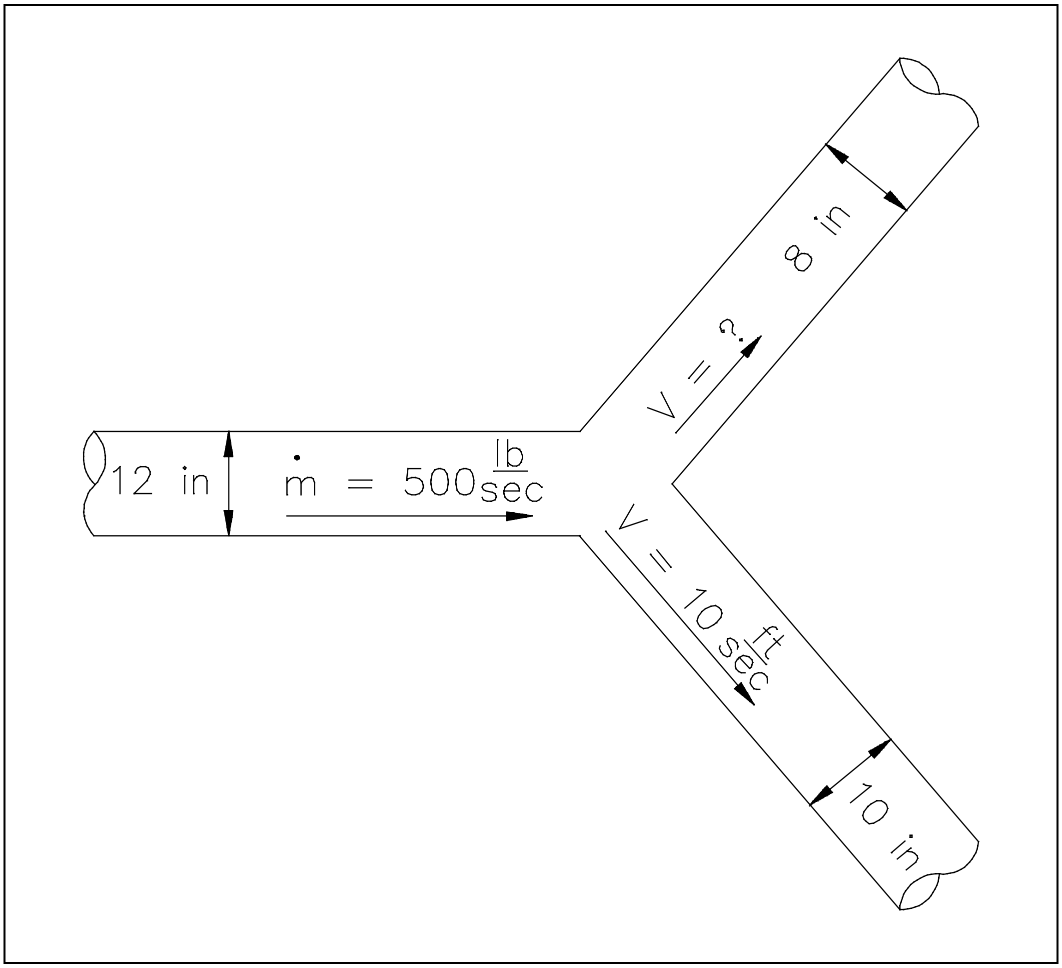


Figure 4 "Y" Configuration for Example Problem

Example: Continuity Equation - Multiple Outlets

A piping system has a "Y" configuration for separating the flow as shown in Figure 4. The diameter of the inlet leg is 12 in., and the diameters of the outlet legs are 8 and 10 in. The velocity in the 10 in. leg is 10 ft/sec. The flow through the main portion is 500 lbm/sec. The density of water is 62.4 lbm/ft^3. What is the velocity out of the 8 in. pipe section?

Solution:

$$A_8 = \pi \left(4 \text{ in.} \ \frac{1 \text{ ft}}{12 \text{ in.}}\right)^2$$

$$= 0.349 \text{ ft}^2$$

$$A_{10} = \pi \left(5 \text{ in.} \ \frac{1 \text{ ft}}{12 \text{ in.}}\right)^2$$

$$= 0.545 \text{ ft}^2$$

$$\Sigma\dot{m}_{\text{inlets}} = \Sigma\dot{m}_{\text{outlets}}$$

$$\dot{m}_{12} = \dot{m}_{10} + \dot{m}_8$$

$$\dot{m}_8 = \dot{m}_{12} - \dot{m}_{10}$$

$$(\rho A v)_8 = \dot{m}_{12} - (\rho A v)_{10}$$

$$v_8 = \frac{\dot{m}_{12} - (\rho A v)_{10}}{(\rho A)_8}$$

$$= \frac{500 \ \frac{\text{lbm}}{\text{sec}} - \left(62.4 \ \frac{\text{lbm}}{\text{ft}^3}\right)(0.545 \text{ ft}^2)\left(10 \ \frac{\text{ft}}{\text{sec}}\right)}{\left(62.4 \ \frac{\text{lbm}}{\text{ft}^3}\right)(0.349 \text{ ft}^2)}$$

$$v_8 = 7.3 \ \frac{\text{ft}}{\text{sec}}$$

Summary

The main points of this chapter are summarized on the next page.

Continuity Equation Summary

- Density changes in a fluid are inversely proportional to temperature changes.
- Buoyancy is the tendency of a body to float or rise when submerged in a fluid.
- The pressure exerted by a column of water is directly proportional to the height of the column and the density of the water.

$$P = \frac{\rho \, h \, g}{g_c}$$

- Pascal's law states that pressure applied to a confined fluid is transmitted undiminished throughout the confining vessel of a system.
- Volumetric flow rate is the volume of fluid per unit time passing a point in a fluid system.
- Mass flow rate is the mass of fluid per unit time passing a point in a fluid system.
- The volumetric flow rate is calculated by the product of the average fluid velocity and the cross-sectional area for flow.

$$\dot{V} = A \, v$$

- The mass flow rate is calculated by the product of the volumetric flow rate and the fluid density.

$$\dot{m} = \rho \, A \, v$$

- The principle of conservation of mass states that all mass flow rates into a control volume are equal to all mass flow rates out of the control volume plus the rate of change of mass within the control volume.
- For a control volume with a single inlet and outlet, the continuity equation can be expressed as follows:

$$\dot{m}_{inlet} = \dot{m}_{outlet}$$

- For a control volume with multiple inlets and outlets, the continuity equation is:

$$\Sigma \, \dot{m}_{inlets} = \Sigma \, \dot{m}_{outlets}$$

LAMINAR AND TURBULENT FLOW

The characteristics of laminar and turbulent flow are very different. To understand why turbulent or laminar flow is desirable in the operation of a particular system, it is necessary to understand the characteristics of laminar and turbulent flow.

EO 1.9 DESCRIBE the characteristics and flow velocity profiles of laminar flow and turbulent flow.

EO 1.10 DEFINE the property of viscosity.

EO 1.11 DESCRIBE how the viscosity of a fluid varies with temperature.

EO 1.12 DESCRIBE the characteristics of an ideal fluid.

EO 1.13 DESCRIBE the relationship between the Reynolds number and the degree of turbulence of the flow.

Flow Regimes

All fluid flow is classified into one of two broad categories or regimes. These two flow regimes are laminar flow and turbulent flow. The flow regime, whether laminar or turbulent, is important in the design and operation of any fluid system. The amount of fluid friction, which determines the amount of energy required to maintain the desired flow, depends upon the mode of flow. This is also an important consideration in certain applications that involve heat transfer to the fluid.

Laminar Flow

Laminar flow is also referred to as streamline or viscous flow. These terms are descriptive of the flow because, in laminar flow, (1) layers of water flowing over one another at different speeds with virtually no mixing between layers, (2) fluid particles move in definite and observable paths or streamlines, and (3) the flow is characteristic of viscous (thick) fluid or is one in which viscosity of the fluid plays a significant part.

Turbulent Flow

Turbulent flow is characterized by the irregular movement of particles of the fluid. There is no definite frequency as there is in wave motion. The particles travel in irregular paths with no observable pattern and no definite layers.

Flow Velocity Profiles

Not all fluid particles travel at the same velocity within a pipe. The shape of the velocity curve (the velocity profile across any given section of the pipe) depends upon whether the flow is laminar or turbulent. If the flow in a pipe is laminar, the velocity distribution at a cross section will be parabolic in shape with the maximum velocity at the center being about twice the average velocity in the pipe. In turbulent flow, a fairly flat velocity distribution exists across the section of pipe, with the result that the entire fluid flows at a given single value. Figure 5 helps illustrate the above ideas. The velocity of the fluid in contact with the pipe wall is essentially zero and increases the further away from the wall.

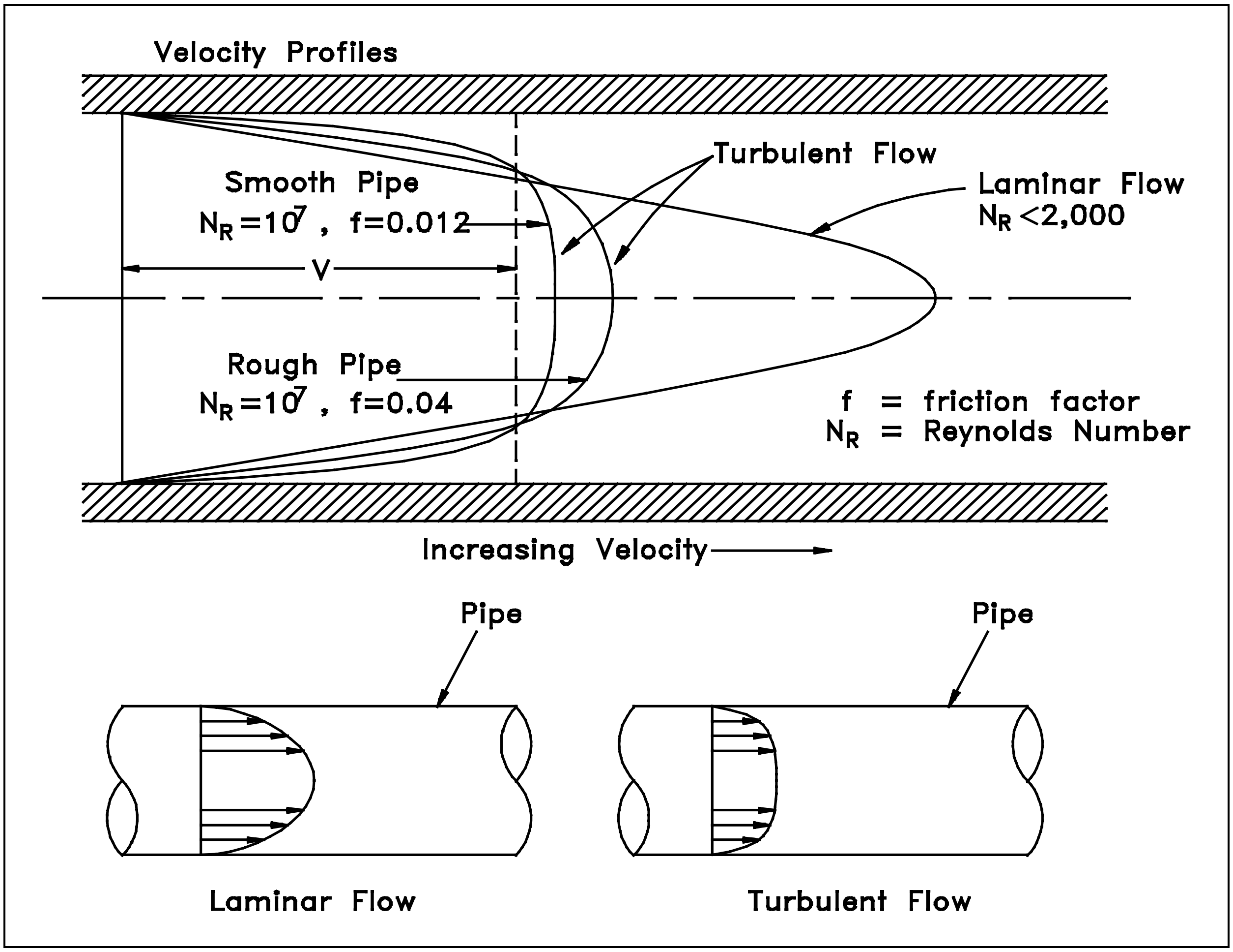


Figure 5 Laminar and Turbulent Flow Velocity Profiles

Note from Figure 5 that the velocity profile depends upon the surface condition of the pipe wall. A smoother wall results in a more uniform velocity profile than a rough pipe wall.

Average (Bulk) Velocity

In many fluid flow problems, instead of determining exact velocities at different locations in the same flow cross-section, it is sufficient to allow a single average velocity to represent the velocity of all fluid at that point in the pipe. This is fairly simple for turbulent flow since the velocity profile is flat over the majority of the pipe cross-section. It is reasonable to assume that the average velocity is the same as the velocity at the center of the pipe.

If the flow regime is laminar (the velocity profile is parabolic), the problem still exists of trying to represent the "average" velocity at any given cross-section since an average value is used in the fluid flow equations. Technically, this is done by means of integral calculus. Practically, the student should use an average value that is half of the center line value.

Viscosity

Viscosity is a fluid property that measures the resistance of the fluid to deforming due to a shear force. Viscosity is the internal friction of a fluid which makes it resist flowing past a solid surface or other layers of the fluid. Viscosity can also be considered to be a measure of the resistance of a fluid to flowing. A thick oil has a high viscosity; water has a low viscosity. The unit of measurement for absolute viscosity is:

μ = absolute viscosity of fluid (lbf-sec/ft^2).

The viscosity of a fluid is usually significantly dependent on the temperature of the fluid and relatively independent of the pressure. For most fluids, as the temperature of the fluid increases, the viscosity of the fluid decreases. An example of this can be seen in the lubricating oil of engines. When the engine and its lubricating oil are cold, the oil is very viscous, or thick. After the engine is started and the lubricating oil increases in temperature, the viscosity of the oil decreases significantly and the oil seems much thinner.

Ideal Fluid

An *ideal fluid* is one that is incompressible and has no viscosity. Ideal fluids do not actually exist, but sometimes it is useful to consider what would happen to an ideal fluid in a particular fluid flow problem in order to simplify the problem.

Reynolds Number

The flow regime (either laminar or turbulent) is determined by evaluating the Reynolds number of the flow (refer to figure 5). The *Reynolds number*, based on studies of Osborn Reynolds, is a dimensionless number comprised of the physical characteristics of the flow. Equation 3-7 is used to calculate the Reynolds number (N_R) for fluid flow.

$$N_R = \rho \ v \ D \ / \ \mu \ g_c \qquad (3\text{-}7)$$

where:

N_R	=	Reynolds number (unitless)
v	=	average velocity (ft/sec)
D	=	diameter of pipe (ft)
μ	=	absolute viscosity of fluid (lbf-sec/ft^2)
ρ	=	fluid mass density (lbm/ft^3)
g_c	=	gravitational constant (32.2 ft-lbm/lbf-sec^2)

For practical purposes, if the Reynolds number is less than 2000, the flow is laminar. If it is greater than 3500, the flow is turbulent. Flows with Reynolds numbers between 2000 and 3500 are sometimes referred to as transitional flows. Most fluid systems in nuclear facilities operate with turbulent flow. Reynolds numbers can be conveniently determined using a Moody Chart; an example of which is shown in Appendix B. Additional detail on the use of the Moody Chart is provided in subsequent text.

Summary

The main points of this chapter are summarized below.

Laminar and Turbulent Flow Summary

- Laminar Flow
  Layers of water flow over one another at different speeds with virtually no mixing between layers.
  The flow velocity profile for laminar flow in circular pipes is parabolic in shape, with a maximum flow in the center of the pipe and a minimum flow at the pipe walls.
  The average flow velocity is approximately one half of the maximum velocity.

- Turbulent Flow
  The flow is characterized by the irregular movement of particles of the fluid.
  The flow velocity profile for turbulent flow is fairly flat across the center section of a pipe and drops rapidly extremely close to the walls.
  The average flow velocity is approximately equal to the velocity at the center of the pipe.

- Viscosity is the fluid property that measures the resistance of the fluid to deforming due to a shear force. For most fluids, temperature and viscosity are inversely proportional.

- An ideal fluid is one that is incompressible and has no viscosity.

- An increasing Reynolds number indicates an increasing turbulence of flow.

BERNOULLI'S EQUATION

Bernoulli's equation is a special case of the general energy equation that is probably the most widely-used tool for solving fluid flow problems. It provides an easy way to relate the elevation head, velocity head, and pressure head of a fluid. It is possible to modify Bernoulli's equation in a manner that accounts for head losses and pump work.

EO 1.14 DESCRIBE the relationship between Bernoulli's equation and the First Law of Thermodynamics.

EO 1.15 DEFINE the term head with respect to its use in fluid flow.

EO 1.16 EXPLAIN the energy conversions that take place in a fluid system between the velocity, elevation, and pressure heads as flow continues through a piping system.

EO 1.17 Given the initial and final conditions of the system, CALCULATE the unknown fluid properties using the simplified Bernoulli equation.

EO 1.18 DESCRIBE the restrictions applied to Bernoulli's equation when presented in its simplest form.

EO 1.19 EXPLAIN how to extend the Bernoulli equation to more general applications.

EO 1.20 RELATE Bernoulli's principle to the operation of a venturi.

General Energy Equation

The conservation of energy principle states that energy can be neither created nor destroyed. This is equivalent to the First Law of Thermodynamics, which was used to develop the general energy equation in the module on thermodynamics. Equation 3-8 is a statement of the general energy equation for an open system.

$$Q + (U + PE + KE + PV)_{in} =$$

$$W + (U + PE + KE + PV)_{out} + (U + PE + KE + PV)_{stored} \qquad (3\text{-}8)$$

where:

Q	=	heat (Btu)
U	=	internal energy (Btu)
PE	=	potential energy (ft-lbf)
KE	=	kinetic energy (ft-lbf)
P	=	pressure (lbf/ft^2)
V	=	volume (ft^3)
W	=	work (ft-lbf)

Simplified Bernoulli Equation

Bernoulli's equation results from the application of the general energy equation and the first law of thermodynamics to a steady flow system in which no work is done on or by the fluid, no heat is transferred to or from the fluid, and no change occurs in the internal energy (i.e., no temperature change) of the fluid. Under these conditions, the general energy equation is simplified to Equation 3-9.

$$(PE + KE + PV)_1 = (PE + KE + PV)_2 \qquad (3\text{-}9)$$

Substituting appropriate expressions for the potential energy and kinetic energy, Equation 3-9 can be rewritten as Equation 3-10.

$$\frac{mgz_1}{g_c} + \frac{mv_1^{\,2}}{2g_c} + P_1V_1 = \frac{mgz_2}{g_c} + \frac{mv_2^{\,2}}{2g_c} + P_2V_2 \qquad (3\text{-}10)$$

where:

m	=	mass (lbm)
z	=	height above reference (ft)
v	=	average velocity (ft/sec)
g	=	acceleration due to gravity (32.17 ft/sec^2)
g_c	=	gravitational constant, (32.17 $ft\text{-}lbm/lbf\text{-}sec^2$)

Note: The factor g_c is only required when the English System of measurement is used and mass is measured in pound mass. It is essentially a conversion factor needed to allow the units to come out directly. No factor is necessary if mass is measured in slugs or if the metric system of measurement is used.

Each term in Equation 3-10 represents a form of energy possessed by a moving fluid (potential, kinetic, and pressure related energies). In essence, the equation physically represents a balance of the KE, PE, PV energies so that if one form of energy increases, one or more of the others will decrease to compensate and vice versa.

Multiplying all terms in Equation 3-10 by the factor g_c/mg results in the form of Bernoulli's equation shown by Equation 3-11.

$$z_1 + \frac{v_1^{\,2}}{2g} + P_1\nu_1\,\frac{g_c}{g} = z_2 + \frac{v_2^{\,2}}{2g} + P_2\nu_2\,\frac{g_c}{g} \qquad (3\text{-}11)$$

Head

Since the units for all the different forms of energy in Equation 3-11 are measured in units of distance, these terms are sometimes referred to as "heads" (pressure head, velocity head, and elevation head). The term *head* is used by engineers in reference to pressure. It is a reference to the height, typically in feet, of a column of water that a given pressure will support. Each of the energies possessed by a fluid can be expressed in terms of head. The elevation head represents the potential energy of a fluid due to its elevation above a reference level. The velocity head represents the kinetic energy of the fluid. It is the height in feet that a flowing fluid would rise in a column if all of its kinetic energy were converted to potential energy. The pressure head represents the flow energy of a column of fluid whose weight is equivalent to the pressure of the fluid.

The sum of the elevation head, velocity head, and pressure head of a fluid is called the total head. Thus, Bernoulli's equation states that the total head of the fluid is constant.

Energy Conversions in Fluid Systems

Bernoulli's equation makes it easy to examine how energy transfers take place among elevation head, velocity head, and pressure head. It is possible to examine individual components of piping systems and determine what fluid properties are varying and how the energy balance is affected.

If a pipe containing an ideal fluid undergoes a gradual expansion in diameter, the continuity equation tells us that as the diameter and flow area get bigger, the flow velocity must decrease to maintain the same mass flow rate. Since the outlet velocity is less than the inlet velocity, the velocity head of the flow must decrease from the inlet to the outlet. If the pipe lies horizontal, there is no change in elevation head; therefore, the decrease in velocity head must be compensated for by an increase in pressure head. Since we are considering an ideal fluid that is incompressible, the specific volume of the fluid will not change. The only way that the pressure head for an incompressible fluid can increase is for the pressure to increase. So the Bernoulli equation indicates that a decrease in flow velocity in a horizontal pipe will result in an increase in pressure.

If a constant diameter pipe containing an ideal fluid undergoes a decrease in elevation, the same net effect results, but for different reasons. In this case the flow velocity and the velocity head must be constant to satisfy the mass continuity equation.

So the decrease in elevation head can only be compensated for by an increase in pressure head. Again, the fluid is incompressible so the increase in pressure head must result in an increase in pressure.

Although the Bernoulli equation has several restrictions placed upon it, there are many physical fluid problems to which it is applied. As in the case of the conservation of mass, the Bernoulli equation may be applied to problems in which more than one flow may enter or leave the system at the same time. Of particular note is the fact that series and parallel piping system problems are solved using the Bernoulli equation.

Example: Bernoulli's Equation

Assume frictionless flow in a long, horizontal, conical pipe. The diameter is 2.0 ft at one end and 4.0 ft at the other. The pressure head at the smaller end is 16 ft of water. If water flows through this cone at a rate of 125.6 ft^3/sec, find the velocities at the two ends and the pressure head at the larger end.

Solution:

$$\dot{V}_1 = A_1 v_1$$

$$v_1 = \frac{\dot{V}_1}{A_1} \qquad v_2 = \frac{\dot{V}_2}{A_2}$$

$$v_1 = \frac{125.6 \ \frac{ft^3}{sec}}{\pi(1 \ ft)^2} \qquad v_2 = \frac{125.6 \ \frac{ft^3}{sec}}{\pi(2 \ ft)^2}$$

$$v_1 = 40 \ \frac{ft}{sec} \qquad v_2 = 10 \ \frac{ft}{sec}$$

$$z_1 + \frac{v_1^{\ 2}}{2g} + P_1 \nu_1 \frac{g_c}{g} = z_2 + \frac{v_2^{\ 2}}{2g} + P_2 \nu_2 \frac{g_c}{g}$$

$$P_2 \nu_2 \frac{g_c}{g} = P_1 \nu_1 \frac{g_c}{g} + (z_1 - z_2) + \frac{v_1^{\ 2} - v_2^{\ 2}}{2g}$$

$$= 16 \ ft + 0 \ ft + \frac{\left(40 \ \frac{ft}{sec}\right)^2 - \left(10 \ \frac{ft}{sec}\right)^2}{2\left(32.17 \ \frac{ft-lbm}{lbf-sec^2}\right)}$$

$$= 39.3 \ ft$$

Restrictions on the Simplified Bernoulli Equation

Practical applications of the simplified Bernoulli Equation to real piping systems is not possible due to two restrictions. One serious restriction of the Bernoulli equation in its present form is that no fluid friction is allowed in solving piping problems. Therefore, Equation 3-10 only applies to ideal fluids. However, in reality, the total head possessed by the fluid cannot be transferred completely from one point to another because of friction. Taking these losses of head into account would provide a much more accurate description of what takes place physically. This is especially true because one purpose of a pump in a fluid system is to overcome the losses in pressure due to pipe friction.

The second restriction on Bernoulli's equation is that no work is allowed to be done on or by the fluid. This restriction prevents two points in a fluid stream from being analyzed if a pump exists between the two points. Since most flow systems include pumps, this is a significant limitation. Fortunately, the simplified Bernoulli equation can be modified in a manner that satisfactorily deals with both head losses and pump work.

Extended Bernoulli

The Bernoulli equation can be modified to take into account gains and losses of head. The resulting equation, referred to as the Extended Bernoulli equation, is very useful in solving most fluid flow problems. In fact, the Extended Bernoulli equation is probably used more than any other fluid flow equation. Equation 3-12 is one form of the Extended Bernoulli equation.

$$z_1 + \frac{v_1^2}{2g} + P_1\nu_1 \frac{g_c}{g} + H_p = z_2 + \frac{v_2^2}{2g} + P_2\nu_2 \frac{g_c}{g} + H_f \qquad (3\text{-}12)$$

where:

z = height above reference level (ft)
v = average velocity of fluid (ft/sec)
P = pressure of fluid (lbf/ft^2)
ν = specific volume of fluid (ft^3/lbm)
H_p = head added by pump (ft)
H_f = head loss due to fluid friction (ft)
g = acceleration due to gravity (ft/sec^2)

The head loss due to fluid friction (H_f) represents the energy used in overcoming friction caused by the walls of the pipe. Although it represents a loss of energy from the standpoint of fluid flow, it does not normally represent a significant loss of total energy of the fluid. It also does not violate the law of conservation of energy since the head loss due to friction results in an equivalent increase in the internal energy (u) of the fluid. These losses are greatest as the fluid flows through entrances, exits, pumps, valves, fittings, and any other piping with rough inner surfaces.

Most techniques for evaluating head loss due to friction are empirical (based almost exclusively on experimental evidence) and are based on a proportionality constant called the friction factor (f), which will be discussed in the next section.

Example: Extended Bernoulli

Water is pumped from a large reservoir to a point 65 feet higher than the reservoir. How many feet of head must be added by the pump if 8000 lbm/hr flows through a 6-inch pipe and the frictional head loss is 2 feet? The density of the fluid is 62.4 lbm/ft^3, and the cross-sectional area of a 6-inch pipe is 0.2006 ft^2.

Solution:

To use the modified form of Bernoulli's equation, reference points are chosen at the surface of the reservoir (point 1) and at the outlet of the pipe (point 2). The pressure at the surface of the reservoir is the same as the pressure at the exit of the pipe, i.e., atmospheric pressure. The velocity at point 1 will be essentially zero.

Using the equation for the mass flow rate to determine the velocity at point 2:

$$\dot{m}_2 = \rho A_2 v_2$$

$$v_2 = \frac{\dot{m}_2}{\rho A_2}$$

$$v_2 = \frac{8000 \ \frac{\text{lbm}}{\text{hr}}}{\left(62.4 \ \frac{\text{lbm}}{\text{ft}^3}\right)(0.2006 \ \text{ft}^2)}$$

$$v_2 = 639 \ \frac{\text{ft}}{\text{hr}} \left(\frac{1 \ \text{hr}}{3600 \ \text{sec}}\right)$$

$$v_2 = 0.178 \ \frac{\text{ft}}{\text{sec}}$$

Now we can use the Extended Bernoulli equation to determine the required pump head.

$$z_1 + \frac{v_1^2}{2g} + P_1\nu_1 \frac{g_c}{g} + H_p = z_2 + \frac{v_2^2}{2g} + P_2\nu_2 \frac{g_c}{g} + H_f$$

$$H_p = (z_2 - z_1) + \left(\frac{v_2^2 - v_1^2}{2g}\right) + (P_2 - P_1)\ \nu\ \frac{g_c}{g} + H_f$$

$$= 65\ \text{ft} + \frac{\left(0.178\ \frac{\text{ft}}{\text{sec}}\right)^2 - \left(0\ \frac{\text{ft}}{\text{sec}}\right)^2}{2\left(32.17\ \frac{\text{ft-lbm}}{\text{lbf-sec}^2}\right)} + 0\ \text{ft} + 2\ \text{ft}$$

$$H_p = 67\ \text{ft}$$

The student should note that the solution of this example problem has a numerical value that "makes sense" from the data given in the problem. The total head increase of 67 ft. is due primarily to the 65 ft. evaluation increase and the 2 ft. of friction head.

Application of Bernoulli's Equation to a Venturi

Many plant components, such as a venturi, may be analyzed using Bernoulli's equation and the continuity equation. A venturi is a flow measuring device that consists of a gradual contraction followed by a gradual expansion. An example of a venturi is shown in Figure 6. By measuring the differential pressure between the inlet of the venturi (point 1) and the throat of the venturi (point 2), the flow velocity and mass flow rate can be determined based on Bernoulli's equation.

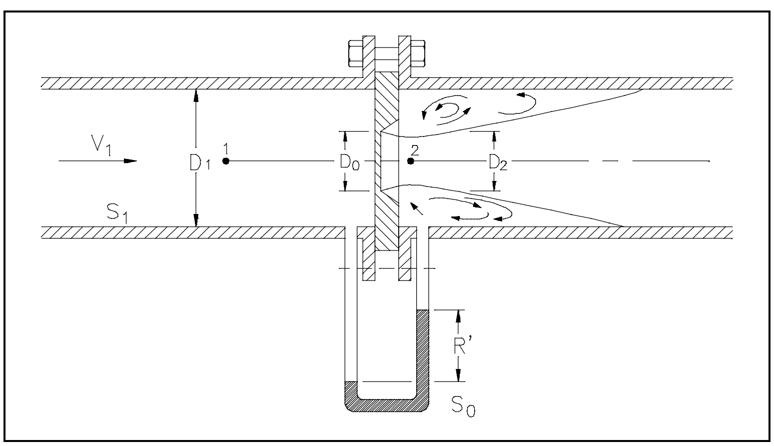


Figure 6 Venturi Meter

Bernoulli's equation states that the total head of the flow must be constant. Since the elevation does not change significantly, if at all, between points 1 and 2, the elevation head at the two points will be essentially the same and will cancel out of the equation. So Bernoulli's equation simplifies to Equation 3-13 for a venturi.

$$\frac{v_1^{\,2}}{2g} + P_1\nu_1 \frac{g_c}{g} = \frac{v_2^{\,2}}{2g} + P_2\nu_2 \frac{g_c}{g} \tag{3-13}$$

Applying the continuity equation to points 1 and 2 allows us to express the flow velocity at point 1 as a function of the flow velocity at point 2 and the ratio of the two flow areas.

$$\rho_1 A_1 v_1 = \rho_2 A_2 v_2$$

$$v_1 = \frac{\rho_2 A_2 v_2}{\rho_1 A_1}$$

$$v_1 = v_2 \frac{A_2}{A_1}$$

Using algebra to rearrange Equation 3-13 and substituting the above result for v_1 allows us to solve for v_2.

$$\frac{v_2^{\,2} - v_1^{\,2}}{2g} = (P_1 - P_2)\ \nu \frac{g_c}{g}$$

$$v_2^{\,2} - \left(v_2 \frac{A_2}{A_1}\right)^2 = (P_1 - P_2)\ 2\nu g_c$$

$$v_2^{\,2}\left(1 - \left(\frac{A_2}{A_1}\right)^2\right) = (P_1 - P_2)\ 2\nu g_c$$

$$v_2^{\,2} = \frac{(P_1 - P_2)\, 2\nu g_c}{\left(1 - \left(\frac{A_2}{A_1}\right)^2\right)}$$

$$v_2 = \sqrt{\frac{(P_1 - P_2)\, 2\nu g_c}{\left(1 - \left(\frac{A_2}{A_1}\right)^2\right)}}$$

$$v_2 = \sqrt{P_1 - P_2}\ \sqrt{\frac{2\nu g_c}{\left(1 - \left(\frac{A_2}{A_1}\right)^2\right)}}$$

Therefore the flow velocity at the throat of the venturi and the volumetric flow rate are directly proportional to the square root of the differential pressure.

The pressures at the upstream section and throat are actual pressures, and velocities from Bernoulli's equation without a loss term are theoretical velocities. When losses are considered in the energy equation, the velocities are actual velocities. First, with the Bernoulli equation (that is, without a head-loss term), the theoretical velocity at the throat is obtained. Then by multiplying this by the venturi factor (C_v), which accounts for friction losses and equals 0.98 for most venturis, the actual velocity is obtained. The actual velocity times the actual area of the throat determines the actual discharge volumetric flow rate.

The pressure drop, P_1-P_2, across the venturi can be used to measure the flow rate using a U-tube manometer as shown in Figure 6. The reading, R', of the manometer is proportional to the pressure drop and thus the velocity of the fluid.

Summary

The main points of this chapter are summarized below.

Bernoulli's Equation Summary

- Bernoulli's equation is an application of the First Law of Thermodynamics.
- Bernoulli's equation is an application of the general energy equation to a steady flow system in which no work is done on or by the fluid, no heat is transferred to or from the fluid, and no change occurs in the internal energy of the fluid.
- Head is the term used to describe pressure exerted on or by a fluid.
- As fluid flows in a piping system, changes in elevation, velocity, and pressure heads must be consistent so that Bernoulli's equation is satisfied.
- Bernoulli's equation can be modified to take into account friction losses and pump work.
- A venturi can be used to determine mass flow rates due to changes in pressure and fluid velocity.
- The volumetric flow rate through a venturi is directly proportional to the square root of the differential pressure between the venturi's inlet and its throat.

HEAD LOSS

The head loss that occurs in pipes is dependent on the flow velocity, pipe length and diameter, and a friction factor based on the roughness of the pipe and the Reynolds number of the flow. The head loss that occurs in the components of a flow path can be correlated to a piping length that would cause an equivalent head loss.

EO 1.21 DEFINE the terms head loss, frictional loss, and minor losses.

EO 1.22 DETERMINE friction factors for various flow situations using the Moody chart.

EO 1.23 CALCULATE the head loss in a fluid system due to frictional losses using Darcy's equation.

EO 1.24 CALCULATE the equivalent length of pipe that would cause the same head loss as the minor losses that occur in individual components.

Head Loss

Head loss is a measure of the reduction in the total head (sum of elevation head, velocity head and pressure head) of the fluid as it moves through a fluid system. Head loss is unavoidable in real fluids. It is present because of: the friction between the fluid and the walls of the pipe; the friction between adjacent fluid particles as they move relative to one another; and the turbulence caused whenever the flow is redirected or affected in any way by such components as piping entrances and exits, pumps, valves, flow reducers, and fittings.

Frictional loss is that part of the total head loss that occurs as the fluid flows through straight pipes. The head loss for fluid flow is directly proportional to the length of pipe, the square of the fluid velocity, and a term accounting for fluid friction called the friction factor. The head loss is inversely proportional to the diameter of the pipe.

$$\text{Head Loss} \propto f \, \frac{Lv^2}{D}$$

Friction Factor

The friction factor has been determined to depend on the Reynolds number for the flow and the degree of roughness of the pipe's inner surface.

The quantity used to measure the roughness of the pipe is called the relative roughness, which equals the average height of surface irregularities (ε) divided by the pipe diameter (D).

$$\text{Relative Roughness} = \frac{\varepsilon}{D}$$

The value of the friction factor is usually obtained from the Moody Chart (Figure B-1 of Appendix B). The Moody Chart can be used to determine the friction factor based on the Reynolds number and the relative roughness.

Example:

Determine the friction factor (f) for fluid flow in a pipe that has a Reynolds number of 40,000 and a relative roughness of 0.01.

Solution:

Using the Moody Chart, a Reynolds number of 40,000 intersects the curve corresponding to a relative roughness of 0.01 at a friction factor of 0.04.

Darcy's Equation

The frictional head loss can be calculated using a mathematical relationship that is known as Darcy's equation for head loss. The equation takes two distinct forms. The first form of Darcy's equation determines the losses in the system associated with the length of the pipe.

$$H_f = f \frac{L\ v^2}{D\ 2\ g} \tag{3-14}$$

where:

f = friction factor (unitless)
L = length of pipe (ft)
D = diameter of pipe (ft)
v = fluid velocity (ft/sec)
g = gravitational acceleration (ft/sec^2)

Example: Darcy's Head Loss Equation

A pipe 100 feet long and 20 inches in diameter contains water at 200°F flowing at a mass flow rate of 700 lbm/sec. The water has a density of 60 lbm/ft^3 and a viscosity of 1.978 x 10^{-7} lbf-sec/ft^2. The relative roughness of the pipe is 0.00008. Calculate the head loss for the pipe.

Solution:

The sequence of steps necessary to solve this problem is first to determine the flow velocity. Second, using the flow velocity and the fluid properties given, calculate the Reynolds number. Third, determine the friction factor from the Reynolds number and the relative roughness. Finally, use Darcy's equation to determine the head loss.

$$\dot{m} = \rho A v$$

$$v = \frac{\dot{m}}{\rho A}$$

$$= \frac{700 \ \frac{\text{lbm}}{\text{sec}}}{\left(60 \ \frac{\text{lbm}}{\text{ft}^3}\right) \pi \ (10 \ \text{in})^2 \ \frac{1 \ \text{ft}^2}{144 \ \text{in}^2}}$$

$$v = 5.35 \ \frac{\text{ft}}{\text{sec}}$$

$$N_R = \frac{\rho v D}{\mu \ g_c}$$

$$N_R = \frac{\left(60 \ \frac{\text{lbm}}{\text{ft}^3}\right)\left(5.35 \ \frac{\text{ft}}{\text{sec}}\right)\left(20 \ \text{in} \ \frac{1 \ \text{ft}}{12 \ \text{in}}\right)}{1.978 \ \text{x} \ 10^{-7} \ \frac{\text{lbf}-\text{sec}}{\text{ft}^2} \quad 32.17 \frac{\text{ft}-\text{lbm}}{\text{lbf}-\text{sec}^2}} = 8.4 \ \text{x} \ 10^7$$

Use the Moody Chart for a Reynolds number of 8.4 x 10^7 and a relative roughness of 0.00008.

$$f = 0.012$$

$$H_f = f \ \frac{L}{D} \ \frac{v^2}{2g}$$

$$= (0.012) \ \frac{100 \ \text{ft}}{20 \ \text{in} \left(\frac{1 \ \text{ft}}{12 \ \text{in}}\right)} \ \cdot \ \frac{\left(5.35 \ \frac{\text{ft}}{\text{sec}}\right)^2}{2\left(32.17 \ \frac{\text{ft}}{\text{sec}^2}\right)}$$

$$H_f = 0.32 \ \text{ft}$$

Minor Losses

The losses that occur in pipelines due to bends, elbows, joints, valves, etc. are sometimes called *minor losses*. This is a misnomer because in many cases these losses are more important than the losses due to pipe friction, considered in the preceding section. For all minor losses in turbulent flow, the head loss varies as the square of the velocity. Thus a convenient method of expressing the minor losses in flow is by means of a loss coefficient (k). Values of the loss coefficient (k) for typical situations and fittings is found in standard handbooks. The form of Darcy's equation used to calculate minor losses of individual fluid system components is expressed by Equation 3-15.

$$H_f = k \frac{v^2}{2g} \tag{3-15}$$

Equivalent Piping Length

Minor losses may be expressed in terms of the equivalent length (L_{eq}) of pipe that would have the same head loss for the same discharge flow rate. This relationship can be found by setting the two forms of Darcy's equation equal to each other.

$$f \frac{L\ v^2}{D\ 2\ g} = k \frac{v^2}{2\ g}$$

This yields two relationships that are useful.

$$L_{eq} = k \frac{D}{f} \tag{3-16}$$

$$k = f \frac{L_{eq}}{D} \tag{3-17}$$

Typical values of L_{eq}/D for common piping system components are listed in Table 1. The equivalent length of piping that will cause the same head loss as a particular component can be determined by multiplying the value of L_{eq}/D for that component by the diameter of the pipe. The higher the value of L_{eq}/D, the longer the equivalent length of pipe.

TABLE 1

Typical Values of $\frac{L_{eq}}{D}$

Item	$\frac{L_{eq}}{D}$
Globe Valve	
Conventional	400
Y-Pattern	160
Gate Valve	
Fully Open	10
75% Open	35
50% Open	150
25% Open	900
Standard Tee	
Flow through Run	10
Flow through Branch	60
90° Standard Elbow	30
45° Standard Elbow	16
Return Bend	50

Example:

A fully-open gate valve is in a pipe with a diameter of 10 inches. What equivalent length of pipe would cause the same head loss as the gate valve?

Solution:

From Table 1, we find that the value of L_{eq}/D for a fully-open gate valve is 10.

$$\begin{aligned} L_{eq} &= (L/D)\ D \\ &= 10\ (10\ \text{inches}) \\ &= 100\ \text{inches} \end{aligned}$$

By adding the equivalent lengths of all components to the actual length of pipe in a system we can obtain the L_{eq} value for the entire piping system.

Summary

The main points of this chapter are summarized below.

Head Loss Summary

- Head loss is the reduction in the total head (sum of potential head, velocity head, and pressure head) of a fluid caused by the friction present in the fluid's motion.
- Frictional loss is that part of the total head loss that occurs as the fluid flows through straight pipes.
- Minor losses are the head losses that occur due to bends, elbows, joints, valves, and other components. Any time the flow experiences a change in direction or a change in cross-sectional area, it will experience a head loss.
- The friction factor for fluid flow can be determined using a Moody Chart if the relative roughness of the pipe and the Reynolds number of the flow can be determined.
- Darcy's equation can be used to calculate frictional losses.
- A special form of Darcy's equation can be used to calculate minor losses.
- The length of pipe that would cause the same head loss as a valve or fitting can be determined by multiplying the value of L/D for the component found in handbooks or vendor manuals by the diameter of the pipe.

NATURAL CIRCULATION

Natural circulation is the circulation of fluid within piping systems or open pools that is due to the density changes caused by temperature differences. Natural circulation does not require any mechanical devices to maintain flow.

EO 1.25 DEFINE natural circulation and forced circulation.

EO 1.26 DEFINE thermal driving head.

EO 1.27 DESCRIBE the conditions necessary for natural circulation to exist.

EO 1.28 EXPLAIN the relationship between flow rate and temperature difference in natural circulation flow.

EO 1.29 DESCRIBE how the operator can determine whether natural circulation exists in the reactor coolant system and other heat removal systems.

EO 1.30 DESCRIBE how to enhance natural circulation flow.

Forced and Natural Circulation

In the previous chapters on fluid flow, it was explained that any time that fluid flows there is some friction associated with the movement, which will cause head loss. It was pointed out that this head loss is commonly compensated for in piping systems by pumps that do work on the fluid, compensating for the head loss due to friction. Circulation of fluid in systems by pumps is referred to as *forced circulation.*

It is possible to design some fluid systems in a manner that does not require the presence of pumps to provide circulation. The head required to compensate for the head losses is created by density gradients and elevation changes. Flow that occurs under these circumstances is called *natural circulation.*

Thermal Driving Head

Thermal driving head is the force that causes natural circulation to take place. It is caused by the difference in density between two bodies or areas of fluid.

Consider two equal volumes of the same type of fluid. If the two volumes are not at the same temperature, then the volume with the higher temperature will also have a lower density and, therefore, less mass. Since the volume at the higher temperature will have a lower mass, it will also have less force exerted on it by gravity. This difference in the force of gravity exerted on the fluid will tend to cause the hotter fluid to rise and the colder fluid to sink.

This effect is seen in many places. One example of this is a hot air balloon. The force causing a hot air balloon to rise is a result of a difference in density between the hot air inside the balloon and the cooler air surrounding it.

Heat added to the air in the balloon adds energy to the molecules of air. The movement of the air molecules increases and the air molecules take up more space. The air molecules inside the balloon take up more space than the same amount of air molecules outside the balloon. This means the hot air is less dense and lighter than the surrounding air. Since the air in the balloon is less dense, gravity has less effect on it. The result is that the balloon weighs less than the surrounding air. Gravity pulls cooler air down into the space occupied by the balloon. The downward movement of the cooler air forces the balloon out of the space previously occupied, and the balloon rises.

Conditions Required for Natural Circulation

Natural circulation will only occur if the correct conditions exist. Even after natural circulation has begun, removal of any one of these conditions will cause the natural circulation to stop. The conditions for natural circulation are as follows.

1. A temperature difference exists (heat source and heat sink exists).

2. The heat source is at a lower elevation than the heat sink.

3. The fluids must be in contact with each other.

There must be two bodies of fluid at different temperatures. This could also be one body of fluid with areas of different temperatures. The difference in temperature is necessary to cause a density difference in the fluid. The density difference is the driving force for natural circulation flow.

The difference in temperature must be maintained for the natural circulation to continue. Addition of heat by a heat source must exist at the high temperature area. Continuous removal of heat by a heat sink must exist at the low temperature area. Otherwise the temperatures would eventually equalize, and no further circulation would occur.

The heat source must be at a lower elevation than the heat sink. As shown by the example of the balloon, a warmer fluid is less dense and will tend to rise, and a cooler fluid is more dense and will tend to sink. To take advantage of the natural movement of warm and cool fluids, the heat source and heat sink must be at the proper elevations.

The two areas must be in contact so that flow between the areas is possible. If the flow path is obstructed or blocked, then natural circulation cannot occur.

Example of Natural Circulation Cooling

Natural circulation is frequently the primary means of cooling for pool-type reactors and for irradiated fuel assemblies stored in pools of water after removal from the reactor. The heat source is the fuel assembly. The heat sink is the bulk of the water in the pool.

Water at the bottom of a fuel assembly absorbs energy generated by the assembly. The water increases in temperature and decreases in density. Gravity pulls cooler (more dense) water into the bottom of the assembly displacing the warmer water. The warmer (lighter) water is forced to give up its position to the cooler (heavier) water. The warmer (lighter) water rises higher in the assembly. As water travels up the length of the assembly, it absorbs more energy. The water becomes lighter and lighter being continuously forced upward by more dense water moving in below it. In turn, the cooler water absorbs energy from the assembly and is also forced to rise as natural circulation flow continues. Water exiting the top of the fuel assembly gives up its energy as it mixes with the bulk of the water in the pool. The bulk of the water in the pool is commonly cooled by circulation through heat exchangers in a separate process.

Flow Rate and Temperature Difference

The thermal driving head that causes natural circulation is due to the density change caused by a temperature difference. In general, the greater the temperature difference between the hot and cold areas of fluid, the greater the thermal driving head and the resulting flow rate. However, it is good practice to keep the hot fluid subcooled to prevent a change of phase from occurring. It is possible to have natural circulation take place in two-phase flow, but it is usually more difficult to maintain flow.

Various parameters can be used to indicate or verify natural circulation is occurring. This is dependent on plant type. For instance for a pressurized water reactor (PWR) selected Reactor Coolant System (RCS) parameters that would be used are as follows.

1. RCS ΔT (T_{Hot} - T_{Cold}) should be 25-80% of the full power value and either steady or slowly decreasing. This indicates that the decay heat is being removed from the system at an adequate rate to maintain or reduce core temperatures.

2. RCS Hot and Cold leg temperatures should be steady or slowly decreasing. Again, this indicates that heat is being removed and the decay heat load is decreasing as expected.

3. Steam generator steam pressure (secondary side pressure) should be following RCS temperature. This verifies that the steam generator is removing heat from the RCS coolant.

If natural circulation for a PWR is in progress or is imminent, several actions can be performed to ensure or enhance core cooling capabilities. First, pressurizer level can be maintained greater than 50%. Secondly, maintain the RCS subcooled by 15°F or greater.

Both of these actions will help ensure steam/vapor pockets are not formed in the RCS where they would restrict RCS flow. Thirdly, maintain steam generator water level ≥ normal range. This provides an adequate heat sink to ensure heat removal is sufficient to prevent boiling of the RCS.

Summary

The main points of this chapter are listed below.

Natural Circulation Flow Summary

- Natural circulation flow is circulation of a fluid without the use of mechanical devices.

- Forced circulation flow is circulation of a fluid through a system by pumps.

- Thermal driving head is the driving force for natural circulation caused by the difference in density between two areas of fluid.

- Three items are necessary to support natural circulation:

  - There must be a heat sink and a heat source.
  - The heat source must be located below the heat sink.
  - Flowpaths must exist between the warm fluid and the cold fluid.

- Generally, the greater the temperature difference, the higher the natural circulation flow rate.

- Natural circulation in a PWR can be verified by monitoring:

  - RCS ΔT - 25%-80% full power value
  - T_{Hot} / T_{Cold} - steady or slowly decreasing
  - S/G steam pressure - tracking RCS temperature

- Natural circulation in a PWR can be enhanced by:

  - maintain pressurizer level >50%
  - maintain RCS ≥ 15°F subcooling
  - maintain adequate heat sink, S/G level ≥ normal range

TWO-PHASE FLUID FLOW

Water at saturation conditions may exist as both a fluid and a vapor. This mixture of steam and water can cause unusual flow characteristics within fluid systems.

EO 1.31 DEFINE two-phase flow.

EO 1.32 DESCRIBE two-phase flow including such phenomena as bubbly, slug, and annular flow.

EO 1.33 DESCRIBE the problems associated with core flow oscillations and flow instability.

EO 1.34 DESCRIBE the conditions that could lead to core flow oscillation and instability.

EO 1.35 DESCRIBE the phenomenon of pipe whip.

EO 1.36 DESCRIBE the phenomenon of water hammer.

Two-Phase Fluid Flow

All of the fluid flow relationships discussed previously are for the flow of a single phase of fluid whether liquid or vapor. At certain important locations in fluid flow systems the simultaneous flow of liquid water and steam occurs, known as *two-phase flow*. These simple relationships used for analyzing single-phase flow are insufficient for analyzing two-phase flow.

There are several techniques used to predict the head loss due to fluid friction for two-phase flow. Two-phase flow friction is greater than single-phase friction for the same conduit dimensions and mass flow rate. The difference appears to be a function of the type of flow and results from increased flow speeds. Two-phase friction losses are experimentally determined by measuring pressure drops across different piping elements. The two-phase losses are generally related to single-phase losses through the same elements.

One accepted technique for determining the two-phase friction loss based on the single-phase loss involves the two-phase friction multiplier (R), which is defined as the ratio of the two-phase head loss divided by the head loss evaluated using saturated liquid properties.

$$R = \frac{H_f,\ \text{two-phase}}{H_f,\ \text{saturated liquid}} \tag{3-18}$$

where:

R	=	two-phase friction multiplier (no units)
H_f , two-phase	=	two-phase head loss due to friction (ft)
H_f , saturated liquid	=	single-phase head loss due to friction (ft)

The friction multiplier (R) has been found to be much higher at lower pressures than at higher pressures. The two-phase head loss can be many times greater than the single-phase head loss.

Although a wide range of names has been used for two-phase flow patterns, we shall define only three types of flow. The flow patterns to be used are defined as follows:

1. Bubbly flow: there is dispersion of vapor bubbles in a continuum of liquid.

2. Slug flow: in bubbly flow, the bubbles grow by coalescence and ultimately become of the same order of diameter as the tube. This generates the typical bullet-shaped bubbles that are characteristic of the slug-flow regime.

3. Annular flow: the liquid is now distributed between a liquid film flowing up the wall and a dispersion of droplets flowing in the vapor core of the flow.

Flow Instability

Unstable flow can occur in the form of flow oscillations or flow reversals. *Flow oscillations* are variations in flow due to void formations or mechanical obstructions from design and manufacturing. A flow oscillation in one reactor coolant channel sometimes causes flow oscillations in the surrounding coolant channels due to flow redistribution. Flow oscillations are undesirable for several reasons. First, sustained flow oscillations can cause undesirable forced mechanical vibration of components. This can lead to failure of those components due to fatigue. Second, flow oscillations can cause system control problems of particular importance in liquid-cooled nuclear reactors because the coolant is also used as the moderator. Third, flow oscillations affect the local heat transfer characteristics and boiling. It has been found through testing that the critical heat flux (CHF) required for departure from nucleate boiling (DNB) can be lowered by as much as 40% when flow is oscillating. This severely reduces the thermal limit and the power density along the length of the reactor core. Again, it has been found through testing that flow oscillations are not a significant problem for some pressurized water reactors unless power is above 150% for the normal flow conditions. Flow oscillations can be a problem during natural circulation operations because of the low flow rates present.

During natural circulation, the steam bubbles formed during a flow oscillation may have enough of an effect to actually cause complete flow reversal in the affected channel.

Both the flow oscillations and flow reversals lead to a very unstable condition since the steam blankets formed on heated surfaces directly affect the ability to transfer heat away from those surfaces.

Pipe Whip

If a pipe were to rupture, the reaction force created by the high velocity fluid jet could cause the piping to displace and cause extensive damage to components, instrumentation, and equipment in the area of the rupture. This characteristic is similar to an unattended garden hose or fire hose "whipping" about unpredictably. This type of failure is analyzed to minimize damage if pipe whip were to occur in the vicinity of safety-related equipment.

Water Hammer

Water hammer is a liquid shock wave resulting from the sudden starting or stopping of flow. It is affected by the initial system pressure, the density of the fluid, the speed of sound in the fluid, the elasticity of the fluid and pipe, the change in velocity of the fluid, the diameter and thickness of the pipe, and the valve operating time.

During the closing of a valve, kinetic energy of the moving fluid is converted into potential energy. Elasticity of the fluid and pipe wall produces a wave of positive pressure back toward the fluid's source. When this wave reaches the source, the mass of fluid will be at rest, but under tremendous pressure. The compressed liquid and stretched pipe walls will now start to release the liquid in the pipe back to the source and return to the static pressure of the source. This release of energy will form another pressure wave back to the valve. When this shockwave reaches the valve, due to the momentum of the fluid, the pipe wall will begin to contract. This contraction is transmitted back to the source, which places the pressure in the piping below that of the static pressure of the source. These pressure waves will travel back and forth several times until the fluid friction dampens the alternating pressure waves to the static pressure of the source. Normally, the entire hammer process takes place in under one second.

The initial shock of suddenly stopped flow can induce transient pressure changes that exceed the static pressure. If the valve is closed slowly, the loss of kinetic energy is gradual. If it is closed quickly, the loss of kinetic energy is very rapid. A shock wave results because of this rapid loss of kinetic energy. The shock wave caused by water hammer can be of sufficient magnitude to cause physical damage to piping, equipment, and personnel. Water hammer in pipes has been known to pull pipe supports from their mounts, rupture piping, and cause pipe whip.

Pressure Spike

A *pressure spike* is the resulting rapid rise in pressure above static pressure caused by water hammer. The highest pressure spike attained will be at the instant the flow changed and is governed by the following equation.

$$\Delta P = \frac{\rho_c \, \Delta v}{g_c}$$

where:

ΔP = Pressure spike $\left(\frac{lbf}{ft^2}\right)$

ρ = Density of the fluid $\left(\frac{lbm}{ft^3}\right)$

c = Velocity of the pressure wave (Speed of sound in the fluid) $\left(\frac{ft}{sec}\right)$

Δv = Change in velocity of the fluid $\left(\frac{ft}{sec}\right)$

g_c = Gravitational constant 32.17 $\left(\frac{lbm-ft}{lbf-sec^2}\right)$

Example: Pressure spike

Water at a density of 62.4 lbm/ft^3 and a pressure of 120 psi is flowing through a pipe at 10 ft/sec. The speed of sound in the water is 4780 ft/sec. A check valve suddenly closed. What is the maximum pressure of the fluid in psi?

Solution:

$$P_{Max} = P_{Static} + \Delta P_{Spike}$$

$$P_{Max} = 120\frac{lbf}{in^2} + \frac{\rho_c\ \Delta V}{g_c}$$

$$P_{Max} = 120\frac{lbf}{in^2} + \frac{62.4\frac{lbm}{ft^3}\quad 4780\frac{ft}{sec}\quad 10\frac{ft}{sec}}{32.17\ \frac{lbm\ ft}{lbf\ sec^2}}$$

$$P_{Max} = 120\frac{lbf}{in^2} + 92{,}631\frac{lbf}{ft^2}\left(\frac{ft^2}{144\ in^2}\right)$$

$$P_{Max} = 120\frac{lbf}{in^2} + 643.3\frac{lbf}{in^2}$$

$$P_{Max} = 763.3\ psi$$

Steam Hammer

Steam hammer is similar to water hammer except it is for a steam system. *Steam hammer* is a gaseous shock wave resulting from the sudden starting or stopping of flow. Steam hammer is not as severe as water hammer for three reasons:

1. The compressibility of the steam dampens the shock wave
2. The speed of sound in steam is approximately one third the speed of sound in water.
3. The density of steam is approximately 1600 times less than that of water.

The items of concern that deal with steam piping are thermal shock and water slugs (i.e., condensation in the steam system) as a result of improper warm up.

Operational Considerations

Water and steam hammer are not uncommon occurrences in industrial plants. Flow changes in piping systems should be done slowly as part of good operator practice. To prevent water and steam hammer, operators should ensure liquid systems are properly vented and ensure gaseous or steam systems are properly drained during start-up. When possible, initiate pump starts against a closed discharge valve, and open the discharge valve slowly to initiate system flow. If possible, start-up smaller capacity pumps before larger capacity pumps. Use warm-up valves around main stream stop valves whenever possible. If possible, close pump discharge valves before stopping pumps. Periodically verify proper function of moisture traps and air traps during operation.

Summary

The main points from this chapter are summarized below.

Two-Phase Fluid Flow Summary

- The combination of liquid and vapor flowing through a pipe is called two-phase flow.

- Types of two-phase flow include:

  - Bubbly flow: there is a dispersion of vapor bubbles in a continuum of liquid.
  - Slug flow: the bubbles grow by coalescence and ultimately become of the same order of diameter as the tube, generating bullet shaped bubbles.
  - Annular flow: the liquid is distributed between a liquid film flowing up the wall and a dispersion of droplets flowing in the vapor core of the flow.

- Core flow oscillations and instabilities can cause:

  - undesirable mechanical vibration of components.
  - a reduction in the heat flux required to cause DNB.
  - interruptions to actual circulation flow.

- Flow oscillations and instabilities can occur during the following conditions:

  - core is outside design conditions, power > 150%
  - mechanical failure, causing flow blockage
  - inadequate core cooling during natural circulation, such that boiling is occurring

- Pipe whip is the displacement of piping created by the reaction forces of a high velocity fluid jet following a pipe rupture.

- Water hammer is a liquid shock wave resulting from a sudden starting or stopping of flow.

CENTRIFUGAL PUMPS

Centrifugal pumps are one of the most common components found in fluid systems. In order to understand how a fluid system containing a centrifugal pump operates, it is necessary to understand the head and flow relationships for a centrifugal pump.

EO 1.37 DEFINE the terms net positive suction head and cavitation.

EO 1.38 CALCULATE the new volumetric flow rate, head, or power for a variable speed centrifugal pump using the pump laws.

EO 1.39 DESCRIBE the effect on system flow and pump head for the following changes:
a. Changing pump speeds
b. Adding pumps in parallel
c. Adding pumps in series

Energy Conversion in a Centrifugal Pump

Fluid entering a centrifugal pump is immediately directed to the low pressure area at the center or eye of the impeller. As the impeller and blading rotate, they transfer momentum to incoming fluid. A transfer of momentum to the moving fluid increases the fluid's velocity. As the fluid's velocity increases its kinetic energy increases. Fluid of high kinetic energy is forced out of the impeller area and enters the volute.

The volute is a region of continuously increasing cross-sectional area designed to convert the kinetic energy of the fluid into fluid pressure. The mechanism of this energy conversion is the same as that for subsonic flow through the diverging section of a nozzle. The mathematical analysis of flow through the volute is based on the general energy equation, the continuity equation, and the equation relating the internal properties of a system. The key parameters influencing the energy conversion are the expanding cross-sectional area of the volute, the higher system back pressure at the discharge of the volute, and the incompressible, subsonic flow of the fluid. As a result of the interdependence of these parameters, the fluid flow in the volute, similar to subsonic flow in a diverging nozzle, experiences a velocity decrease and a pressure increase.

Operating Characteristics of a Centrifugal Pump

Normally, a centrifugal pump produces a relatively low pressure increase in the fluid. This pressure increase can be anywhere from several dozen to several hundred psid across a centrifugal pump with a single stage impeller. The term PSID (**P**ounds Force Per **S**quare **I**nch **D**ifferential) is equivalent to ΔP. In this context, it is the pressure difference between the suction and discharge of a pump. PSID can also be used to describe a pressure drop across a system component (strainers, filters, heat exchangers, valves, demineralizers, etc.). When a centrifugal pump is operating at a constant speed, an increase in the system back pressure on the flowing stream causes a reduction in the magnitude of volumetric flow rate that the centrifugal pump can maintain.

Analysis of the relationship between the volumetric flow rate ($\dot{V}$) that a centrifugal pump can maintain and the pressure differential across the pump (ΔP_{pump}) is based on various physical characteristics of the pump and the system fluid. Variables evaluated by design engineers to determine this relationship include the pump efficiency, the power supplied to the pump, the rotational speed, the diameter of the impeller and blading, the fluid density, and the fluid viscosity. The result of this complicated analysis for a typical centrifugal pump operating at one particular speed is illustrated by the graph in Figure 7.

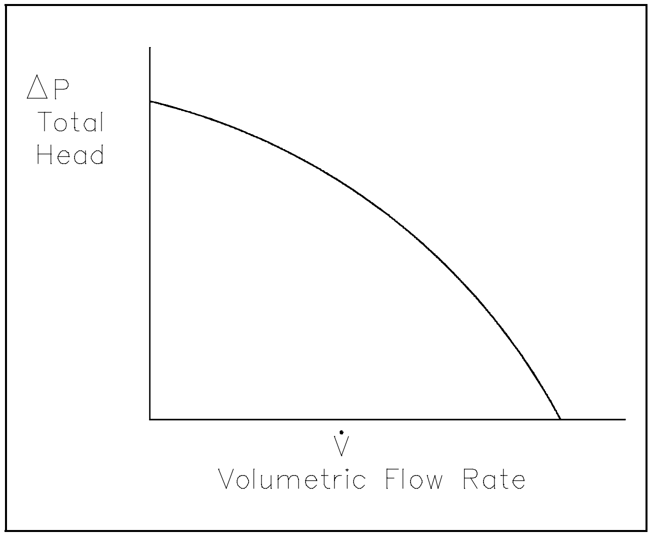


Figure 7 Typical Centrifugal Pump Characteristic Curve

Pump head, on the vertical axis, is the difference between system back pressure and the inlet pressure of the pump (ΔP_{pump}). Volumetric flow rate ($\dot{V}$), on the horizontal axis, is the rate at which fluid is flowing through the pump. The graph assumes one particular speed (N) for the pump impeller.

Cavitation

When the liquid being pumped enters the eye of a centrifugal pump, the pressure is significantly reduced. The greater the flow velocity through the pump the greater this pressure drop. If the pressure drop is great enough, or if the temperature of the liquid is high enough, the pressure drop may be sufficient to cause the liquid to flash to steam when the local pressure falls below the saturation pressure for the fluid that is being pumped. These vapor bubbles are swept along the pump impeller with the fluid. As the flow velocity decreases the fluid pressure increases. This causes the vapor bubbles to suddenly collapse on the outer portions of the impeller. The formation of these vapor bubbles and their subsequent collapse is *cavitation*.

Cavitation can be a very serious problem for centrifugal pumps. Some pumps can be designed to operate with limited amounts of cavitation. Most centrifugal pumps cannot withstand cavitation for significant periods of time; they are damaged by erosion of the impeller, vibration, or some other cavitation-induced problem.

Net Positive Suction Head

It is possible to ensure that cavitation is avoided during pump operation by monitoring the net positive suction head of the pump. *Net positive suction head* (NPSH) for a pump is the difference between the suction pressure and the saturation pressure of the fluid being pumped. NPSH is used to measure how close a fluid is to saturated conditions. Equation 3-19 can be used to calculate the net positive suction head available for a pump. The units of NPSH are feet of water.

$$NPSH = P_{suction} - P_{saturation} \tag{3-19}$$

where:

$P_{suction}$ = suction pressure of the pump
$P_{saturation}$ = saturation pressure for the fluid

By maintaining the available NPSH at a level greater than the NPSH required by the pump manufacturer, cavitation can be avoided.

Pump Laws

Centrifugal pumps generally obey what are known as the pump laws. These laws state that the flow rate or capacity is directly proportional to the pump speed; the discharge head is directly proportional to the square of the pump speed; and the power required by the pump motor is directly proportional to the cube of the pump speed. These laws are summarized in the following equations.

$$\dot{V} \propto n \tag{3-20}$$

$$H_p \propto n^2 \tag{3-21}$$

$$p \propto n^3 \tag{3-22}$$

where:

n = speed of pump impeller (rpm)
$\dot{V}$ = volumetric flow rate of pump (gpm or ft^3/hr)
H_p = head developed by pump (psid or feet)
p = pump power (kW)

Using these proportionalities, it is possible to develop equations relating the condition at one speed to those at a different speed.

$$\dot{V}_1 \left(\frac{n_2}{n_1}\right) = \dot{V}_2 \tag{3-23}$$

$$H_{p_1} \left(\frac{n_2}{n_1}\right)^2 = H_{p_2} \tag{3-24}$$

$$p_1 \left(\frac{n_2}{n_1}\right)^3 = p_2 \tag{3-25}$$

Example: Pump Laws

A cooling water pump is operating at a speed of 1800 rpm. Its flow rate is 400 gpm at a head of 48 ft. The power of the pump is 45 kW. Determine the pump flow rate, head, and power requirements if the pump speed is increased to 3600 rpm.

Solution:

Flow rate

$$\dot{V}_2 = \dot{V}_1 \left(\frac{n_2}{n_1}\right)$$

$$= (400\ \text{gpm}) \left(\frac{3600\ \text{rpm}}{1800\ \text{rpm}}\right)$$

$$= 800\ \text{gpm}$$

Head

$$H_{p_2} = H_{p_1} \left(\frac{n_2}{n_1} \right)^2$$

$$= 48 \text{ ft} \left(\frac{3600 \text{ rpm}}{1800 \text{ rpm}} \right)^2$$

$$= 192 \text{ ft}$$

Power

$$P_2 = P_1 \left(\frac{n_2}{n_1} \right)^3$$

$$= 45 \text{ kW} \left(\frac{3600 \text{ rpm}}{1800 \text{ rpm}} \right)^3$$

$$= 360 \text{ kW}$$

It is possible to develop the characteristic curve for the new speed of a pump based on the curve for its original speed. The technique is to take several points on the original curve and apply the pump laws to determine the new head and flow at the new speed. The pump head versus flow rate curve that results from a change in pump speed is graphically illustrated in Figure 8.

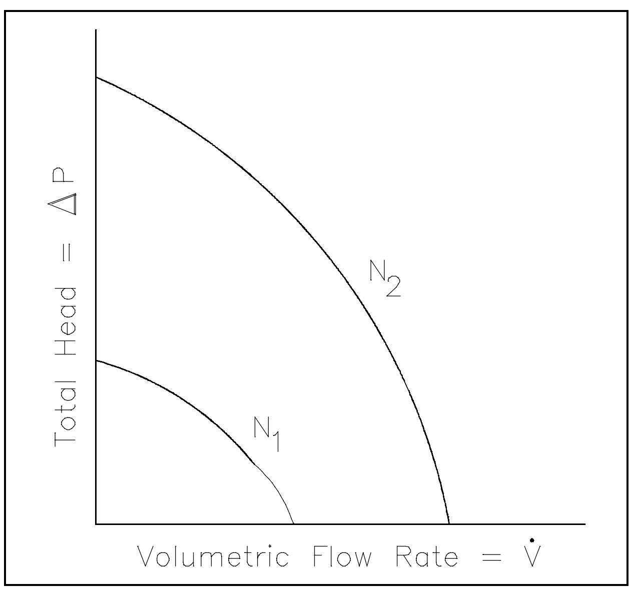


Figure 8 Changing Speeds for Centrifugal Pump

System Characteristic Curve

In the chapter on head loss, it was determined that both frictional losses and minor losses in piping systems were proportional to the square of the flow velocity. Since flow velocity is directly proportional to the volumetric flow rate, the system head loss must be directly proportional to the square of the volumetric flow rate. From this relationship, it is possible to develop a curve of system head loss versus volumetric flow rate. The head loss curve for a typical piping system is in the shape of a parabola as shown in Figure 9.

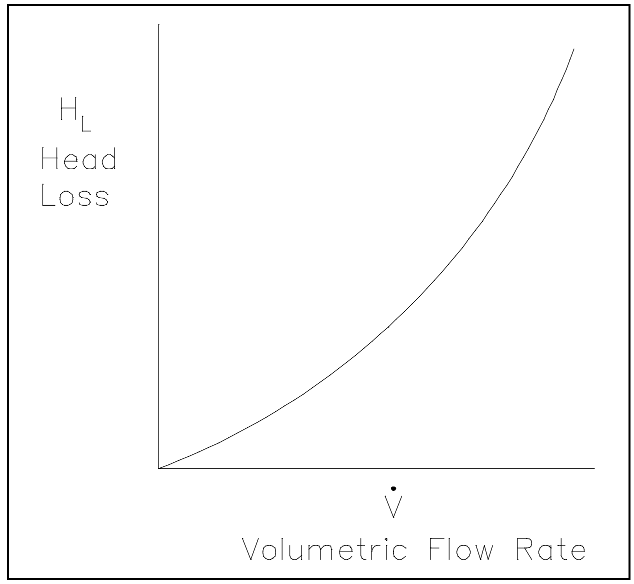


Figure 9 Typical System Head Loss Curve

System Operating Point

The point at which a pump operates in a given piping system depends on the flow rate and head loss of that system. For a given system, volumetric flow rate is compared to system head loss on a system characteristic curve. By graphing a system characteristic curve and the pump characteristic curve on the same coordinate system, the point at which the pump must operate is identified. For example, in Figure 10, the operating point for the centrifugal pump in the original system is designated by the intersection of the pump curve and the system curve (h_{Lo}).

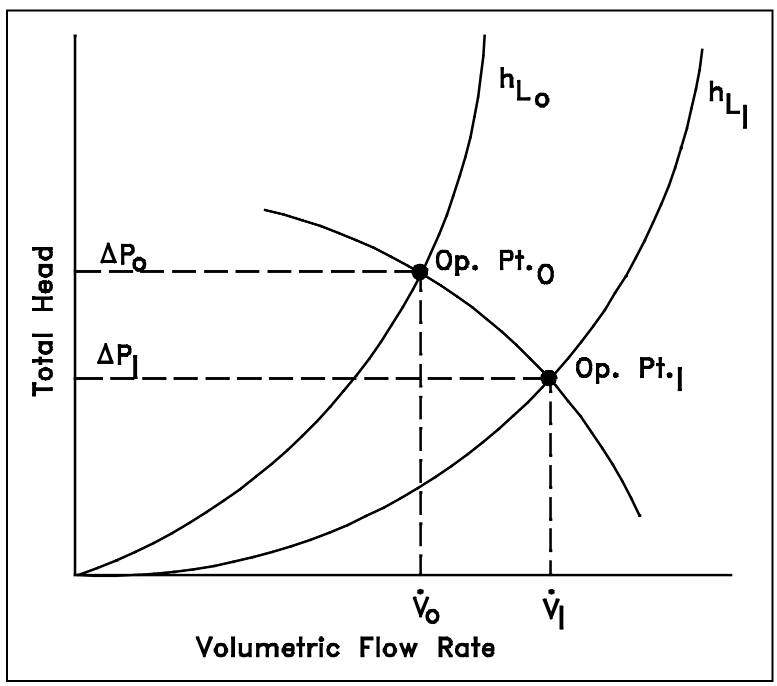


Figure 10 Operating Point for a Centrifugal Pump

The system has a flow rate equal to $\dot{V}_o$ and a total system head loss equal to ΔP_o. In order to maintain the flow rate ($\dot{V}_o$), the pump head must be equal to ΔP_o. In the system described by the system curve (h_{L1}), a valve has been opened in the system to reduce the system's resistance to flow. For this system, the pump maintains a large flow rate ($\dot{V}_1$) at a smaller pump head (ΔP_1).

System Use of Multiple Centrifugal Pumps

A typical centrifugal pump has a relatively low number of moving parts and can be easily adapted to a variety of prime movers. These prime movers include AC and DC electric motors, diesel engines, steam turbines, and air motors. Centrifugal pumps are typically small in size and can usually be built for a relatively low cost. In addition, centrifugal pumps provide a high volumetric flow rate with a relatively low pressure.

In order to increase the volumetric flow rate in a system or to compensate for large flow resistances, centrifugal pumps are often used in parallel or in series. Figure 11 depicts two identical centrifugal pumps operating at the same speed in parallel.

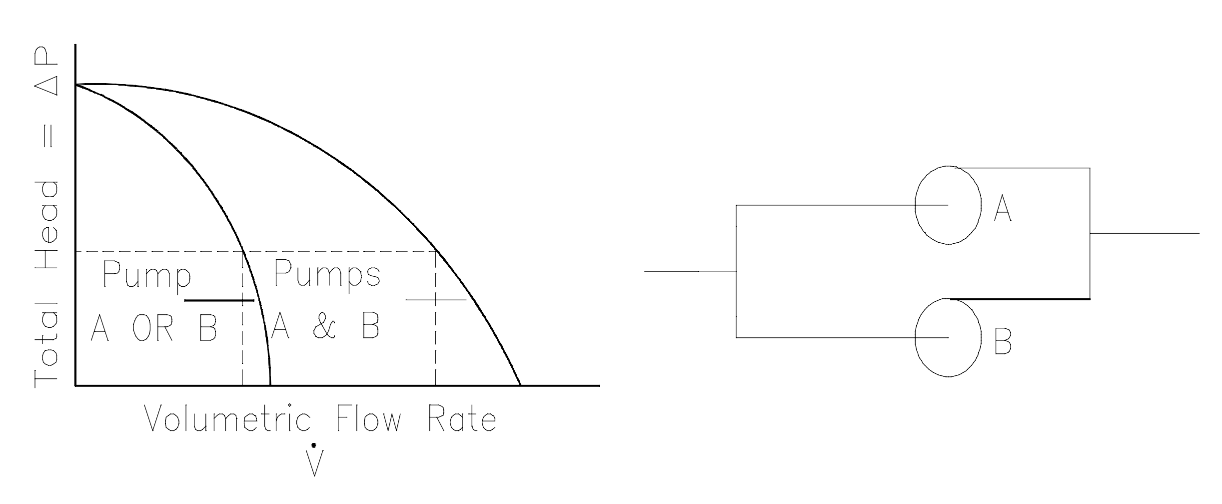


Figure 11 Pump Characteristic Curve for Two Identical Centrifugal Pumps Used in Parallel

Centrifugal Pumps in Parallel

Since the inlet and the outlet of each pump shown in Figure 11 are at identical points in the system, each pump must produce the same pump head. The total flow rate in the system, however, is the sum of the individual flow rates for each pump.

When the system characteristic curve is considered with the curve for pumps in parallel, the operating point at the intersection of the two curves represents a higher volumetric flow rate than for a single pump and a greater system head loss. As shown in Figure 12, a greater system head loss occurs with the increased fluid velocity resulting from the increased volumetric flow rate. Because of the greater system head, the volumetric flow rate is actually less than twice the flow rate achieved by using a single pump.

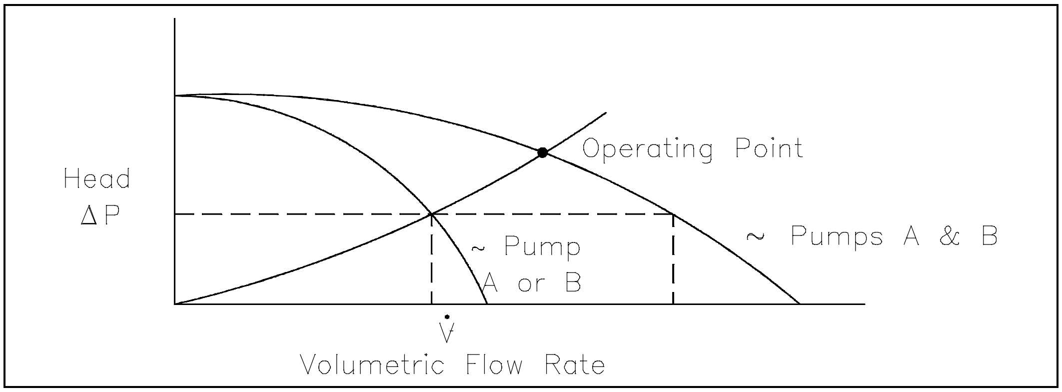


Figure 12 Operating Point for Two Parallel Centrifugal Pumps

Centrifugal Pumps in Series

Centrifugal pumps are used in series to overcome a larger system head loss than one pump can compensate for individually. As illustrated in Figure 13, two identical centrifugal pumps operating at the same speed with the same volumetric flow rate contribute the same pump head. Since the inlet to the second pump is the outlet of the first pump, the head produced by both pumps is the sum of the individual heads. The volumetric flow rate from the inlet of the first pump to the outlet of the second remains the same.

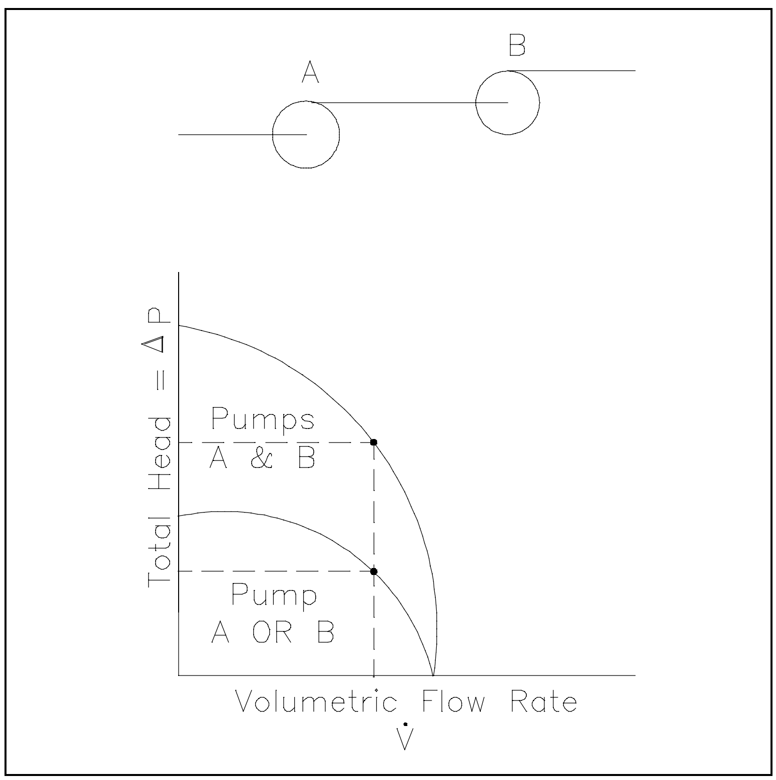


Figure 13 Pump Characteristic Curve for Two Identical Centrifugal Pumps Used in Series

As shown in Figure 14, using two pumps in series does not actually double the resistance to flow in the system. The two pumps provide adequate pump head for the new system and also maintain a slightly higher volumetric flow rate.

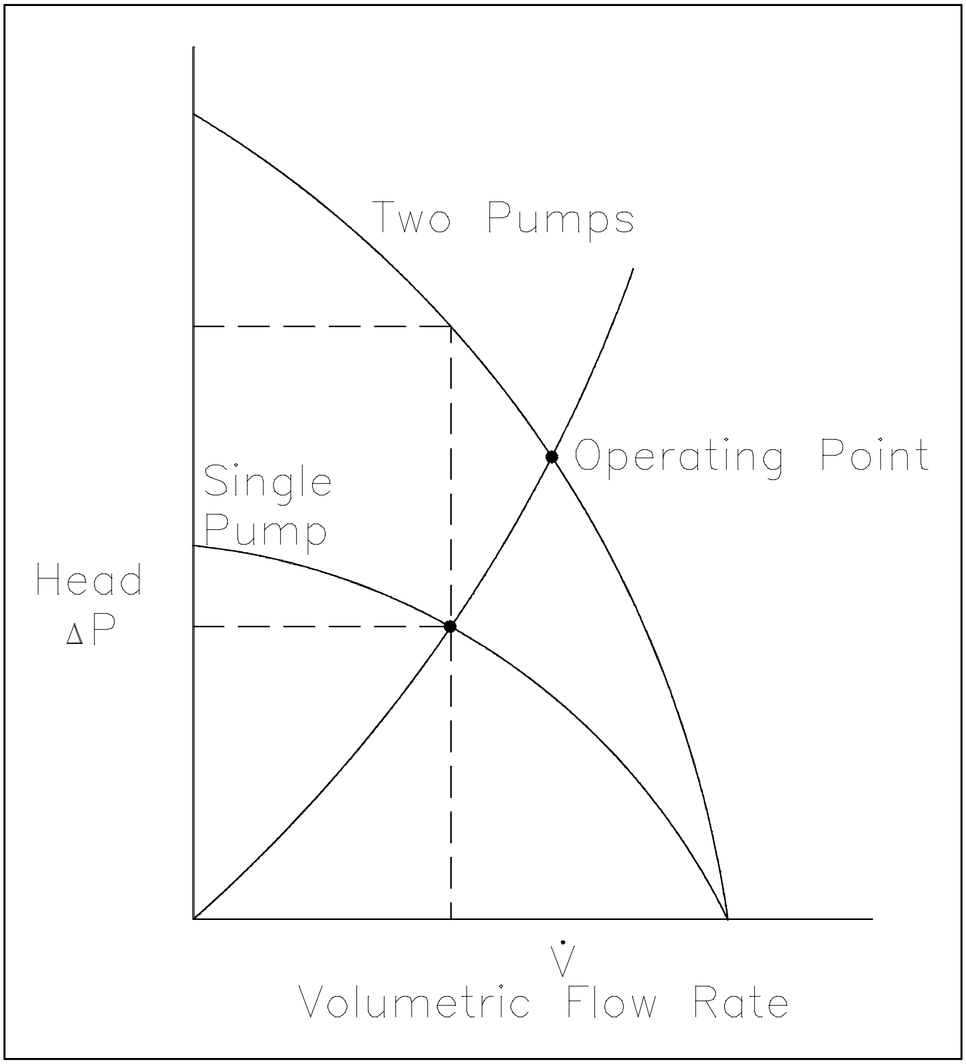


Figure 14 Operating Point for Two Centrifugal Pumps in Series

Summary

The main points from this chapter are summarized below.

Centrifugal Pumps Summary

- Net positive suction head is the difference between the pump suction pressure and the saturation pressure for the fluid.
- Cavitation is the formation and subsequent collapse of vapor bubbles on the impeller of a pump as the local pressure falls below and then rises above the saturation pressure of the fluid being pumped.
- The pump laws can be used to determine the effect of varying the speed of a centrifugal pump on the flow, head, and power.

$$\dot{V}_1 \left(\frac{n_2}{n_1}\right) = \dot{V}_2$$

$$H_{p_1} \left(\frac{n_2}{n_1}\right)^2 = H_{p_2}$$

$$p_1 \left(\frac{n_2}{n_1}\right)^3 = p_2$$

- The combined pump curve for two centrifugal pumps in parallel can be determined by adding the individual flows for any given head.
- The combined pump curve for two centrifugal pumps in series can be determined by adding the individual heads for any given flow.
- The operating point (head and flow) of a system can be determined by plotting the pump curve and the system head loss curve on the same axes. The system will operate at the intersection of the two curves.

Appendix B

Fluid Flow

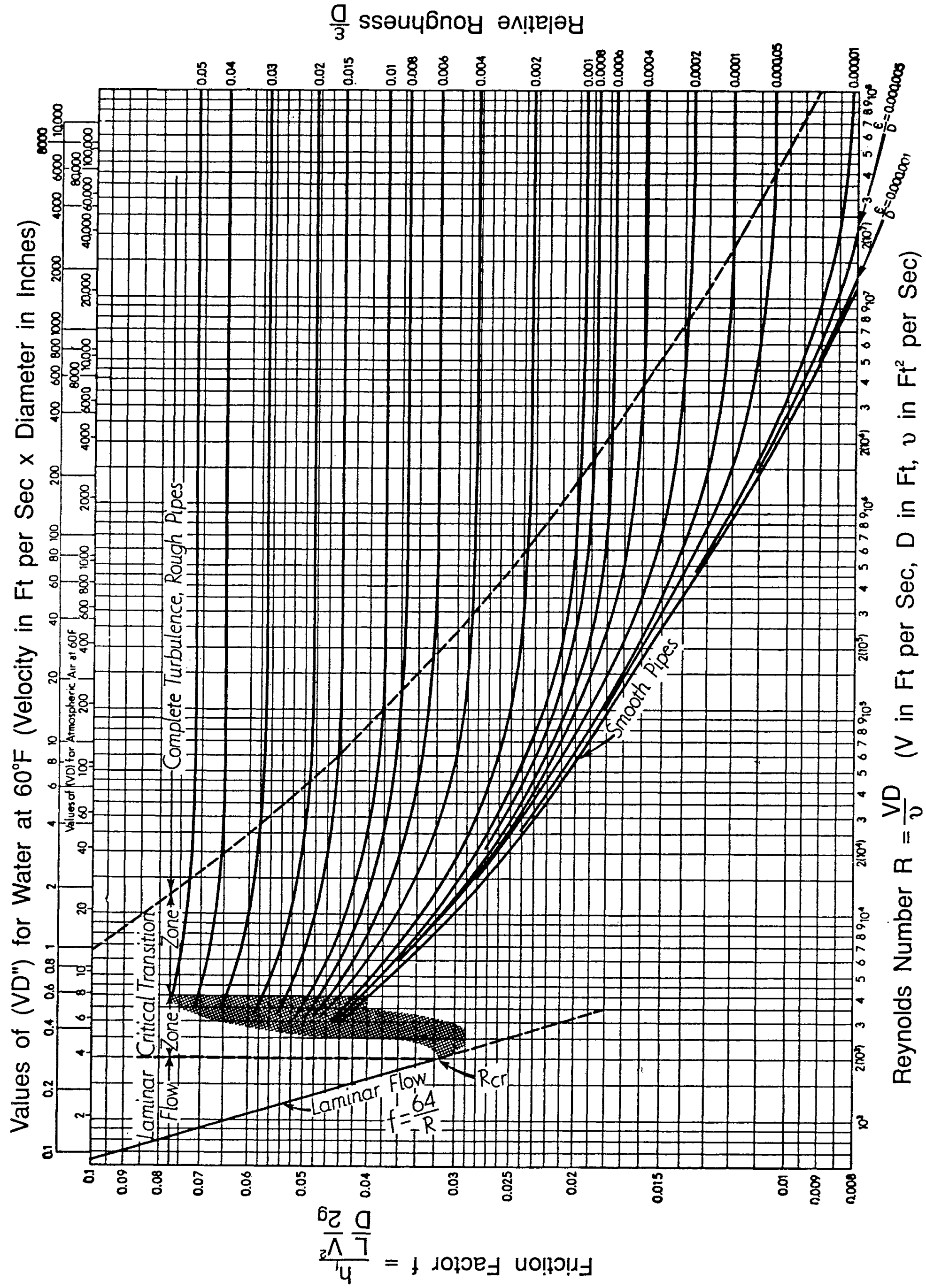


Figure B-1 Moody Chart

end of text.

CONCLUDING MATERIAL

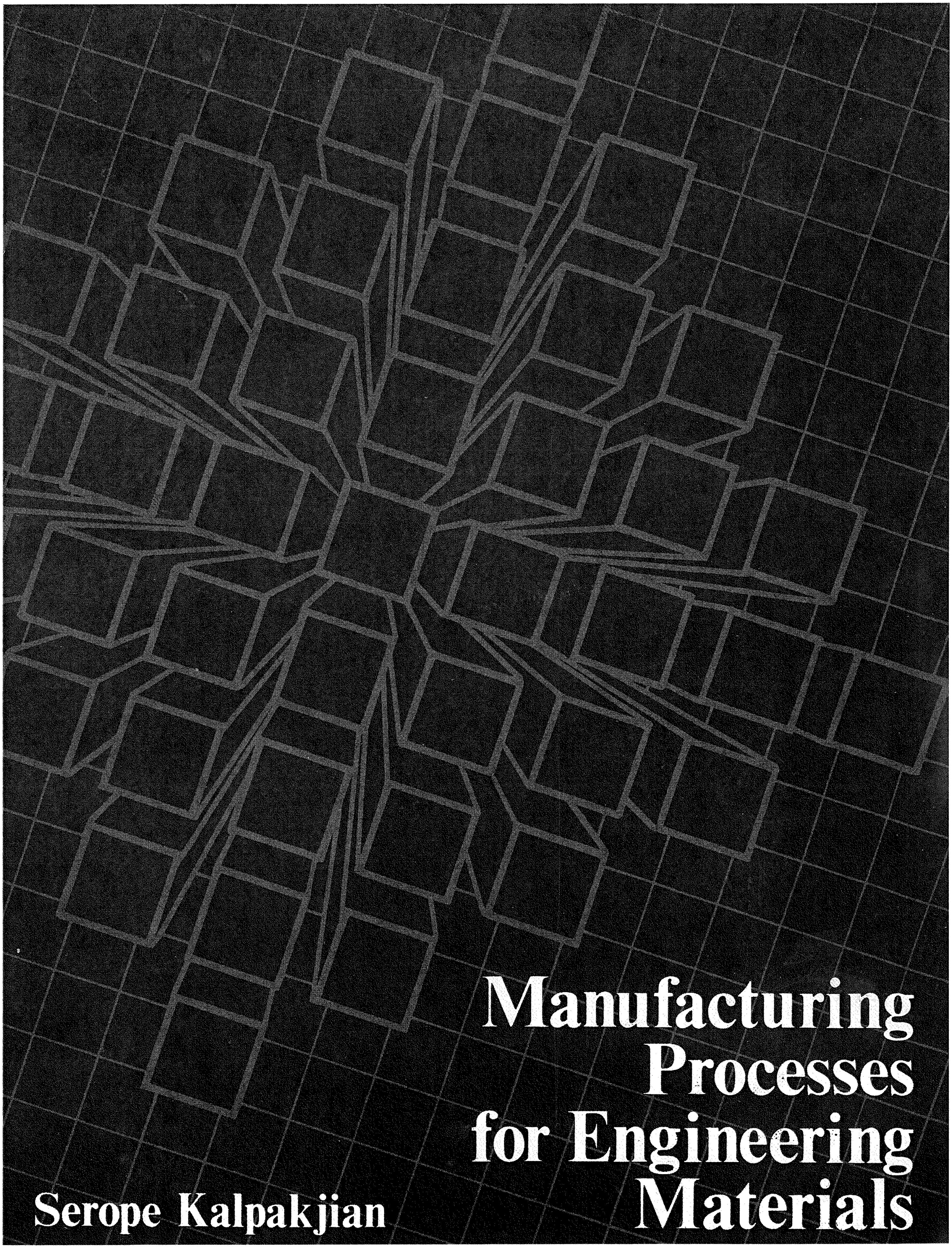
Review activities:

DOE - ANL-W, BNL, EG&G Idaho,
EG&G Mound, EG&G Rocky Flats,
LLNL, LANL, MMES, ORAU, REECo,
WHC, WINCO, WEMCO, and WSRC.

Preparing activity:

DOE - NE-73
Project Number 6910-0018/3

Exhibit C


Manufacturing
Processes
for Engineering
Materials
Serope Kalpakjian

Serope Kalpakjian
Illinois Institute of Technology

Manufacturing Processes for Engineering Materials

ADDISON-WESLEY PUBLISHING COMPANY
Reading, Massachusetts
Menlo Park, California
London
Amsterdam
Don Mills, Ontario
Sydney

Sponsoring Editor: Thomas Robbins
Production Manager: Martha K. Morong
Production Editor: Laura Skinger
Designer: Catherine L. Dorin
Design Coordinator: Patricia O. Williams
Cover Designer: T. A. Philbrook
Art Editors: Susanah H. Michener and Marcia S. Strykowski
Copy Editor: Deborah A. Shukis
Illustrator: Parkway Illustrated Press
Manufacturing Supervisor: Ann E. DeLacey

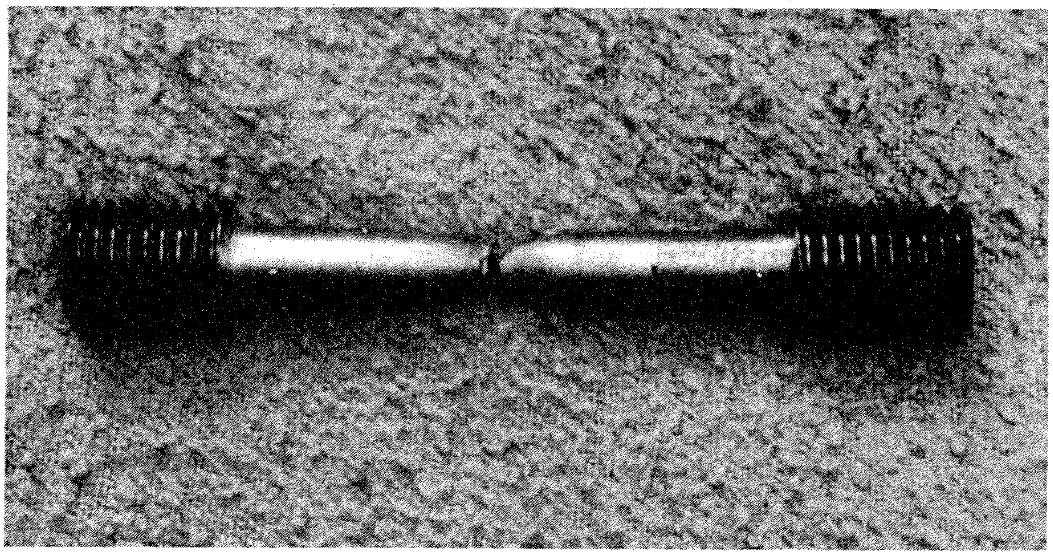
Library of Congress Cataloging in Publication Data

Kalpakjian, Serope, 1928–
Manufacturing processes for engineering materials.

Bibliography: p.
Includes index.
1. Manufacturing processes. I. Title.
TS183.K34 1984 670 83-15456
ISBN 0-201-11690-1

ABCDEFGHIJ—DO—8987654

2 Fundamentals of the Mechanical Behavior of Materials



2.1 INTRODUCTION

The manufacturing methods and techniques by which materials can be shaped into useful products were outlined in Chapter 1. In manufacturing, one of the most important groups of processes is *plastic deformation*, namely, shaping materials by applying forces in various ways (also known as deformation processing). It includes bulk deformation (forging, rolling, extrusion, rod and wire drawing) and sheet-forming processes (bending, drawing, spinning, and general pressworking).

This chapter deals with the fundamental aspects of the mechanical behavior of materials during deformation. Individual topics are deformation modes, stresses, forces, effects of rate of deformation and temperature, hardness, residual stresses, yield criteria, and various methods of analyzing metalworking processes in order to determine force and power requirements.

In stretching a piece of metal to make an object such as a fender of an automobile or a length of wire, the material is subjected to tension. A solid cylindrical piece of metal is forged in the making of a turbine disk, thus the material is subjected to compression. Sheet metal undergoes shearing stresses when, for instance, a hole is punched in it. A piece of plastic tubing is expanded by internal pressure to make a bottle, thus subjecting the material to tension in various directions.

In all of these processes, the material is subjected to one or more of the basic modes of deformation shown in Fig. 2.1, namely, tension, compression, and shear. The degree of deformation to which the material is subjected is defined as *strain*. For tension or compression the *engineering strain*, or *nominal strain*, is defined as

$$e = \frac{\ell - \ell_o}{\ell_o}. \tag{2.1}$$

We note that in tension the strain is positive and in compression it is negative. The *shear strain* is defined as

$$\gamma = \frac{a}{b} \tag{2.2}$$

In order to change the geometry of the bodies or elements in Fig. 2.1, forces must be applied to them as shown by the arrows. The determination of these forces as a function of strain is very important in manufacturing processes. The forces have to be known in order to design the proper equipment, to select the tool and die materials for proper strength, and to determine whether or not a specific metalworking operation can be accomplished on certain equipment.

Thus, the relation between a force and the deformation it produces is an essential parameter in manufacturing. In this relationship, strain is an *independent* variable, whereas the force required to cause it is a *dependent* variable. The basic material property tests that are performed to establish these relationships are covered in the sections below.

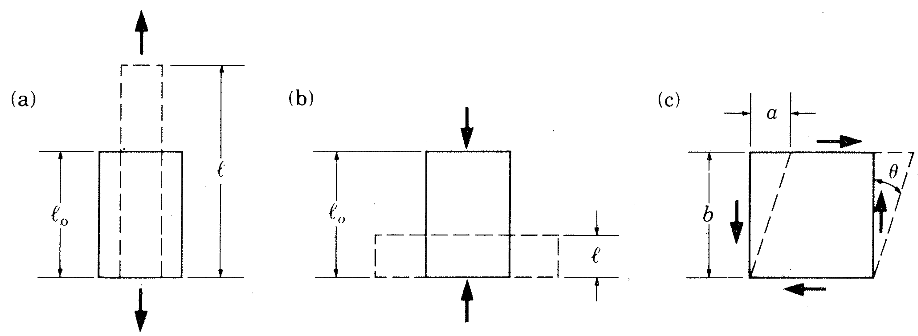


FIGURE 2.1 Types of strain. (a) Tensile. (b) Compressive. (c) Shear. All deformation processes in manufacturing involve strains of these types. Tensile strains are involved in stretching sheet metal to make car bodies, compressive strains in forging metals to make turbine disks, and shear strains in making holes by punching.

2.2 THE TENSION TEST

The tension test, because of its relative simplicity, is the most common test for determining the *strength-deformation characteristics* of materials. It involves the preparation of a test specimen (according to *ASTM* specifications) and testing it under tension on any of a variety of available testing machines.

The specimen (Fig. 2.2a) has an original length ℓ_o and an original cross-sectional area A_o. Although most specimens are solid and round, flat sheet or tubular specimens are also tested under tension. The original length is the distance between *gage marks* on the specimen and is generally 2 in. (50 mm). Longer lengths may be used for larger specimens such as structural members.

Typical results from a tension test are shown in Fig. 2.2(b). The *engineering stress*, or *nominal stress*, is defined as the ratio of the applied load to the original area,

$$\sigma = \frac{P}{A_o} \tag{2.3}$$

and the engineering strain is

$$e = \frac{\ell - \ell_o}{\ell_o}.$$

When the load is applied, the specimen elongates proportionately up to the *yield point* Y. If the load is removed within this range, the specimen returns to its original length. This is the range of *linear elastic behavior*. The *modulus of elasticity*, or *Young's modulus*, E is defined as

$$E = \frac{\sigma}{e}. \tag{2.4}$$

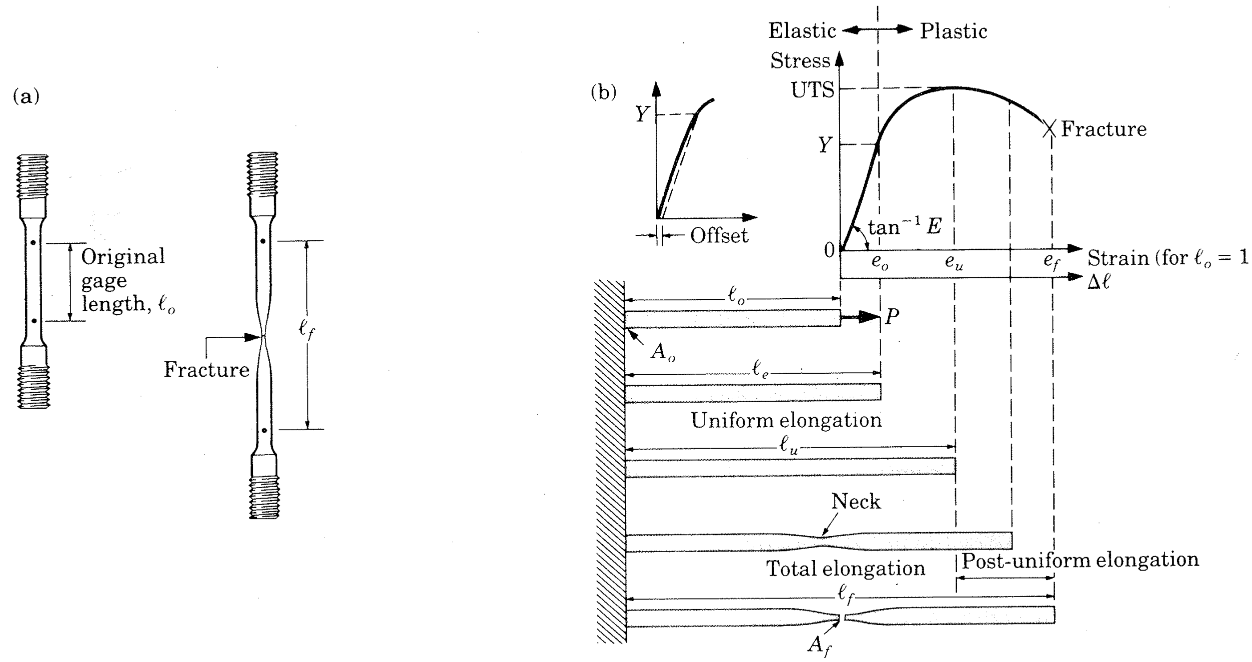


FIGURE 2.2 (a) Original and final shape of a standard tensile test specimen. (b) Outline of a tensile test sequence showing different stages in the elongation of the specimen. See also Fig. 10.6(a). Tensile test data are important in calculating forces and power requirements, and predicting the behavior of materials in manufacturing processes.

This linear relationship between stress and strain is known as *Hooke's law* (after R. Hooke, 1635–1703), the more generalized forms of which are given in Section 2.8.3.

The elongation of the specimen is accompanied by a contraction of its lateral dimensions. The absolute value of the ratio of the lateral strain to longitudinal strain is known as *Poisson's ratio*, ν, (after S. D. Poisson, 1781–1840). Some typical values for a variety of materials are given in Table 2.1.

The area under the stress–strain curve up to the yield point Y of a material is known as the *modulus of resilience*,

$$\text{Modulus of resilience} = \frac{Ye_o}{2} = \frac{Y^2}{2E}. \tag{2.5}$$

This area has the units of *energy per unit volume* and indicates the *specific energy* that the material can store elastically (Table 2.2).

With increasing load, the specimen begins to yield, that is, it begins to undergo *plastic (permanent) deformation* and the relationship between stress and strain is no longer linear. Since the rate of change in the slope of the stress–strain curve beyond the yield point is rather small for most materials, the determination of Y may be difficult.

The usual practice is to define the yield stress as the point on the curve that is offset by a strain of (usually) 0.2%, or 0.002 (Fig. 2.2b). Other offset strains may also

TABLE 2.1
MODULUS OF ELASTICITY *E* AND POISSON'S RATIO *ν* FOR VARIOUS MATERIALS

	E		
METALS	psi × 10^6	GPa	*ν*
Aluminum and its alloys	10–11.5	69–79	0.31–0.34
Cast irons	15–22	105–150	0.21–0.30
Copper and its alloys	15–21.5	105–150	0.33
Ductile iron	22–25	150–170	0.28
Lead and its alloys	2	14	0.43
Magnesium and its alloys	6–6.5	41–45	0.35
Malleable iron	26–27	180–185	0.27
Molybdenum	47	325	0.32
Nickel and its alloys	26–31	180–214	0.31
Steel (plain carbon)	29	200	0.33
Steel (austenitic stainless)	27.5–29	190–200	0.28
Tantalum and its alloys	21.5–27	150–186	0.35
Titanium and its alloys	11.5–19	80–130	0.31–0.34
Tungsten	58	400	0.27
NONMETALLIC			
Acrylics	0.2–0.5	1.4–3.4	0.35–0.40
Epoxies	0.5–2.5	3.5–17	0.34
Nylons	0.2–0.4	1.4–2.8	0.32–0.40
Rubbers	0.0015–0.015	0.01–0.1	0.5
Concrete	3–5	20–35	0.12–0.25
Glass and porcelain	10	70	0.24
Diamond	120–150	820–1050	
Graphite (bulk)	1	7	

TABLE 2.2
MODULUS OF RESILIENCE, MODULUS OF ELASTICITY, AND YIELD STRESS FOR VARIOUS METALS

	MODULUS OF RESILIENCE, in.-lb/in^3	*E*, psi × 10^6	*Y*, psi × 10^3
Copper, annealed	3	16	10
Lead	0.25	2	1
Magnesium	8.5	6	10
Medium-carbon steel, annealed	27.5	29	40
Spring steel	385	29	150
Titanium alloy, annealed	570	16	135

be used and should be specified in reporting the yield stress. The terms *elastic limit* and *proportional limit* can also be used to specify the point where the stress and strain are no longer proportional.

As the specimen continues to elongate under increasing load, its cross-sectional area decreases uniformly throughout its length. The load (hence the engineering stress) reaches a maximum and then begins to decrease. The maximum stress is known as the *tensile strength* or *ultimate tensile strength* (*UTS*) of the material.

When the specimen is loaded beyond its UTS, it begins to *neck* (Fig. 2.2) and the elongation is no longer uniform. That is, the change in the cross-sectional area of the specimen is no longer uniform along the length of the specimen but is concentrated locally in a "neck" formed in the specimen (known as *necking*, or *necking down*). As the test progresses, the engineering stress drops further and the specimen finally fractures in the necked region. The final stress level (marked by an × in Fig. 2.2b) is known as *breaking* or *fracture stress.*

2.2.1 DUCTILITY

Ultimate tensile strength is a practical measure of the overall strength of a material. Likewise, the strain at fracture is a measure of its *ductility*, that is, how large a strain the material withstands before failure. Note from Fig. 2.2(b) that until the UTS is reached, elongation is uniform. The strain up to the UTS is known as *uniform strain.* The elongation at fracture is known as the *total elongation.* This is measured between the original gage marks after the two pieces of the broken specimen are placed together, as shown in Fig. 2.2(a).

Two quantities commonly used to define ductility in a tension test are percent elongation and reduction of area. Percent *elongation* is defined as

$$\%\ \text{elongation} = \frac{\ell_f - \ell_o}{\ell_o} \times 100 \tag{2.6}$$

and is based on the total elongation.

Necking is a *local* phenomenon. If we put a series of gage marks at different points on the specimen, pull, and break it under tension and then calculate the percent elongation for each pair of gage marks, we find that with decreasing gage length the percent elongation increases (Fig. 2.3). The closest pair of gage marks undergo the largest elongation because they are closest to the necked region. However, the curves do not approach zero elongation with increasing gage length. This is because the specimens have all undergone a finite uniform and permanent elongation before fracture. It is thus important to report gage length in conjunction with elongation data, whereas other tensile properties are generally independent of gage length.

A second measure of ductility is *reduction of area*, defined as

$$\%\ \text{reduction of area} = \frac{A_o - A_f}{A_o} \times 100. \tag{2.7}$$

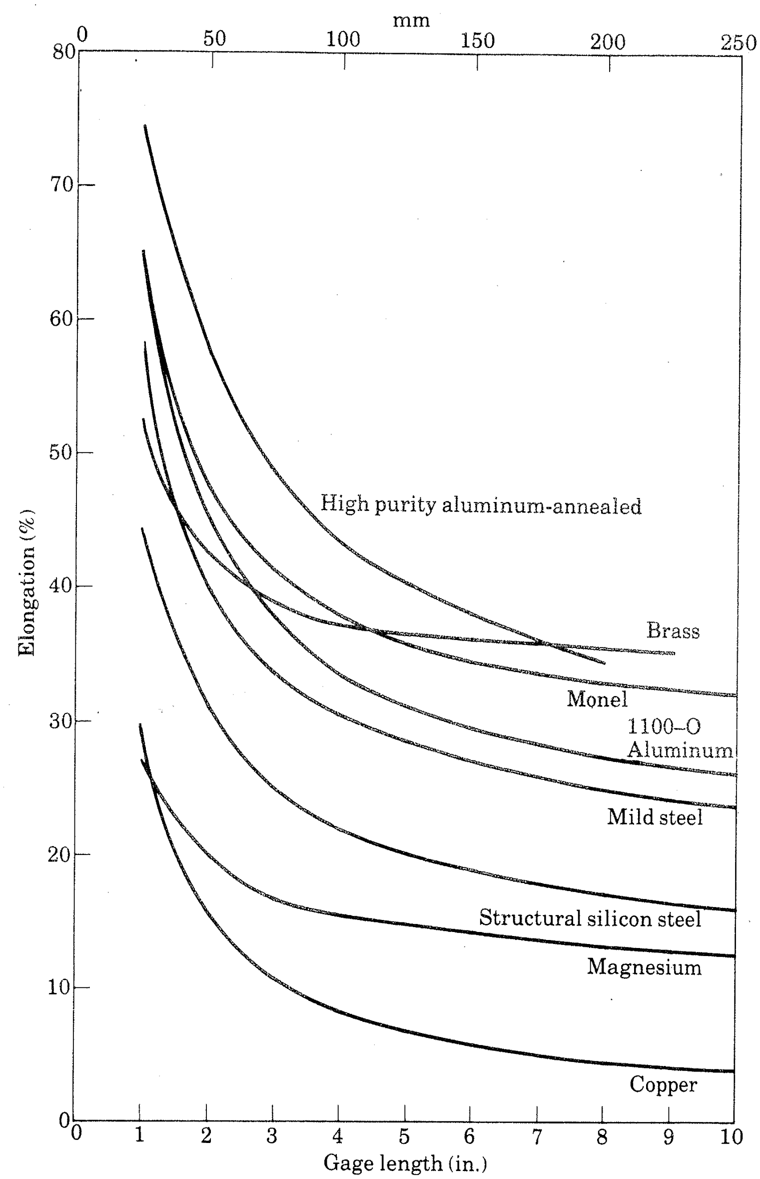


FIGURE 2.3 Total elongation in a tensile test as a function of original gage length for various metals. Because necking is a local phenomenon, elongation decreases with gage length. Standard gage length is usually 2 in. (50 mm), although shorter ones can be used if larger specimens are not available.

Thus, a material that necks down to a point at fracture, such as a glass rod at elevated temperature, has 100% reduction of area.

The elongation and reduction of area are generally related to each other for many common engineering metals and alloys, as shown in Fig. 2.4. Elongation ranges approximately between 10 and 60%, for most materials, while values between 20 and 90% are typical for reduction of area. *Thermoplastics* (Chapter 10) and *superplastic* materials (Section 2.2.7) exhibit much higher ductility. Brittle materials, by definition, have little or no ductility. Examples are glass at room temperature or a piece of chalk.

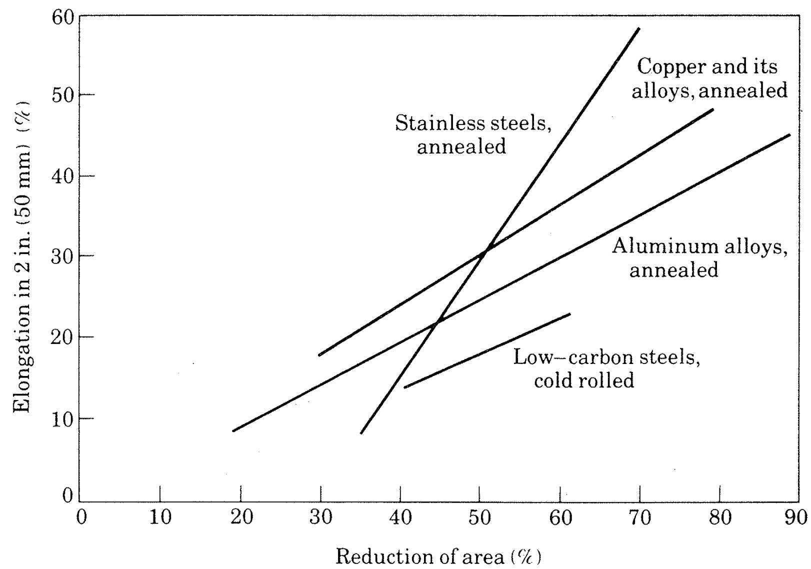


FIGURE 2.4 Approximate relationship between elongation and reduction of area for different groups of metals. Both elongation and reduction of area are measures of the ductility of a material.

● **Illustrative Problem 2.1**

A tension test specimen has an original gage length of 50 mm and a diameter of 4 mm. The maximum load during the test is 10 kN. The final gage length is 80 mm and the diameter of the necked region is 3 mm. Calculate the UTS, percent elongation, and reduction of area.

SOLUTION.

$$\text{UTS} = \frac{P}{A_o} = \frac{10{,}000}{\pi \frac{4^2}{4}} = 796 \text{ N/mm}^2 = 796 \text{ MPa}.$$

$$\text{Elongation} = \frac{80 - 50}{50} \times 100 = 60\%.$$

$$A_o = \pi \frac{4^2}{4} = 12.57 \text{ mm}^2.$$

$$A_f = \pi \frac{3^2}{4} = 7.07 \text{ mm}^2.$$

$$\text{Reduction of area} = \frac{12.57 - 7.07}{12.57} \times 100 = 43.8\%. \quad ●$$

2.2.2 TRUE STRESS AND TRUE STRAIN

It is apparent that, since stress is defined as the ratio of force to area, *true stress* should be defined as

$$\sigma = \frac{P}{A}, \tag{2.8}$$

where A is the actual (hence true) or instantaneous area supporting the load.

TABLE 2.3
COMPARISON OF ENGINEERING AND TRUE STRAINS IN TENSION

e	0.01	0.05	0.1	0.2	0.5	1	2	5	10
ϵ	0.01	0.049	0.095	0.18	0.4	0.69	1.1	1.8	2.4

Likewise, the complete tension test may be regarded as a series of incremental tension tests where, for each succeeding increment, the original specimen is a little longer than the previous one. We can now define *true strain* (or *natural* or *logarithmic strain*) ϵ as

$$\epsilon = \int_{\ell_o}^{\ell} \frac{d\ell}{\ell} = \ln\left(\frac{\ell}{\ell_o}\right). \tag{2.9}$$

Note that, for small values of engineering strain, $e = \epsilon$ since the value of $\ln(1 + e) = \epsilon$. For larger strains, however, the values diverge rapidly, as seen in Table 2.3.

The volume of a metal specimen in the plastic region of the test remains constant (Section 2.8.5). Hence, the true strain within the uniform elongation range can be expressed as

$$\epsilon = \ln\left(\frac{\ell}{\ell_o}\right) = \ln\left(\frac{A_o}{A}\right) = \ln\left(\frac{D_o}{D}\right)^2 = 2\ln\left(\frac{D_o}{D}\right). \tag{2.10}$$

Once necking begins, the true strain at any point in the specimen can be calculated from the change in cross-sectional area at that point. Thus, by definition, the largest strain is at the narrowest region of the neck.

We have seen that, at small strains, the engineering and true strains are very close and hence either one can be used in calculations. However, for the large strains encountered in metalworking, the true strain should be used. This is the true measure of the strain and can be illustrated by the following two examples.

Assume that a tension specimen is elongated to twice its original length. This involves a deformation equivalent to that of compressing a specimen to one half of its original height. Using the subscripts t and c for tension and compression, respectively, it is found that $\epsilon_t = 0.69$ and $\epsilon_c = -0.69$, whereas $e_t = 1$ and $e_c = -0.5$. Thus, true strain is a correct measure of strain.

For the second example, assume that a specimen 10 mm in height is compressed to a final thickness of zero. Thus, $\epsilon_c = -\infty$, whereas $e_c = -1$. The fact is that we have deformed the specimen infinitely, which is exactly what the value of the true strain indicates.

It is obvious from these examples that true strains are consistent with the actual physical phenomenon and that the engineering strains are not.

2.2.3 TRUE STRESS—TRUE STRAIN CURVES

The relation between engineering and true values for stress and strain can now be used to construct true stress–true strain curves from a curve such as that in Fig. 2.2.

The procedure for this construction is given in Illustrative Problem 2.2 below.

A typical true stress–true strain curve is shown in Fig. 2.5. For convenience, such a curve can be approximated by the equation

$$\sigma = K\epsilon^n. \tag{2.11}$$

Note that this equation indicates neither the elastic region nor the yield point Y of the material. These quantities, however, are easily available from the engineering stress–strain curve. (Since the strains at the yield point are very small, the difference between the true and engineering yield stress is negligible for metals. This is because the difference in the cross-sectional areas A_o and A at yielding is very small.)

When the curve in Fig. 2.5 is plotted on a log–log scale, it is composed of two straight lines. One straight line indicates the elastic region and the other straight line, with the slope n, indicates the plastic region, Fig. 2.6. The slope n is known as the *strain-hardening exponent*, and K (the true stress at a true strain of one) is known as the *strength coefficient*. The agreement between the data points and Eq. (2.11) may not always be as good as shown in Fig. 2.6; however, in most cases the curve can be drawn with a reasonably representative value of n.

Values of K and n for a variety of engineering materials are given in Table 2.4. The true stress–true strain curves for some materials are given in Fig. 2.7. Some differences between Table 2.4 and these curves may exist because of different sources of data and test conditions.

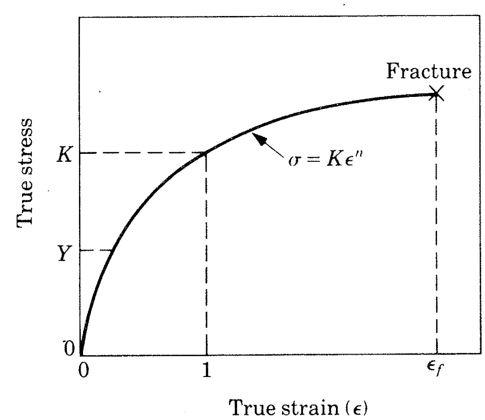


FIGURE 2.5 True stress-true strain curve in tension. Note that, unlike in an engineering stress-strain curve, the slope is always positive, and that the slope decreases with increasing strain. Whereas in the elastic range stress and strain are proportional, in the plastic range the curve can be approximated by the power expression shown in the figure.

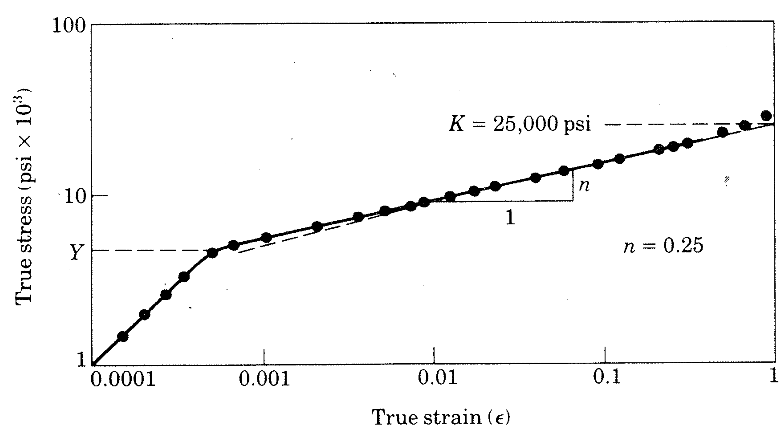


FIGURE 2.6 True stress-true strain curve in tension for 1100–O aluminum plotted on log-log scale. Note the large difference in the slopes in the elastic and plastic ranges. *Source*: After R. M. Caddell and R. Sowerby, *Bull. Mech. Eng. Educ.*, vol. 8, 1969, pp. 31–43.

TABLE 2.4
TYPICAL VALUES FOR *K* AND *n* IN EQ. (2.11) AT ROOM TEMPERATURE

	K psi × 10^3	*K* MPa	*n*
Aluminum, 1100-O	26	180	0.20
2024-T4	100	690	0.16
5052-O	30	210	0.13
6061-O	30	205	0.20
6061-T6	60	410	0.05
7075-O	58	400	0.17
Brass, 60-39-1 Pb, annealed	115	800	0.33
70-30, annealed	130	895	0.49
85-15, cold-rolled	84	580	0.34
Bronze (phosphor), annealed	105	720	0.46
Cobalt-base alloy, heat-treated	300	2070	0.50
Copper, annealed	46	315	0.54
Molybdenum, annealed	105	725	0.13
Muntz metal, annealed	115	800	0.50
Steel, low-carbon annealed	77	530	0.26
1045 hot-rolled	140	965	0.14
1112 annealed	110	760	0.19
1112 cold-rolled	110	760	0.08
4135 annealed	147	1015	0.17
4135 cold-rolled	160	1100	0.14
4340 annealed	93	640	0.15
17-4 P-H annealed	175	1200	0.05
52100 annealed	210	1450	0.07
302 stainless, annealed	190	1300	0.30
304 stainless, annealed	185	1275	0.45
410 stainless, annealed	140	960	0.10
Vanadium, annealed	112	770	0.35

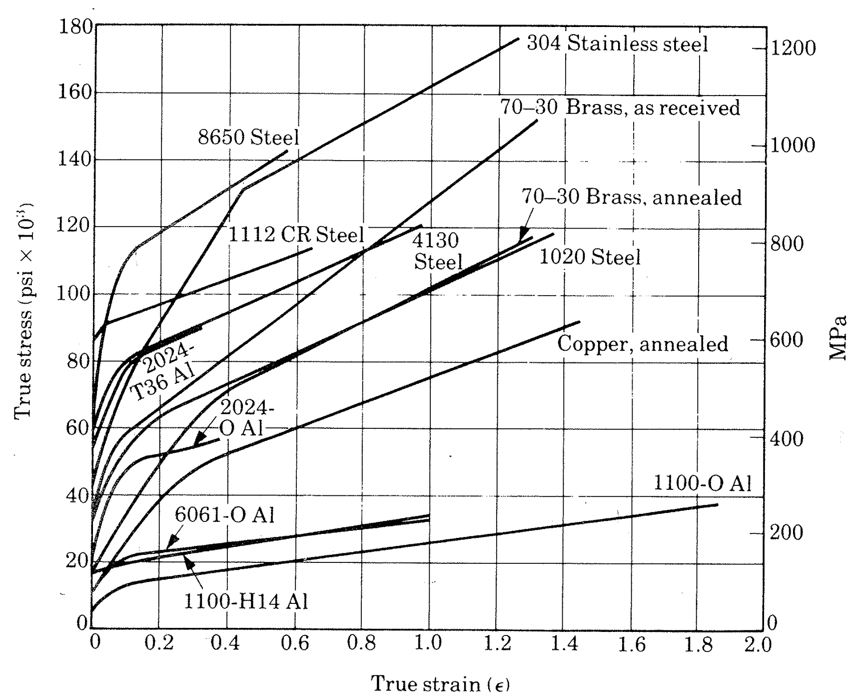


FIGURE 2.7 True stress-true strain curves in tension at room temperature for various metals. The point of intersection of each curve at the ordinate is the yield stress *Y*; thus the elastic portions of the curves are not indicated. Compare with Fig. 2.5. The area under the curve is toughness. See also Fig. 10.8. *Source*: S. Kalpakjian.

The area under the true stress–true strain curve is known as *toughness* and can be expressed as

$$\text{Toughness} = \int_o^{\epsilon_f} \sigma \, d\epsilon, \tag{2.12}$$

where ϵ_f is the true strain at fracture. Toughness is defined as the energy per unit volume (specific energy) that has been dissipated up to fracture.

It is important to remember that this specific energy pertains only to that volume of material at the narrowest region of the neck. Any volume of material away from the neck has undergone less strain and, hence, has dissipated less energy.

It should also be pointed out that here toughness is different from the concept of fracture toughness as treated in textbooks on fracture mechanics. Fracture mechanics (the study of the initiation and propagation of cracks in a solid medium) is beyond the scope of this text. It is of limited relevance to metalworking processes, except in die design and die life.

● Illustrative Problem 2.2

The following data are taken from a stainless steel tension test specimen:

$A_o = 0.056 \text{ in}^2$

$A_f = 0.016 \text{ in}^2$

$\ell_o = 2 \text{ in.}$

Elongation in 2 in. = 49%

LOAD, P, lb	EXTENSION, $\Delta\ell$, in.
1600	0
2500	0.02
3000	0.08
3600	0.20
4200	0.40
4500	0.60
4600 (max.)	0.86
3300 (fracture)	0.98

Draw the true stress–true strain curve for the material.

SOLUTION. The load–extension curve is shown in Fig. P2.1. In order to determine the true stress and true strain, the following relationships are used:

$$\text{True stress } \sigma = \frac{P}{A}$$

$$\text{True strain } \epsilon = \ln\left(\frac{\ell}{\ell_o}\right) \text{ up to necking}$$

$$\epsilon_f = \ln\left(\frac{A_o}{A_f}\right) \text{ at fracture}$$

$$\ell = \ell_o + \Delta\ell \text{ up to necking}$$

Assuming that the volume of the specimen remains constant, we have

$$A_o \ell_o = A\ell.$$

Thus,

$$A = \frac{A_o \ell_o}{\ell} = \frac{(0.056)(2)}{\ell} = \frac{0.112}{\ell} \text{ in}^2.$$

Using these relationships, we obtain the following data:

$\Delta\ell$, in.	ℓ, in.	ϵ	A, in^2	TRUE STRESS, psi
0	2.00	0	0.056	28,600
0.02	2.02	0.01	0.0555	45,000
0.08	2.08	0.039	0.054	55,800
0.20	2.20	0.095	0.051	70,800
0.40	2.40	0.182	0.047	90,000
0.60	2.60	0.262	0.043	104,500
0.86	2.86	0.357	0.039	117,300
0.98	2.98	1.253	0.016	206,000

The true stress and true strain are plotted (solid line) in Fig. P2.2. The point at necking

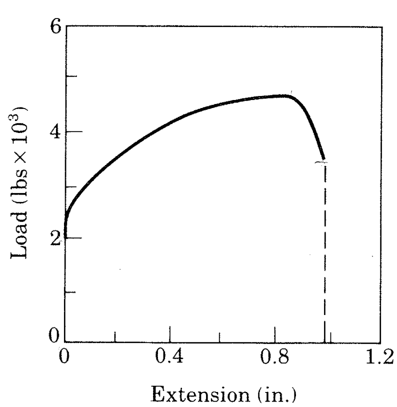


FIGURE P2.1

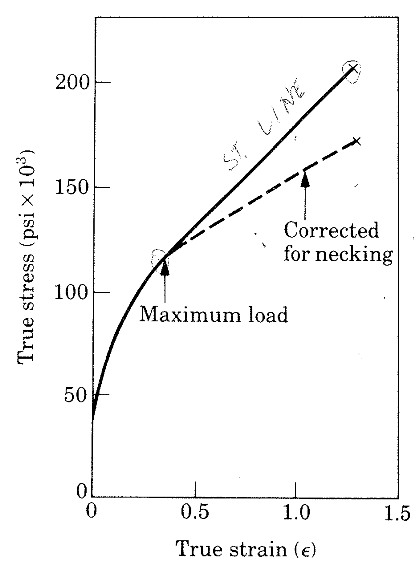


FIGURE P2.2

is connected to the point at fracture by a straight line because there are no data on the instantaneous areas after necking begins. The correction on this curve (broken line) is explained in Section 2.8. ●

2.2.4 INSTABILITY IN THE TENSION TEST

As noted previously the onset of necking in a tension test corresponds to the ultimate tensile strength (UTS) of the material. The slope of the load–elongation curve at this point is zero, and it is here that instability begins. That is, the specimen begins to neck and cannot support the load because the neck is becoming smaller in cross-sectional area.

Using the following relationships

$$\epsilon = \ln\left(\frac{A_o}{A}\right),$$

$$A = A_o e^{-\epsilon},$$

and

$$P = \sigma A = \sigma A_o e^{-\epsilon},$$

one finds that at necking,

$$\frac{dP}{d\epsilon} = A_o\left(\frac{d\sigma}{d\epsilon} e^{-\epsilon} - \sigma e^{-\epsilon}\right) = 0.$$

Hence,

$$\frac{d\sigma}{d\epsilon} = \sigma.$$

However,

$$\sigma = K\epsilon^n,$$

and consequently,

$$\begin{aligned} nK\epsilon^{n-1} &= K\epsilon^n \\ \epsilon &= n. \end{aligned} \tag{2.13}$$

Thus, the true strain at the onset of necking (termination of uniform elongation) is numerically equal to the strain hardening exponent, n. Hence, the higher the value of n, the greater the strain to which a piece of material can be stretched before necking begins.

It can be seen from Table 2.4 that metals such as annealed copper, brass, and stainless steel can be stretched uniformly to a greater extent than the other materials listed. These observations are covered in greater detail in Chapters 7 and 10 because of their relevance to sheet-forming processes.

Instability in a tension test can be viewed as a phenomenon where two competing processes are simultaneously taking place. As the load on the specimen is increased, its cross-sectional area becomes smaller, this being more pronounced in the region where necking begins. On the other hand, with increasing strain, the material is becoming stronger due to strain hardening. Since the load on the specimen is the product of area and strength, instability sets in when the rate of decrease in area is greater than the rate of increase in strength. This is also known as *geometric softening*.

● **Illustrative Problem 2.3**

A material has a true stress–true strain curve given by

$$\sigma = 100{,}000\epsilon^{0.5} \text{ psi.}$$

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

SOLUTION. Since the necking strain corresponds to the maximum load and the necking strain for this material is

$$\epsilon = n = 0.5,$$

we have as the true ultimate tensile strength

$$\sigma = Kn^n$$

$$\sigma = 100{,}000\,(0.5)^{0.5} = 70{,}710 \text{ psi.}$$

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

$$P = \sigma A = \sigma A_0 e^{-0.5},$$

where σ is the true ultimate tensile strength. Hence,

$$P = (70{,}710)(0.606)(A_o) = 42{,}850A_o \text{ lb.}$$

Since UTS $= P/A_o$,

$$\text{UTS} = 42{,}850 \text{ psi.}$$ ●

2.2.5 TYPES OF STRESS–STRAIN CURVES

Every material has a differently shaped stress–strain curve. Its shape depends on its composition and many other factors to be treated in detail later. Some of the major types of curves are shown in Fig. 2.8.

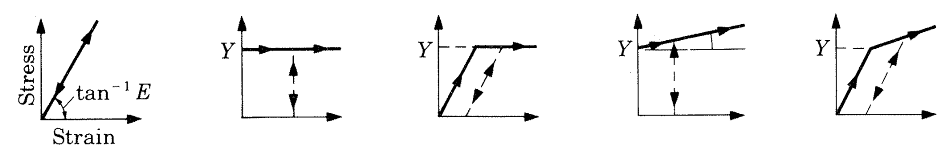


FIGURE 2.8 Schematic illustration of various types of idealized stress-strain curves. (a) Perfectly elastic. (b) Rigid, perfectly plastic. (c) Elastic, perfectly plastic. (d) Rigid, linearly strain-hardening. (e) Elastic, linearly strain hardening. The broken lines and arrows indicate unloading and reloading during the test. Most engineering metals exhibit a behavior similar to curve (e). See also Fig. 10.9.

A perfectly elastic material behaves like a spring with stiffness E. The behavior of brittle materials, such as glass, ceramics, and some cast irons, may be represented by such a curve (Fig. 2.8a). There is a limit to the stress the material can sustain, after which it breaks.

A rigid, perfectly plastic material has, by definition, an infinite value of E. Once the stress reaches the yield stress Y, it continues to undergo deformation at the same stress level. When the load is released, the material has undergone permanent deformation, with no elastic recovery, as seen in Fig. 2.8(b).

An elastic, perfectly plastic material is a combination of the first two; it undergoes elastic recovery when the load is released (Fig. 2.8c).

A rigid, linearly strain-hardening material requires an increasing stress level to undergo further strain. Thus, its *flow stress* (magnitude of the stress required to maintain plastic deformation at a given strain) increases with increasing strain. It has no elastic recovery upon unloading (Fig. 2.8d). See Fig. 2.51 for flow stress.

An elastic, linearly strain-hardening curve (Fig. 2.8e) is an approximation of the behavior of most engineering materials, with the modification that the plastic portion of the curve has a decreasing slope with increasing strain (i.e., Fig. 2.5).

Remember that the slopes of the elastic portions of the models in Fig. 2.8 are highly exaggerated and that the initial slopes are actually very steep. This can be seen by noting the magnitude of the modulus of elasticity of engineering materials. Also note that some of these curves can be expressed by Eq. (2.11) by changing the value of n (Fig. 2.9) or by other equations of a similar nature.

2.2.6 EFFECTS OF TEMPERATURE

In this and subsequent sections, the various factors that have an influence on the shape of stress–strain curves will be discussed. The first factor is temperature. Although somewhat difficult to generalize, increasing temperature usually increases ductility and toughness and lowers the modulus of elasticity, yield stress, and ultimate tensile strength. These effects are shown in Figs. 2.10 and 2.11.

Temperature also affects the strain-hardening exponent n of most metals, in that n decreases with increasing temperature (Fig. 2.12). Depending on the type of material and its composition and level of impurities, elevated temperatures can have other

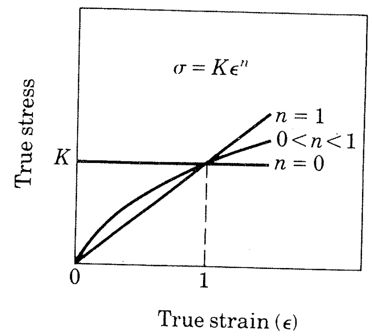


FIGURE 2.9 The effect of strain-hardening exponent n on the shape of true stress-true strain curves. When $n = 1$, the material is elastic, and when $n = 0$ it is rigid, perfectly plastic. See Table 2.4 for n values.

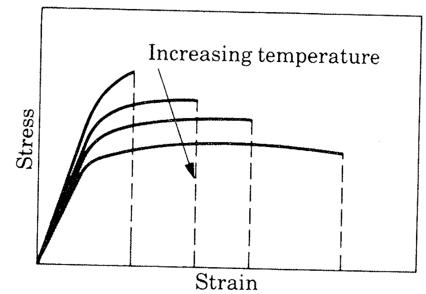


FIGURE 2.10 Typical effects of temperature on engineering stress-strain curves. Temperature affects the modulus of elasticity, yield stress, ultimate tensile strength, and toughness of materials. See also Fig. 10.12.

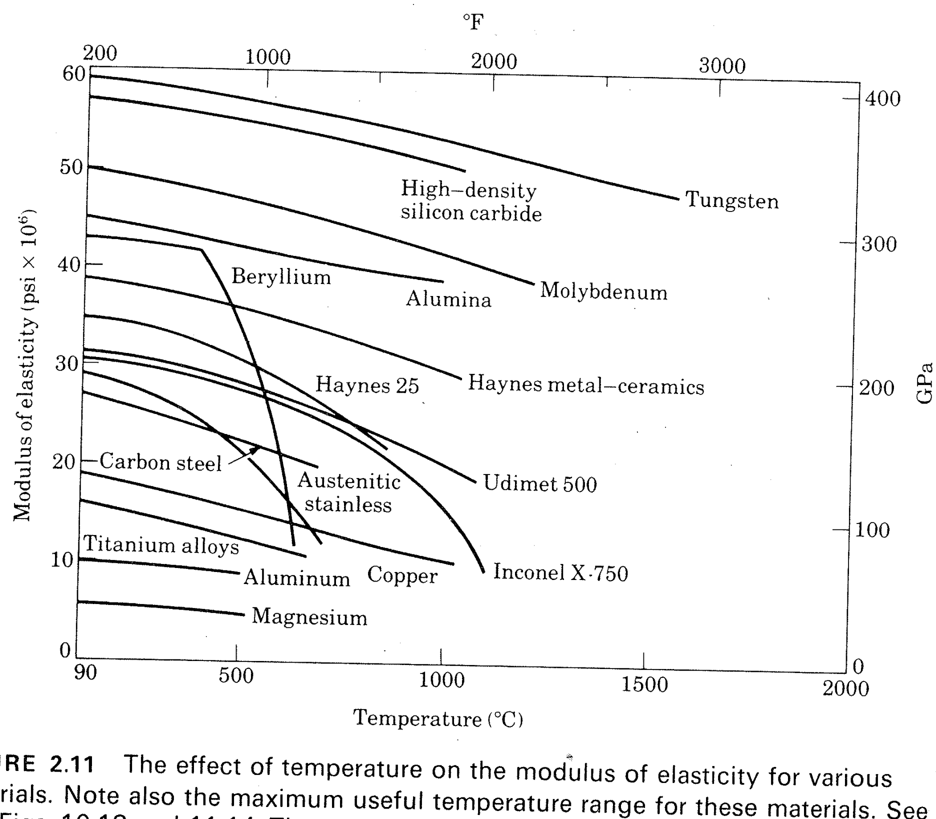


FIGURE 2.11 The effect of temperature on the modulus of elasticity for various materials. Note also the maximum useful temperature range for these materials. See also Figs. 10.13 and 11.14. The properties of these materials are described in Chapter 3.

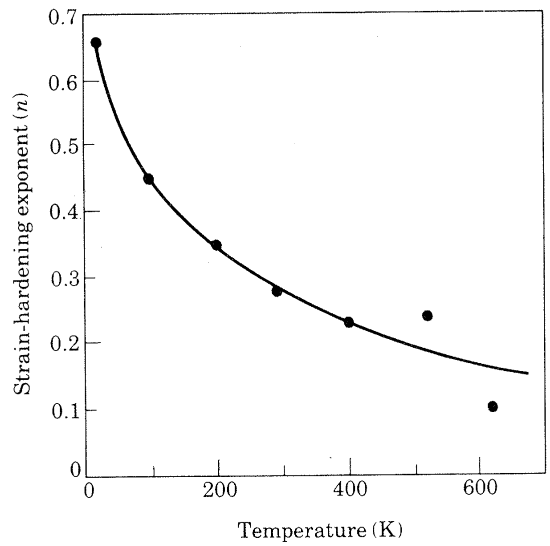


FIGURE 2.12 The effect of temperature on the strain-hardening exponent n for pure aluminum. *Source*: After R. P. Carreker and W. R. Hibbard, Jr., *Trans. TMS-AIME*, vol. 209, 1957, pp. 1157–1163.

significant effects, as detailed in Chapter 3. However, the influence of temperature is best discussed in conjunction with strain rate for the reasons explained below.

2.2.7 STRAIN RATE AND ITS EFFECTS

Depending on the particular manufacturing operation and equipment, a piece of material may be formed at low or at high speeds. In performing a tension test, the specimen can be strained at different rates to simulate the actual deformation process.

Whereas the *deformation rate* may be defined as the speed (in feet per minute or meters per second, for instance) at which a tension test is being carried out (e.g., the rate at which a rubber band is stretched), the *strain rate* is a function of the geometry of the specimen.

The *engineering strain rate* $\dot{e}$ is defined as

$$\dot{e} = \frac{de}{dt} = \frac{d\left(\frac{\ell - \ell_o}{\ell_o}\right)}{dt} = \frac{1}{\ell_o} \cdot \frac{d\ell}{dt} = \frac{v}{\ell_o} \tag{2.14}$$

and the *true strain rate* $\dot{\epsilon}$ as

$$\dot{\epsilon} = \frac{d\epsilon}{dt} = \frac{d\left(\ln\left(\frac{\ell}{\ell_o}\right)\right)}{dt} = \frac{1}{\ell} \cdot \frac{d\ell}{dt} = \frac{v}{\ell}, \tag{2.15}$$

where v is the speed of deformation (e.g., the speed of the jaws of the testing machine in which the specimen is clamped).

Although the deformation rate v and engineering strain rate $\dot{e}$ are equivalent, the true strain rate $\dot{\epsilon}$ is not identical to v. Thus, in a tension test with v constant, the true strain rate decreases as the specimen becomes longer. In order to maintain a constant $\dot{\epsilon}$, the speed must be increased accordingly. However, for small changes in length of the specimen during a test, this difference is not significant.

Typical deformation speeds employed in various metalworking processes, and the strain rates involved, are shown in Table 2.5. Note that there are considerable differences in the magnitudes. Because of this wide range, strain rates are generally quoted in orders of magnitude, such as 10^2 per second, 10^4 per second, etc.

The typical effects of temperature and strain rate on the strength of metals are shown in Figs. 2.13 and 2.14. These figures clearly indicate that increasing strain rate increases strength, and that the sensitivity of the strength to the strain rate increases with temperature. Note, however, that for the materials in Fig. 2.13 this effect is relatively small at room temperature; this may not be true for other metals and alloys. For instance, the yield stress at room temperature in Fig. 2.14 shows considerable sensitivity to strain rate.

It can be seen from Fig. 2.13 that the same strength can be obtained either at low temperature, low strain rate or at high temperature, high strain rate. This is important in estimating the resistance of materials to deformation when processing them at various strain rates and temperatures.

The effect of strain rate on strength also depends on the particular level of strain; this effect increases with strain (Fig. 2.15). The strain rate has also been found to affect the strain hardening exponent n in that it decreases with increasing strain rate (Fig. 2.16).

The effect of strain rate on the strength of materials is generally expressed by

$$\sigma = C\dot{\epsilon}^m, \tag{2.16}$$

where C is the *strength coefficient*, similar to K in Eq. (2.11), and m is the *strain-rate sensitivity exponent* of the material.

Equation (2.16) can be plotted in a manner similar to Fig. 2.6, with temperature as a parameter, as seen in Fig. 2.13. Note that the value of m increases with increasing temperature (Fig. 2.17).

A general range of values for m is as follows: up to 0.05 for cold working, 0.05 to 0.4 for hot working, and 0.3 to 0.85 for superplastic materials. For a Newtonian fluid,

TABLE 2.5
TYPICAL RANGES OF STRAIN, DEFORMATION SPEED, AND STRAIN RATES IN METALWORKING PROCESSES

PROCESS	TRUE STRAIN	DEFORMATION SPEED, m/s	STRAIN RATE, s^{-1}
Cold working			
Forging, rolling	0.1–0.5	0.1–100	1–10^3
Wire and tube drawing	0.05–0.5	0.1–100	1–10^4
Explosive forming	0.05–0.2	10–100	10–10^5
Hot working and warm working			
Forging, rolling	0.1–0.5	0.1–30	1–10^3
Extrusion	2–5	0.1–1	10^{-1}–10^2
Machining	1–10	0.1–100	10^3–10^6
Sheet-metal forming	0.1–0.5	0.05–2	1–10^2
Superplastic forming	0.2–3	10^{-4}–10^{-2}	10^{-4}–10^{-2}

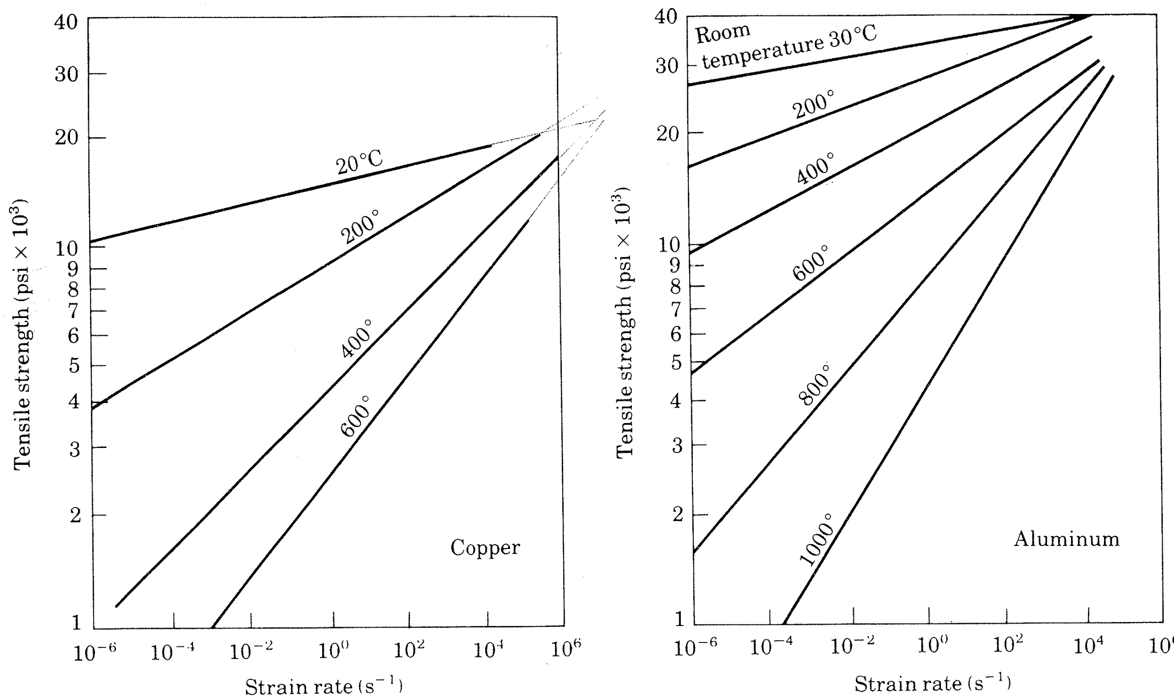


FIGURE 2.13 The effect of strain rate on the ultimate tensile strength of copper and aluminum. Note that as temperature increases, the slope increases. Thus, tensile strength becomes more and more sensitive to strain rate as temperature increases. *Source*: After J. H. Hollomon.

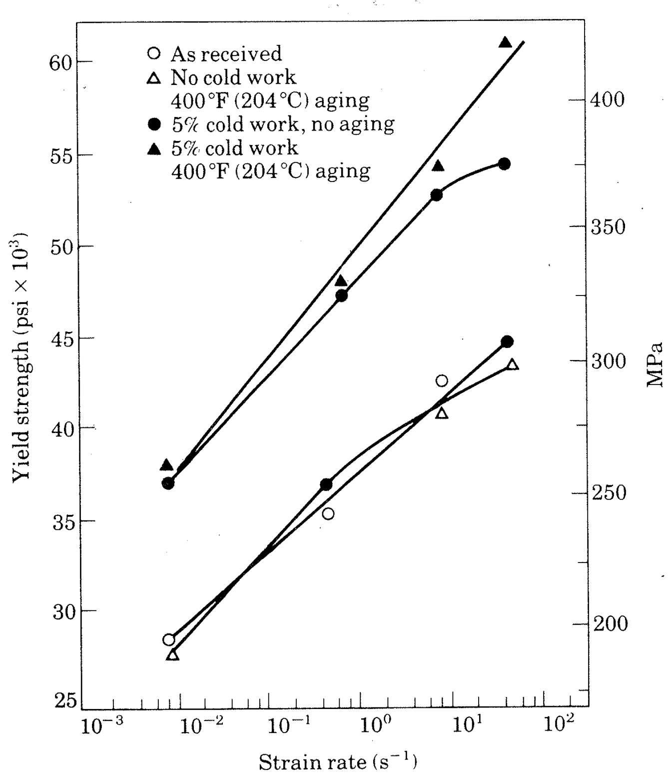


FIGURE 2.14 The effect of strain rate on the yield strength Y for cold-rolled aluminum-killed steel (Section 5.5). Note the wide range of strain rates. *Source*: After A. Saxena and D. A. Chatfield, SAE Paper 760209, 1976.

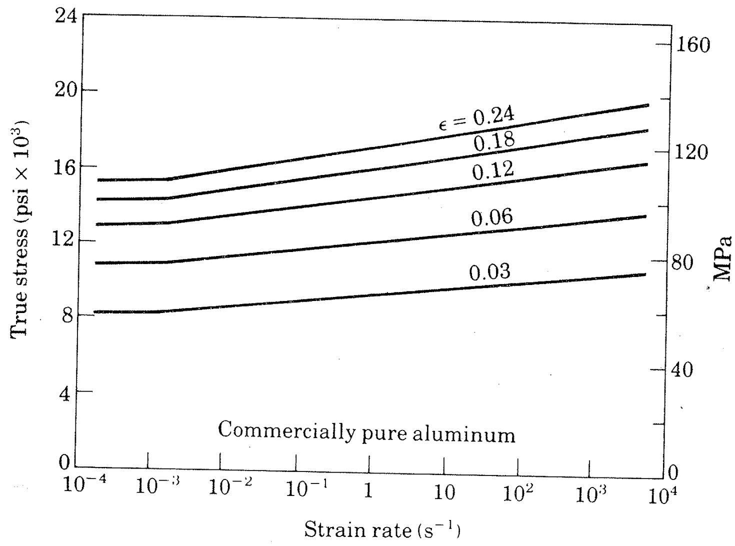


FIGURE 2.15 True stress as a function of strain rate at various strains, for commercially pure aluminum. The true stress required to sustain plastic deformation at a particular strain is known as the flow stress. *Source*: After U. S. Lindholm, *Behavior of Materials Under Dynamic Loading*, ASME, 1965, p. 51.

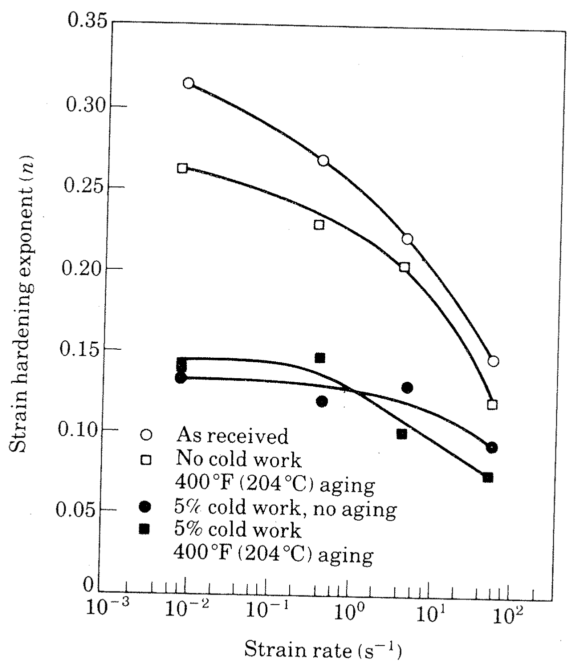


FIGURE 2.16 The effect of strain rate on the strain hardening exponent n for cold-rolled rimmed steel (Section 5.5). *Source*: After A. Saxena and D. A. Chatfield, SAE Paper 760209, 1976.

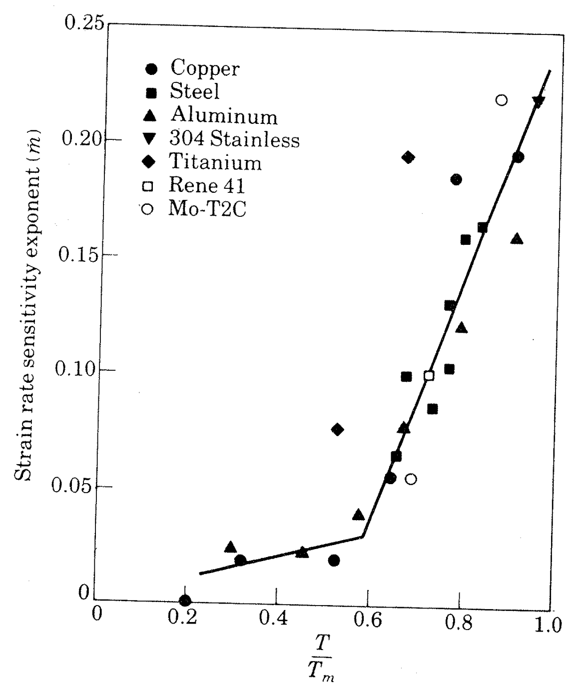


FIGURE 2.17 Dependence of the strain-rate sensitivity exponent m on the homologous temperature T/T_m for various materials. T is the testing temperature and T_m is the melting point of the metal, both on the absolute scale. The transition in the slopes of the curve occurs at about the recrystallization temperature of the metals. See also Fig. 3.17. *Source*: After F. W. Boulger, DMIC Report 226, Battelle Mem. Inst., 1966, pp. 13–37.

TABLE 2.6
APPROXIMATE RANGE OF VALUES FOR *C* AND *m* IN EQ: (2.16) FOR VARIOUS ANNEALED METALS AT TRUE STRAINS RANGING FROM 0.2 TO 1.0. (After T. Altan and F. W. Boulger, *Jr. Eng. Ind.* 95:1009–1019, 1973.)

MATERIAL	TEMPERATURE, °C	*C* psi × 10^3	*C* MPa	*m*
Aluminum	200–500	12–2	82–14	0.07–0.23
Aluminum alloys	200–500	45–5	310–35	0–0.20
Copper	300–900	35–3	240–20	0.06–0.17
Copper alloys (brasses)	200–800	60–2	415–14	0.02–0.3
Lead	100–300	1.6–0.3	11–2	0.1–0.2
Magnesium	200–400	20–2	140–14	0.07–0.43
Steel				
Low-carbon	900–1200	24–7	165–48	0.08–0.22
Medium-carbon	900–1200	23–7	160–48	0.07–0.24
Stainless	600–1200	60–5	415–35	0.02–0.4
Titanium	200–1000	135–2	930–14	0.04–0.3
Titanium alloys	200–1000	130–5	900–35	0.02–0.3
Ti-6Al-4V*	815–930	9.5–1.6	65–11	0.50–0.80
Zirconium	200–1000	120–4	830–27	0.04–0.4

* At a strain rate of 2×10^{-4}/s.

Note: As temperature increases, *C* decreases and *m* increases. As strain increases, *C* increases and *m* may increase or decrease, or it may become negative within certain ranges of temperature and strain.

where the shear stress increases linearly with rate of shear, the value of m is 1. Some specific values for C and m are given in Table 2.6.

The term *superplastic* refers to the capability of some materials to undergo large uniform elongation prior to failure. The elongation may be on the order of a few hundred to as much as 2000%. An extreme example is hot glass. Among other materials exhibiting superplastic behavior are polymers at elevated temperatures, very-fine-grain alloys of zinc-aluminum and titanium alloys.

The magnitude of m has a significant effect on necking in a tension test. Experimental observations show that, with high m values, the material stretches to a greater length before it fails. This is an indication that necking is delayed with increasing m. When necking is about to begin (i.e., the cross-sectional area of the specimen in the neck region becomes smaller) its strength, with respect to the rest of the specimen, increases due to strain hardening. However, the strain rate in the neck region is also higher than the rest of the specimen because the material is elongating faster there. Since the material in the necked region is becoming stronger as it is strained at a higher rate, this region exhibits a greater resistance to necking. This increased resistance to necking thus depends on the magnitude of m.

As the test progresses, necking becomes more diffuse and the specimen becomes longer before fracture. Thus, total elongation increases with increasing m value (Fig. 2.18). As expected, the elongation after necking (postuniform elongation) also

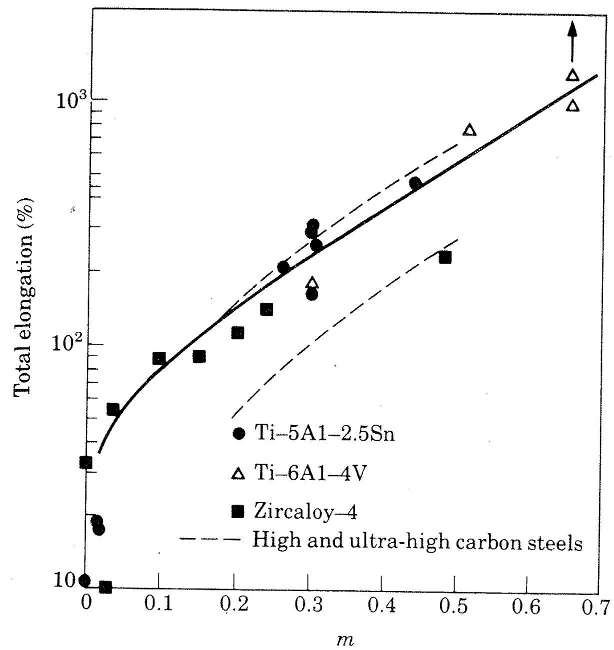


FIGURE 2.18 The effect of strain-rate sensitivity exponent m on the total elongation for various metals. Note that elongation at high values of m approaches 1000%. This phenomenon is utilized in superplastic forming of metals, described in Section 7.13. *Source*: After D. Lee and W. A. Backofen, *Trans. TMS-AIME*, vol. 239, 1967, p. 1034.

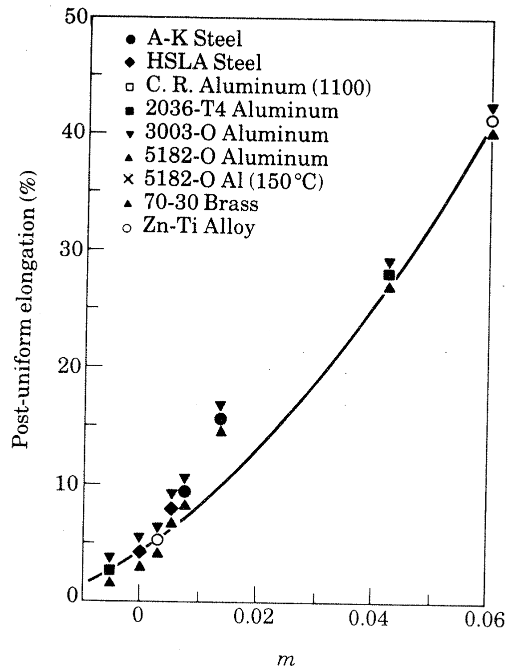


FIGURE 2.19 The effect of strain-rate sensitivity exponent m on the post-uniform (after necking) elongation for various metals. See Fig. 2.2(b). *Source*: After A. K. Ghosh, *Jr. Eng. Mat. Tech.*, vol. 99, 1977, pp. 264–274.

increases with increasing m (Fig. 2.19). It has also been observed that the value of m decreases with metals of increasing strength (Fig. 2.20).

The effect of strain rate on ductility is difficult to generalize, as can be seen from Fig. 2.21. Some of these phenomena are explained in Chapter 3. Since the formability of materials depends largely on their ductility, it is important to recognize the effect of temperature and strain rate on ductility. Generally, higher strain rates have an adverse effect on the ductility of materials. (The increase in ductility due to the strain-rate sensitivity of materials has been exploited in superplastic forming, as will be described in Section 7.13.)

2.2.8 EFFECTS OF HYDROSTATIC PRESSURE

Note that the tests described thus far have been carried out at ambient pressure. Tests have been performed under hydrostatic conditions, with pressures ranging up to 10^5 psi (10^3 MPa). Important observations have been made concerning the effects of high hydrostatic pressure on the behavior of materials:

a. Hydrostatic pressure substantially increases the strain at fracture (Fig. 2.22).

b. It has little or no effect on the shape of the true stress–true strain curve, but only extends it (Fig. 2.23).

c. It has no effect on the strain or the maximum load at which necking begins.

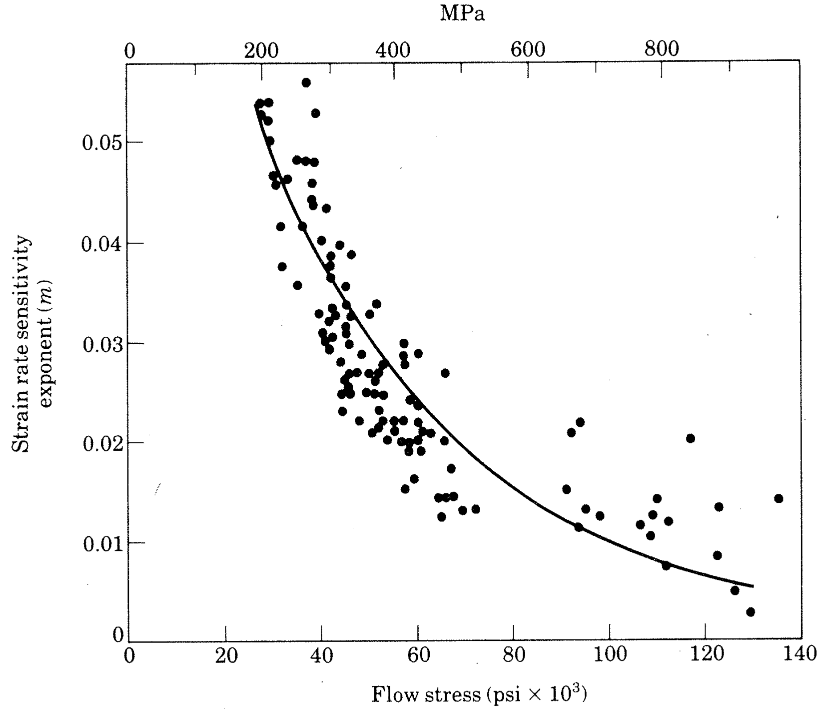


FIGURE 2.20 Relationship between flow stress for hot- and cold-rolled low-carbon steels and their *m* value at room temperature. Stronger steels are less sensitive to strain rate. Flow stress in this figure is the same as the uniaxial yield stress. *Source*: After A. Saxena and D. A. Chatfield, SAE Paper 760209, 1976.

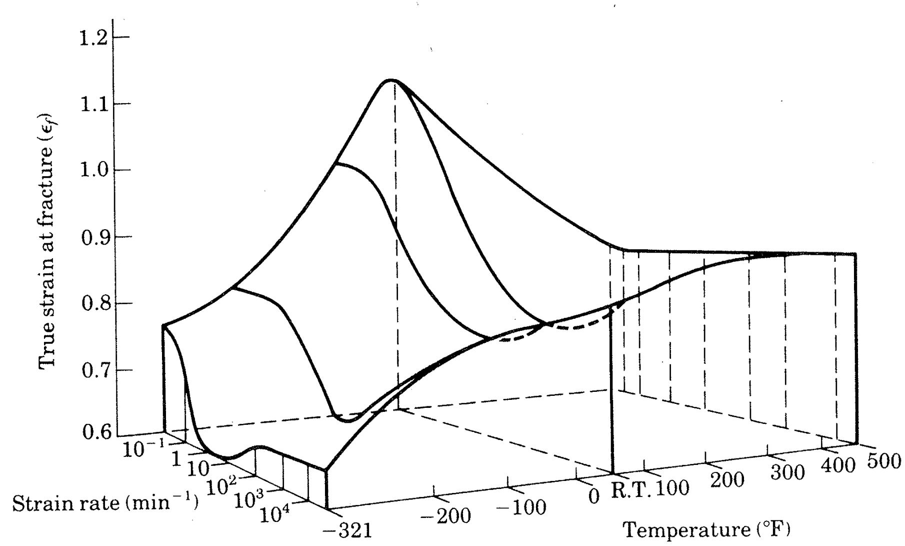


FIGURE 2.21 True strain at fracture for 303 stainless steel as a function of strain rate and temperature. This curve shows that it is difficult to generalize the combined effects of temperature and strain rate on the ductility of metals. *Source*: After G. W. Form and W. M. Baldwin, Jr., *Trans. ASM,* vol. 48, 1956, pp. 474–485.

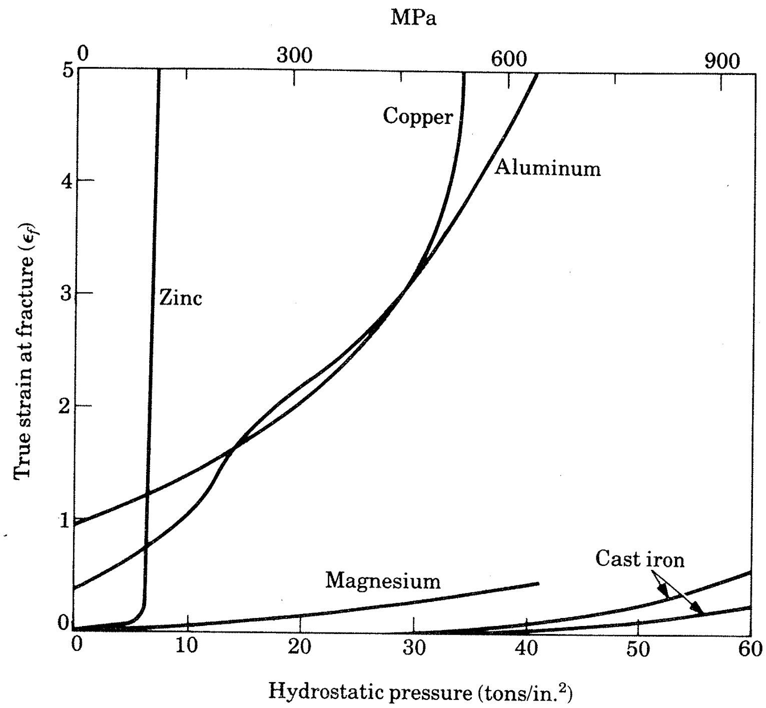


FIGURE 2.22 The effect of hydrostatic pressure on true strain at fracture in tension for various metals. Even cast iron becomes ductile under high pressure. *Source*: After H. Ll. D. Pugh and D. Green, *Proc. Inst. Mech. Eng.*, vol. 179, Part 1, No. 12, 1964–1965.

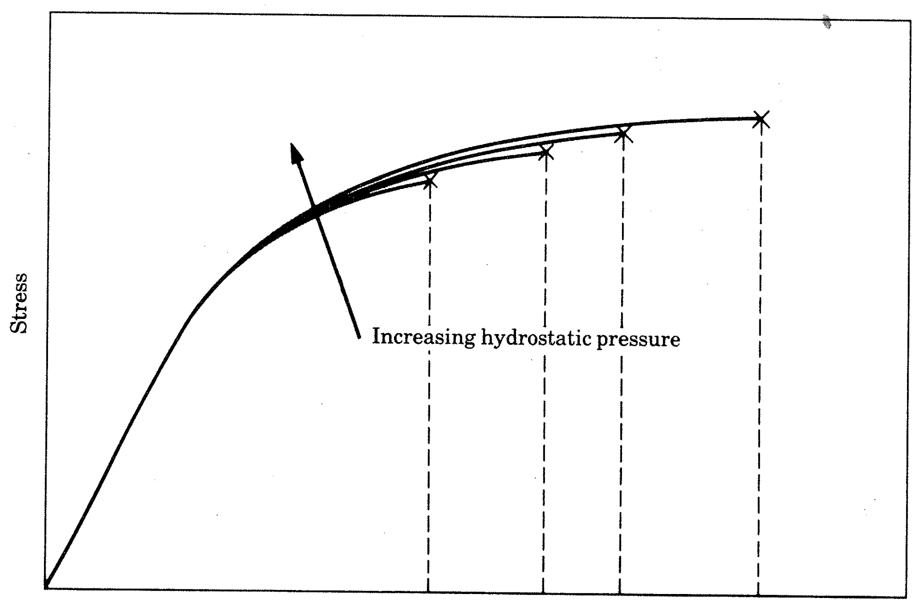


FIGURE 2.23 Schematic illustration of the effect of hydrostatic pressure on true stress-true strain curves. Note that pressure does not change the shape of the curve, it simply extends it. Ductility (true strain at fracture) and toughness of metals increase substantially with hydrostatic pressure, an observation that is significant in metalworking processes. See, for instance, Section 6.16.4.

The increase in ductility due to hydrostatic pressure has also been observed in other tests, such as compression and torsion. This increase in ductility has been observed not only with ductile metals, but also with brittle metals and nonmetallic materials. Materials such as cast iron, marble, and various rocks have been found to acquire some ductility (or increase in ductility), and thus deform plastically, when subjected to hydrostatic pressure. The level of the pressure required to enhance ductility depends on the material.

Experiments have shown that generally the mechanical properties of metals are not altered after being subjected to hydrostatic pressure.

2.2.9 EFFECTS OF RADIATION

In view of the nuclear applications of many metals and alloys, studies have been conducted on the effects of radiation on material properties. Typical changes in the mechanical properties of steels and other metals exposed to high-energy radiation are increased yield stress and tensile strength and hardness, and decreased ductility and toughness (Fig. 2.24). The magnitudes of these changes depend on the material and its condition, temperature, and level of radiation.

2.2.10 FLEXURE (BEND) TESTS

For materials with insufficient ductility or for brittle materials (certain tool and die materials made of ceramics and carbides), it can be difficult to prepare and test tension-test specimens. The test commonly used for such materials is the 3-point or 4-point flexure, or bend, test, as described in Section 3.7.2.

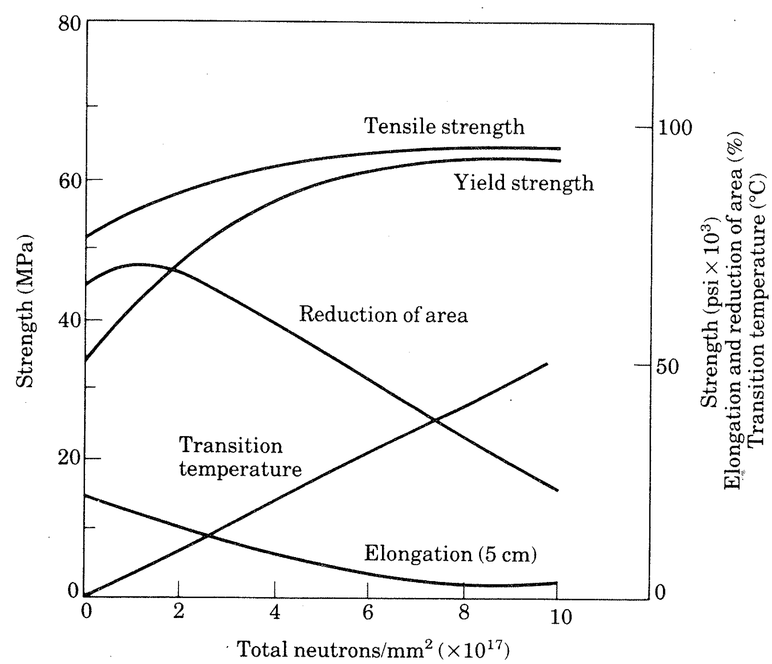


FIGURE 2.24 The effect of radiation on the strength and ductility for a carbon-silicon steel. Such curves are useful in predicting the service behavior of materials for nuclear applications. *Source*: After C. O. Smith, ORSORT, Oak Ridge, TN.

2.3 THE COMPRESSION TEST

Many operations in metalworking, such as forging, rolling, and extrusion, are performed with the workpieces under compressive loads. The compression test, where the specimen is subjected to a compressive load as shown in Fig. 2.1(b), can give useful information for these processes, such as stresses required and the behavior of the material under compression.

However, the deformation shown in Fig. 2.1(b) is ideal. The compression test is usually carried out by compressing a solid cylindrical specimen between two flat platens; the friction between the specimen and the dies is an important factor. Friction causes *barreling* (Fig. 2.25). In other words, friction prevents the top and bottom surfaces from expanding freely.

This situation makes it difficult to obtain relevant data and to construct properly the compressive stress–strain curve for the following reasons:

a. The cross-sectional area of the specimen changes along its height,

b. Friction dissipates energy and this energy is supplied through an increased compressive force (Fig. 2.26). Thus it is difficult to obtain results that are truly indicative of the properties of the material. With effective lubrication or other means (see Chapter 4) it is, of course, possible to minimize friction, and hence barreling, to obtain a reasonably constant cross-sectional area during this test.

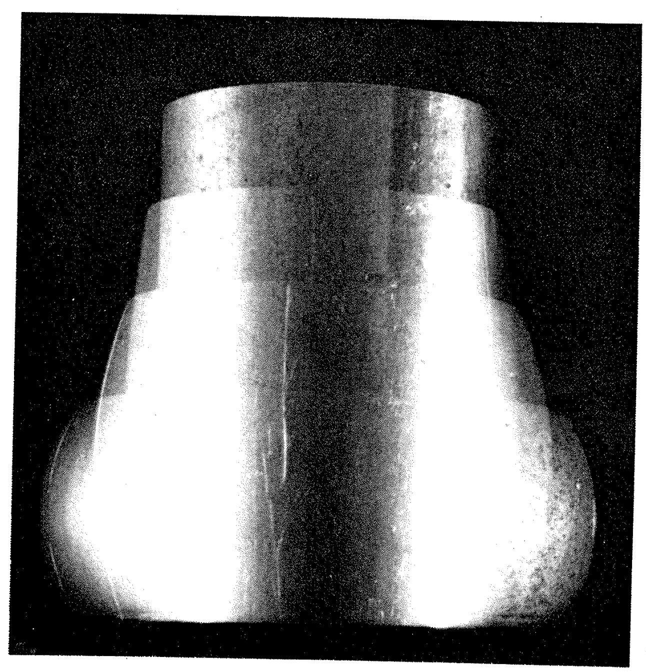


FIGURE 2.25 Barreling in compressing a round solid cylindrical specimen (7075–O aluminum) between flat dies. Barreling is due to friction at the die-specimen interfaces, which retard the free flow of the material. See also Figs. 6.1 and 6.2. *Source*: K. M. Kulkarni and S. Kalpakjian.

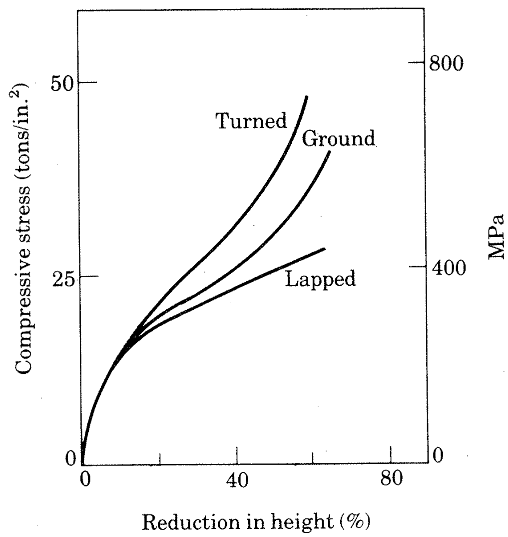


FIGURE 2.26 The effect of surface roughness of dies on the compressive stress for annealed copper. The stress must increase because additional work has to be provided to overcome friction. *Source*: After M. Cook and E. C. Larke, *Jr. Inst. Met.*, vol. 71, 1945, p. 386.

The engineering strain rate $\dot{e}$ in compression is given by

$$\dot{e} = -\frac{v}{h_o}, \tag{2.17}$$

where v is the speed of the die and h_o is the original height of the specimen. The true strain rate $\dot{\epsilon}$ is given by

$$\dot{\epsilon} = -\frac{v}{h}, \tag{2.18}$$

where h is the instantaneous height of the specimen. It can be seen that, if v is constant, the true strain rate increases as the test progresses.

In order to conduct this test at a constant true strain rate, a *cam plastometer* has been designed that, through a cam action, reduces v proportionately as h decreases during the test.

The compression test can also be used to determine the ductility of a metal by observing the cracks that form on the barreled cylindrical surfaces of the specimen. These are discussed in Section 6.7 in greater detail. It should be noted here that hydrostatic pressure has a beneficial effect in delaying the formation of these cracks.

It is apparent that, with a sufficiently ductile material and effective lubrication, compression tests can be carried out uniformly to large strains. This is unlike the tension test, where, even for very ductile materials, necking sets in after relatively little elongation.

Plane Strain Compression Test

Another test is the *plane strain compression test* (Fig. 2.27), which simulates processes such as rolling (see Section 6.10). In this test, the die and workpiece geometries are such that the width of the specimen does not undergo any significant change during

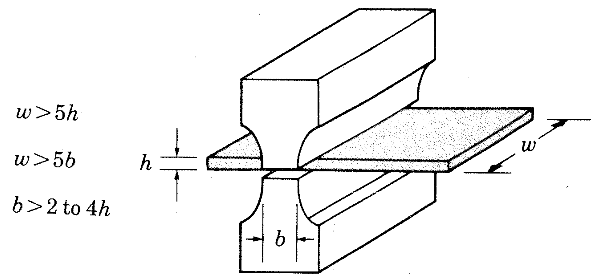


FIGURE 2.27 Schematic illustration of the plane-strain compression test. The dimensional relationships shown should be satisfied for this test to be useful and reproducible. This test gives the yield stress of the material in plane strain, Y'. *Source*: After A. Nadai and H. Ford.

compression; that is, the material under the dies is in the condition of plane strain (see Section 2.8.3).

The yield stress of a material in plane strain Y' can be shown to be

$$Y' = \frac{2}{\sqrt{3}} Y = 1.15Y \tag{2.19}$$

(see Section 2.8.3).

As can be seen from the recommended geometric relationships given in Fig. 2.27, the test parameters must be chosen properly to make the results meaningful. Furthermore, caution should be exercised in test procedures, such as preparing the die surfaces and aligning the dies, lubricating the surfaces, and accurately measuring the load.

When the results of tension and compression tests on the same material are compared, it is found that, for ductile metals, the true stress–true strain curves for both tests coincide (Fig. 2.28). However, this is not true for brittle materials, particularly in regard to ductility.

● Illustrative Problem 2.4

A 40-mm-high cylindrical specimen is being compressed between flat platens at a speed of 0.1 m/s. Calculate the engineering and true strain rates to which the material is being subjected when its height is 10 mm.

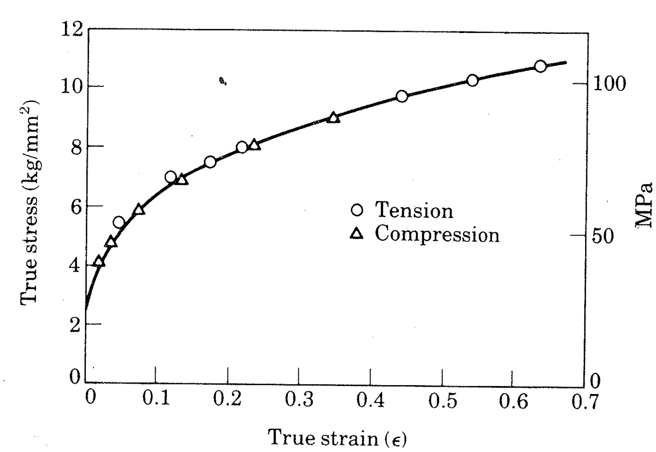


FIGURE 2.28 True stress-true strain curve in tension and compression for aluminum. For ductile metals the curves for tension and compression are identical. This is not true for brittle materials. *Source*: After A. H. Cottrell, *The Mechanical Properties of Matter*, New York: Wiley, 1964, p. 289

SOLUTION. The engineering strain rate is

$$\dot{e} = -\frac{v}{h_o} = -\frac{0.1}{0.040} = -2.5\ \mathrm{s}^{-1}.$$

The true strain rate is

$$\dot{\epsilon} = -\frac{v}{h} = -\frac{0.1}{0.010} = -10\ \mathrm{s}^{-1}.$$

Note the large difference in the numbers obtained. ●

2.3.1 BAUSCHINGER EFFECT

There are situations in deformation processing of materials where a workpiece is first subjected to tension and then to compression, or vice versa. Examples are bending and unbending, roller leveling (see Section 6.12.3), and reverse drawing (see Section 7.9.4). It has been observed that when a metal with a tensile yield stress Y is subjected to tension into the plastic range and then the load is released and applied in compression, the yield stress in compression is lower than that in tension (Fig. 2.29).

This phenomenon is known as the *Bauschinger effect* (after J. Bauschinger, 1886) and is exhibited in varying degrees by all metals and alloys. This effect is also observed when the loading path is reversed, i.e., compression followed by tension. Because of the lowered yield stress in the reverse direction of load application, this phenomenon is also called *strain softening* or *work softening*.

2.3.2 THE DISK TEST

For brittle materials a disk test is available where the disk is subjected to diametral compression. This test is covered in Section 3.7.2.

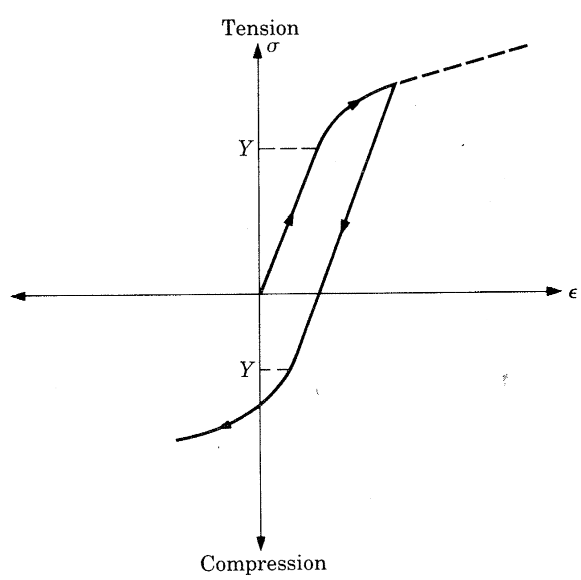


FIGURE 2.29 Schematic illustration of the Bauschinger effect. Arrows show loading and unloading paths. Note the decrease in the yield stress in compression after the specimen has been subjected to tension. The same result is obtained if compression is applied first, followed by tension, whereby the yield stress in tension decreases.

2.4 THE TORSION TEST

Another test method for determination of material properties is the torsion test. In order to obtain an approximately uniform stress and strain distribution along the cross-section, this test is generally carried out on a tubular specimen (Fig. 2.30).

The shear stress τ can be determined from the equation

$$\tau = \frac{T}{2\pi r^2 t}, \tag{2.20}$$

where T is the torque, r is the mean radius, and t is the thickness of the reduced section of the tube.

The shear strain γ is determined from the equation

$$\gamma = \frac{r\phi}{\ell}, \tag{2.21}$$

where ℓ is the length of the reduced section and ϕ is the angle of twist, in radians. With the shear stress and shear strain thus obtained from torsion tests, one can construct the shear stress–shear strain curve of the material (see Section 2.8.7).

In the elastic range, the ratio of the shear stress to shear strain is known as the *shear modulus* or the *modulus of rigidity*, G,

$$G = \frac{\tau}{\gamma}. \tag{2.22}$$

It can be shown that the shear modulus and the modulus of elasticity E are related by the formula

$$G = \frac{E}{2(1 + \nu)}. \tag{2.23}$$

Thus, for most metals E is about 2.6 times G.

Note that unlike tension and compression tests, one does not have to be concerned with changes in the cross-sectional area of the specimen in torsion testing. The

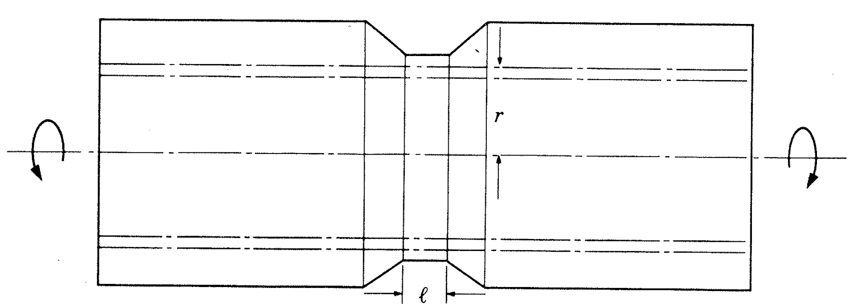


FIGURE 2.30 Notched-tube specimen for torsion testing. Torsion tests do not involve necking (as in tensile testing), or friction (as in compression testing).

shear stress–shear strain curves in torsion increase monotonically, hence they are analogous to true stress–true strain curves.

Torsion tests are performed on solid round bars at elevated temperatures in order to estimate the formability of a metal in forging. The greater the number of twists prior to failure, the greater the forgeability of the material (see Section 6.7).

Tests have also been conducted on round bars that are compressed axially. The maximum shear strain at fracture is measured as a function of the compressive stress. It has been found that the shear strain at fracture increases substantially as the compressive stress increases (Fig. 2.31). This observation again indicates the beneficial effect of a compressive environment on the ductility of materials. Other experiments show that, with tensile normal stresses, the curves in Fig. 2.31 follow a trend downward and to the left, signifying reduced ductility.

The effect of compressive stresses on increasing the maximum shear strain at fracture has also been observed in metal cutting, as treated in Section 8.3.

The normal compressive stress has been found to have no effect on the magnitude of shear stresses required to cause yielding or to continue the deformation, just as hydrostatic pressure has no effect on the general shape of the stress–strain curve (see Fig. 2.23).

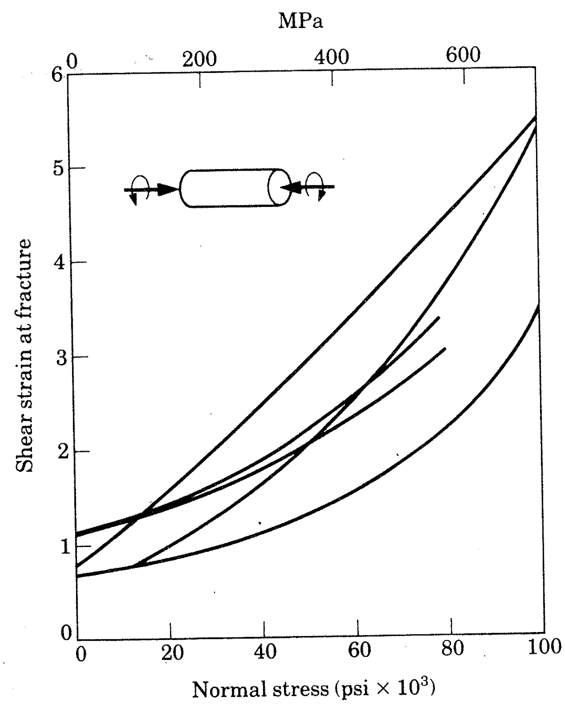


FIGURE 2.31 The effect of axial compressive stress on the shear strain at fracture in torsion for various steels. Note that the effect on ductility is similar to that of hydrostatic pressure, shown in Fig. 2.22. *Source*: Based on data in P. W. Bridgman, *Large Plastic Flow and Fracture*, New York: McGraw-Hill, 1952.

2.5 THE HARDNESS TEST

One of the most common tests for assessment of the mechanical properties of materials is the hardness test. Hardness of a material is generally defined as its resistance to permanent indentation. Less commonly, hardness may also be defined as resistance to scratching or to wear, as treated in Section 4.8.

Various techniques have been developed to measure the hardness of materials using different indenter materials and geometries. However, it has been found that the resistance to indentation depends on the shape of the indenter and the load applied. Thus, hardness is not a fundamental property.

Among the most common standardized hardness tests are the Brinell, Rockwell, Vickers, Knoop, and Scleroscope tests.

2.5.1 BRINELL TEST

In this test, introduced by the Swedish metallurgist J. A. Brinell in 1900, a steel or tungsten carbide ball 10 mm in diameter is pressed against a surface with a load of 500, 1500, or 3000 kg (Fig. 2.32). The Brinell hardness number (HB) is defined as the ratio of the load P to the curved area of indentation,

$$\mathrm{HB} = \frac{2P}{(\pi D)(D - \sqrt{D^2 - d^2})}, \tag{2.24}$$

where D is the diameter of the ball and d is the diameter of the impression in millimeters.

Depending on the condition of the material, different types of impressions are obtained on the surface after performing a Brinell hardness test. Annealed materials generally have a rounded profile, whereas cold-worked (strain-hardened) materials have a sharp profile (Fig. 2.33). The correct method of measuring the indentation diameter d is shown in Figs. 2.33(a) and (b).

Because the indenter (with a finite elastic modulus) also undergoes elastic deformation under the applied load P, hardness measurements may not be as correct as expected. One method of minimizing this effect is to use tungsten carbide balls, which, because of their high modulus of elasticity (see Table 8.7), deform less than steel balls. Also, since harder workpiece materials produce very small impressions, a 1500-kg or 3000-kg load is recommended in order to obtain impressions that are sufficiently large for accurate measurement. Tungsten carbide balls are generally recommended for Brinell hardness numbers higher than 500. In reporting the test results for these high hardnesses the type of ball used should be cited.

Because the impressions made by the same indenter at different loads are not geometrically similar, the Brinell hardness number depends on the load used. Thus the load employed should also be cited with the test results. The Brinell test is generally suitable for materials of low to medium hardness.

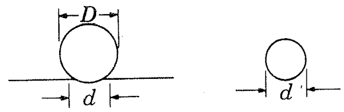
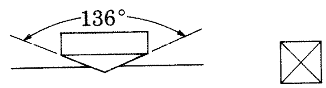
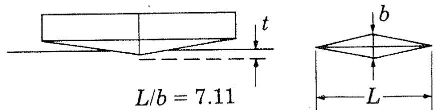
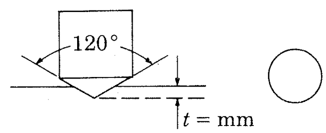
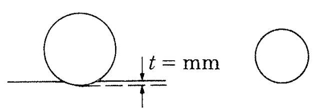
Test	Indenter	Shape of indentation: Side view	Shape of indentation: Top view	Load	Hardness number
Brinell	10-mm steel or tungsten carbide ball			500 kg 1500 kg 3000 kg	
Vickers	Diamond pyramid			1–120 kg	
Knoop	Diamond pyramid	 $L/b = 7.11$ $b/t = 4.00$		25 g–5 kg	
Rockwell				kg	Hardness number
A C D	Diamond cone			60 150 100	HRA HRC HRD $= 100 - 500t$
B F G	$\frac{1}{16}$-in. diameter steel ball			100 60 150	HRB HRF HRG $= 130 - 500t$
E	$\frac{1}{8}$-in. diameter steel ball			100	HRE

FIGURE 2.32 General characteristics of hardness testing methods. The Knoop test is known as a microhardness test because of the light load and small impressions. *Source*: After H. W. Hayden, W. G. Moffatt, and V. Wulff.

Brinelling is a term used to describe permanent indentations on a surface between contacting bodies, for example, a component with a hemispherical protrusion or a ball bearing resting on a flat surface. Under fluctuating loads or vibrations, such as during transportation or vibrating foundations, a permanent indentation may be produced on the flat surface due to dynamic loading.

Meyer Hardness

Whereas the Brinell test is based on the curved area of indentation, the *Meyer test* (after E. Meyer, 1908) uses the projected area of indentation,

$$\text{Meyer hardness} = \frac{4P}{\pi d^2}. \tag{2.25}$$

This equation may be rewritten in the form

$$P = C_1 d^n, \tag{2.26}$$

where C_1 is a constant for the material and n is known as the Meyer hardness exponent, which is a measure of the strain hardening capability of the material. However, the Meyer test is not commonly used.

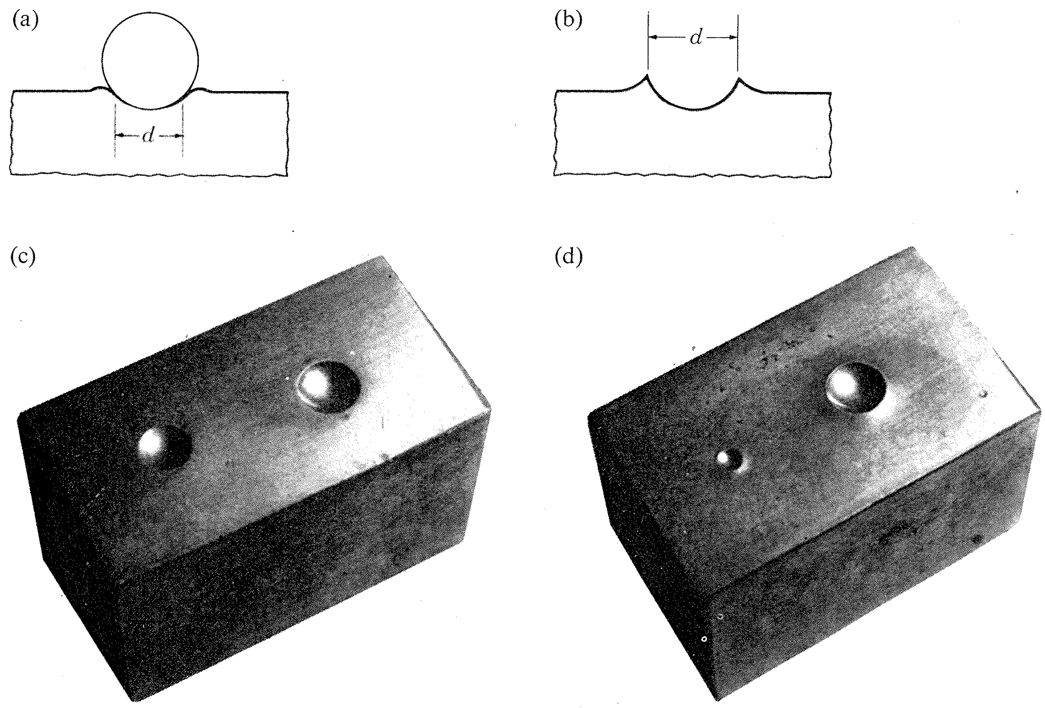


FIGURE 2.33 Indentation geometry for Brinell hardness testing. (a) Annealed metal. (b) Work-hardened metal. (c) Hardness impression on annealed steel. (d) Hardness impression on cold-worked steel. Note the difference of metal flow at the periphery of the impressions.

2.5.2 ROCKWELL TEST

In this test, developed by the metallurgist S. P. Rockwell in 1922, the depth of penetration is measured. The indenter is pressed on the surface, first with a minor load and then a major one. The difference in the depth of penetration is a measure of the hardness.

There are several Rockwell hardness test scales that use different loads and indenter materials and geometries. Some of the more common hardness scales and the indenters used are listed in Fig. 2.32. The Rockwell hardness number, which is read directly from a dial on the testing machine, is expressed as follows: If the hardness number is 55 using the C scale, then it is written as 55 HRC. *Rockwell superficial hardness* tests have also been developed using lighter loads and the same type of indenters.

The Rockwell hardness test is used for a wide range of hardnesses. The test is rapid and can be automated.

2.5.3 VICKERS TEST

The *Vickers* hardness test, also known as the *diamond pyramid hardness* test, uses a pyramid-shaped diamond indenter (see Fig. 2.32) with a load ranging from 1 to

120 kg. The test was developed in England in 1922 by R. Smith and G. Sandlund. The Vickers hardness number (HV) is given by the formula

$$\mathrm{HV} = \frac{1.854P}{L^2}. \tag{2.27}$$

The impressions are typically less than 0.5 mm on the diagonal. The Vickers test gives essentially the same hardness number regardless of the load. It is suitable for testing materials with a wide range of hardness, including very hard steels.

2.5.4 KNOOP TEST

This test uses a diamond indenter in the shape of an elongated pyramid (see Fig. 2.32) with loads ranging generally from 25 g to 5 kg. The *Knoop* hardness number (HK) is given by the formula

$$\mathrm{HK} = \frac{14.2P}{L^2}. \tag{2.28}$$

The Knoop test (after F. Knoop, 1939) is a *microhardness* test because of the light loads it employs and hence is suitable for very small or thin specimens and for brittle materials, such as gem stones, carbides, and glass. This test is also used in measuring the hardness of individual grains in a metal. The size of the indentation is generally in the range of 0.01 to 0.10 mm; thus surface preparation is very important. Because the hardness number obtained depends on the applied load, test results should always cite the load employed.

2.5.5 SCLEROSCOPE

The Scleroscope is an instrument in which a diamond-tipped indenter (hammer), enclosed in a glass tube, is dropped on the specimen from a certain height. The hardness is determined by the rebound of the indenter; the higher the rebound, the harder the specimen. Indentation is slight. Since the instrument is portable it is useful for measuring the hardness of large objects.

2.5.6 MOHS HARDNESS

This test is based on the capability of one material to scratch another. The Mohs hardness (after F. Mohs, 1822) is based on a scale of 10, with 1 for talc and 10 for diamond (hardest substance known). Thus, a material with a higher Mohs hardness can scratch ones with a lower hardness (Fig. 2.34).

The Mohs scale is used generally by mineralogists and geologists. However, some of the materials tested are of interest to manufacturing engineers, as discussed in Chapters 8, 9, and 11. Although the Mohs scale is qualitative, good correlation is obtained with the Knoop hardness. Soft metals have a Mohs scale of 2 to 3, hardened steels about 6, and aluminum oxide (used in grinding wheels) a scale of 9.

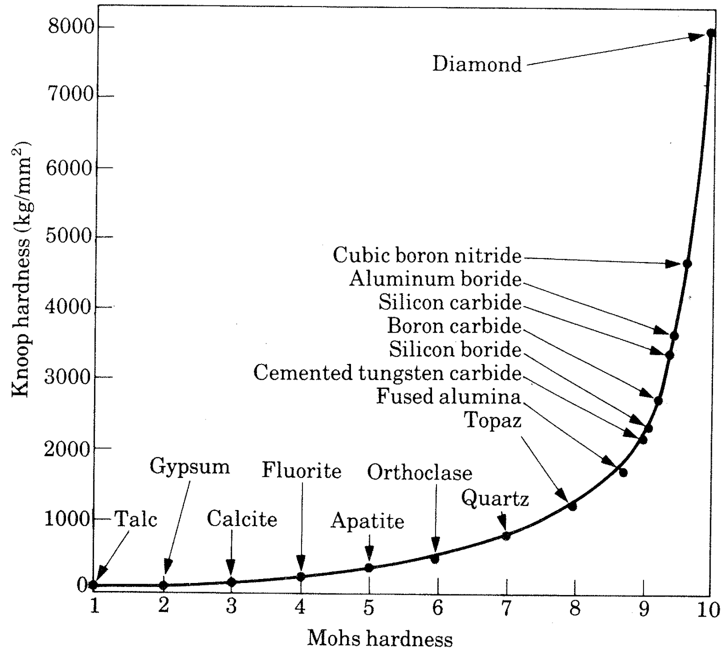


FIGURE 2.34 Comparison of Knoop and Mohs hardness scales for various materials. Diamond is the hardest substance known. Diamonds that are used in manufacturing industries are synthetically made.

2.5.7 HARDNESS AND STRENGTH

Since hardness is the resistance to permanent indentation, it is equivalent to performing a compression test on a small portion of the surface of a material. Thus, one would expect some correlation between hardness and yield stress Y in the form of

$$\text{Hardness} = cY, \tag{2.29}$$

where c is a proportionality constant.

Theoretical studies, based on plane strain slip-line analysis (Section 2.10.2) with a smooth flat punch indenting the surface of a semi-infinite body, have shown that for a perfectly plastic material of yield stress Y the value of c is about 3. This is in reasonably good agreement with experimental data (Fig. 2.35). Note that cold-worked materials (which are close to being perfectly plastic in their behavior) show better agreement than annealed ones. The higher c value for the annealed materials is explained by the fact that, due to strain hardening, the average yield stress they exhibit during indentation is higher than their initial yield stress.

The reason for hardness, as a compression test, giving higher values than the uniaxial yield stress Y of the material can be seen in the following analysis. If we assume that the volume under the indenter is a column of material (Fig. 2.36), then it would exhibit an uniaxial compressive yield stress Y. However, the volume being deformed under the indenter is, in reality, surrounded by a rigid mass (Fig. 2.37). The surrounding mass prevents this volume of material from deforming freely. In fact, this volume is under *triaxial compression.* As will be seen in Section 2.8 on yield criteria, this material requires a normal compressive yield stress that is higher than the uniaxial yield stress of the material.

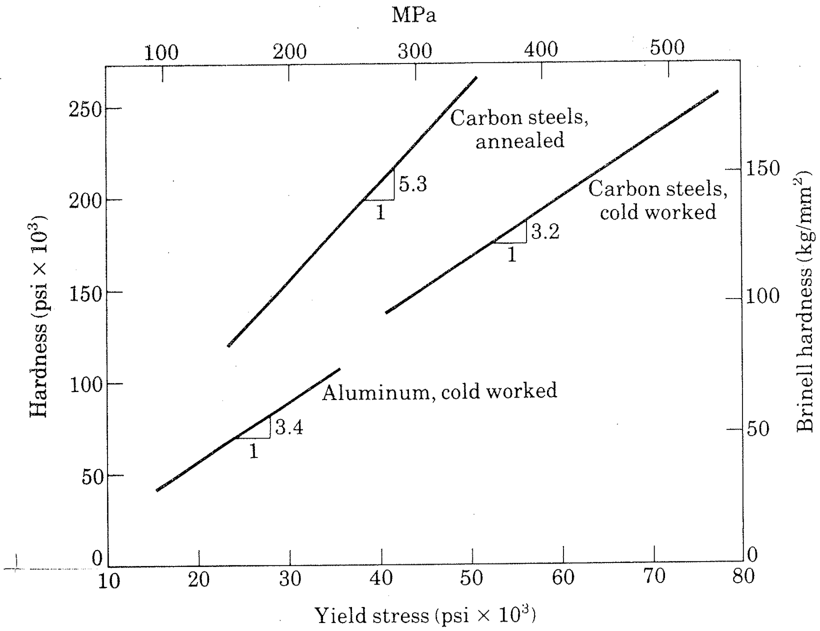


FIGURE 2.35 Relation between Brinell hardness and yield stress for aluminum and steels. For comparison, the Brinell hardness (which is always measured in kg/mm^2) is converted to psi units on the left scale.

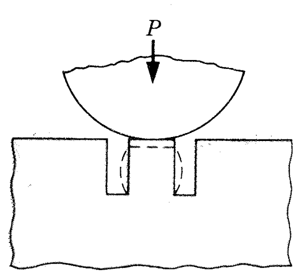


FIGURE 2.36 Simulation of a hardness test on a uniaxial column of metal, showing unconstrained deformation of the metal under the indenter. Because in actual hardness testing the metal is constrained, hardness values are about three times that of the uniaxial yield stress of the metal. *Source*: After E. Orowan.

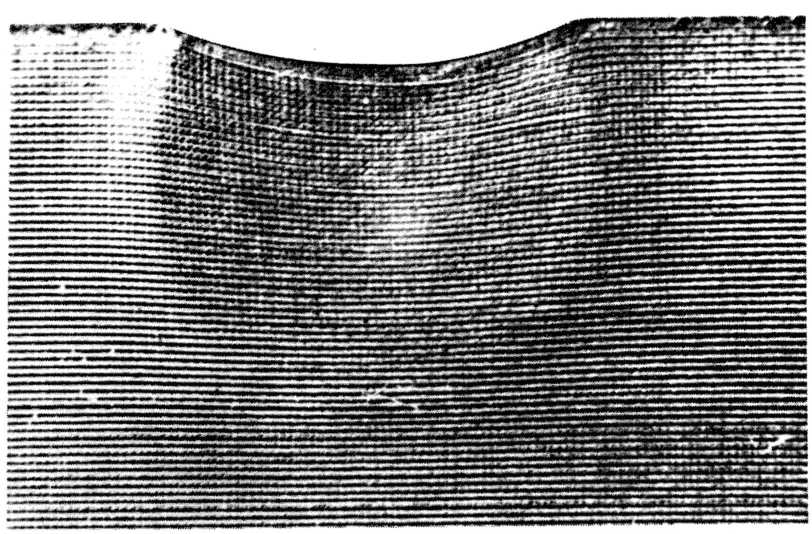


FIGURE 2.37 Bulk deformation in mild steel under a spherical indenter. Note that the depth of the deformed zone is about one order of magnitude larger than the depth of indentation. For a hardness test to be valid, the material should be allowed to fully develop this zone. This is why thinner specimens require smaller impressions. *Source*: Courtesy of M. C. Shaw and C. T. Yang.

More practically, a relationship has also been observed between the ultimate tensile strength (UTS) and Brinell hardness number (HB) for steels, as follows:

$$\text{UTS} = 500(\text{HB}), \tag{2.30}$$

where *UTS* is in psi and HB in kg/mm^2 as measured with a load of 3000 kg. In *SI* units, the relationship is given by

$$\text{UTS} = 3.5(\text{HB}), \tag{2.31}$$

where UTS is in MPa.

2.5.8 HARDNESS TESTING PROCEDURES

For a hardness test to be meaningful and reliable, the *zone of deformation* under the indenter (see Fig. 2.37) 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. It is generally recommended that the location be at least two diameters (of the ball) from the edge of the specimen, and that the thickness should be at least 10 times the depth of penetration of the indenter. These can be appreciated by reviewing Fig. 2.37.

Moreover, the indentation should be sufficiently large to give a representative hardness value for the bulk material. If hardness variations are 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, like those in Vickers or Knoop tests obtained under light loads.

While surface preparation is not critical for the Brinell test and somewhat more important for the Rockwell test, it is important for the other hardness tests because of the small size of the indentations. Surfaces may have to be polished in order to measure correctly the dimensions of the impression.

The values obtained from different hardness tests can be correlated and converted to different scales (Fig. 2.38). Charts are available in the technical literature for such conversions. Care should be exercised in using these charts because of the many variables involved with regard to material characteristics and indentation geometry.

Although most hardness testers require relatively small specimens or parts, portable testers have been developed (in addition to the Scleroscope) for soft metals and nonmetallic materials. These testers, which can be used on large parts, usually operate on the principle of a spring-loaded indenter mounted in a suitable frame. New hardness testing methods are also being developed. One is based on the eddy current principle and utilizes the electrical and magnetic properties of the workpiece material. In another method, the deceleration of an impacting indenter is measured.

Hot hardness tests can also be carried out using conventional testers with certain modifications, such as surrounding the specimen and indenter with a small electric furnace. The hot hardness of materials is important in applications where the materials are subjected to elevated temperatures, such as in cutting tools in machining and dies for metalworking (see Fig. 8.52).

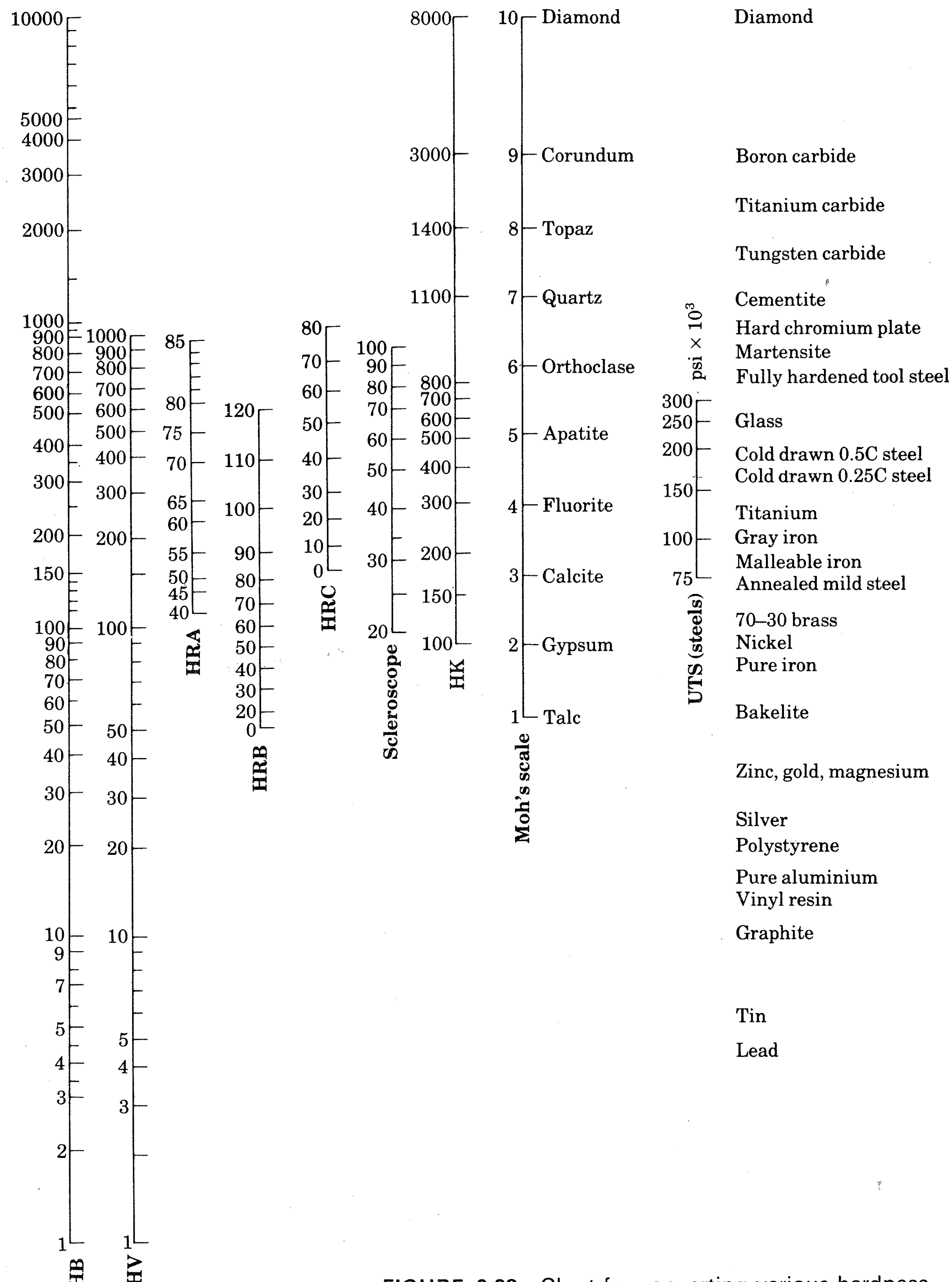


FIGURE 2.38 Chart for converting various hardness scales. Because of the various factors involved, hardness conversions are approximate.

R.J. Reynolds Vapor Exhibit 1034-00165

● **Illustrative Problem 2.5**

A piece of steel is highly deformed at room temperature. Its hardness is found to be 300 HB. Estimate the modulus of resilience for this material in lb/in^2.

SOLUTION. Since the steel has been subjected to large strains at room temperature, we may assume that its stress–strain curve has flattened considerably, thus approaching the shape of a perfectly plastic curve. According to Eq. (2.29) and using a value of $c = 3$, we obtain

$$Y = \frac{300}{3} = 100 \text{ kg/mm}^2 = 142{,}250 \text{ psi}.$$

The modulus of resilience is defined as in Eq. (2.6),

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

For steel, $E = 30 \times 10^6$ psi. Hence,

$$\text{Modulus of resilience} = \frac{(142{,}250)^2}{2 \times 30 \times 10^6} = 337 \text{ in.} \cdot \text{lb/in}^3. \; ●$$

2.6 DEFORMATION-ZONE GEOMETRY

The observations made concerning Fig. 2.37 are important in estimating and calculating forces in metalworking operations. As shown above for hardness testing, the compressive stress required for indentation is, ideally, about 3 times the yield stress Y required for uniaxial compression. It was also noted that:

a. The deformation under the indenter is localized, making the overall deformation highly nonuniform,

b. The deformation zone is relatively small compared to the dimensions of the specimen.

On the other hand, in a simple frictionless compression test with flat dies, the top and bottom surfaces of the specimen are always in contact with the dies and the specimen deforms uniformly.

You can visualize different situations covering the wide range between these two extreme examples of specimen-deformation geometry. The *deformation zones* and the pressures required are shown in Fig. 2.39 for the frictionless condition as obtained from slip-line analysis (Section 2.10.2). Note that the ratio h/L is the important parameter in determining the inhomogeneity of deformation. It is also important to emphasize the frictionless nature of these examples since, as will be shown in Section 2.10.1, friction has a significant effect on forces, particularly at small values of h/L.

Deformation-zone geometry depends on the particular metalworking process and such parameters as the die geometry and percent reduction of the material (Fig. 2.40). Details of these processes and the role of the deformation geometry are covered in Chapter 6.

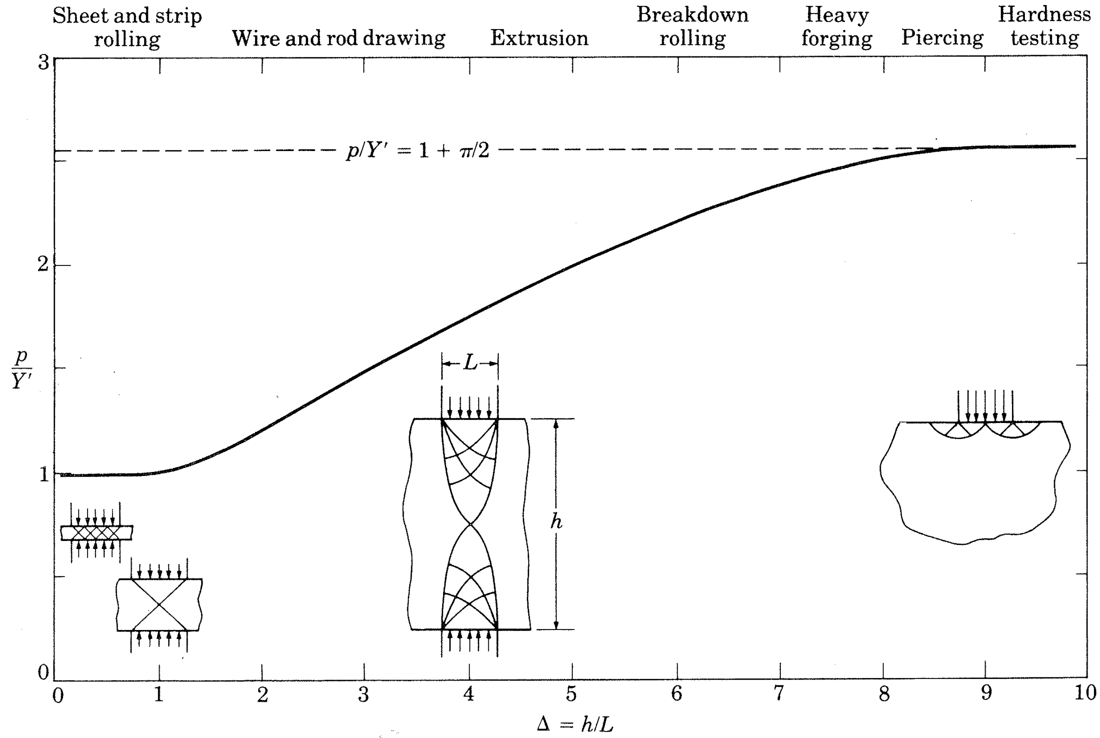


FIGURE 2.39 Die pressures required in frictionless plane-strain conditions for a variety of metalworking operations. The geometric relationship between contact area of the dies and workpiece dimensions is an important factor in predicting forces in plastic deformation of materials. *Source*: After W. A. Backofen, *Deformation Processing*, Reading, Mass.: Addison-Wesley, 1972, p. 135.

2.7 RESIDUAL STRESSES

In this section it will be shown that inhomogeneous deformation, such as that discussed above, leads to *residual stresses*—stresses that remain within a part after it has been deformed and all external forces have been removed.

A typical example of inhomogeneous deformation is the bending of a beam (Fig. 2.41). The bending moment first produces a linear elastic stress distribution. As the moment is increased, the outer fibers begin to yield and, for a typical strain-hardening material, the stress distribution shown in Fig. 2.41(b) is eventually obtained. After the part is bent (permanently, since it has undergone plastic deformation) the moment is removed by unloading. This unloading is equivalent to applying an equal and opposite moment to the beam.

As shown in Fig. 2.8, all recovery is elastic. This means that in Fig. 2.41(c) the moments of the areas *oab* and *oac*, about the neutral axis, must be equal. (For the purposes of this treatment it is assumed that the neutral axis does not shift.)

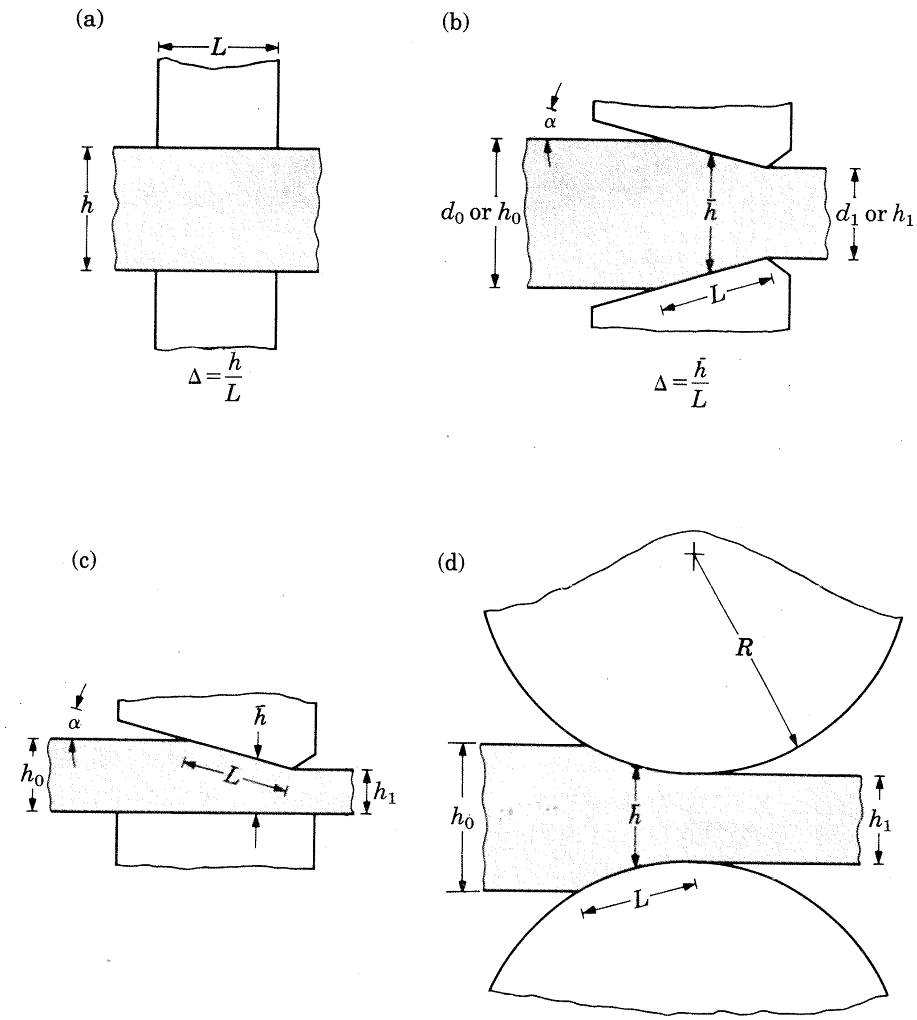


FIGURE 2.40 Examples of plastic deformation processes in plane strain showing the h/L ratio. (a) Indenting with flat dies. This is similar to the cogging operation shown in Fig. 6.18. (b) Drawing or extrusion of strip with a wedge-shaped die, described in Chapter 6. (c) Ironing. See also Fig. 7.46. (d) Rolling, described in Section 6.11. As shown in Fig. 2.39, the larger the h/L ratio, the higher the die pressure. In actual processing, however, the smaller this ratio the greater is the effect of friction at the die-workpiece interfaces. This is because contact area, hence friction, increases with and compression are equal, as they are for ?

The difference between the two stress distributions produces the residual stress pattern within the beam. Note that there are compressive residual stresses in layers *ad* and *oe*, and tensile in layers *do* and *ef*. Since there are no external forces, the residual stress system in the beam must be in static equilibrium.

Although this example involves stresses in one direction only, in most situations in deformation processing the residual stresses are three-dimensional.

In the example described above, the equilibrium of residual stresses may be disturbed by altering the geometry of the beam, such as by removing a layer of material.

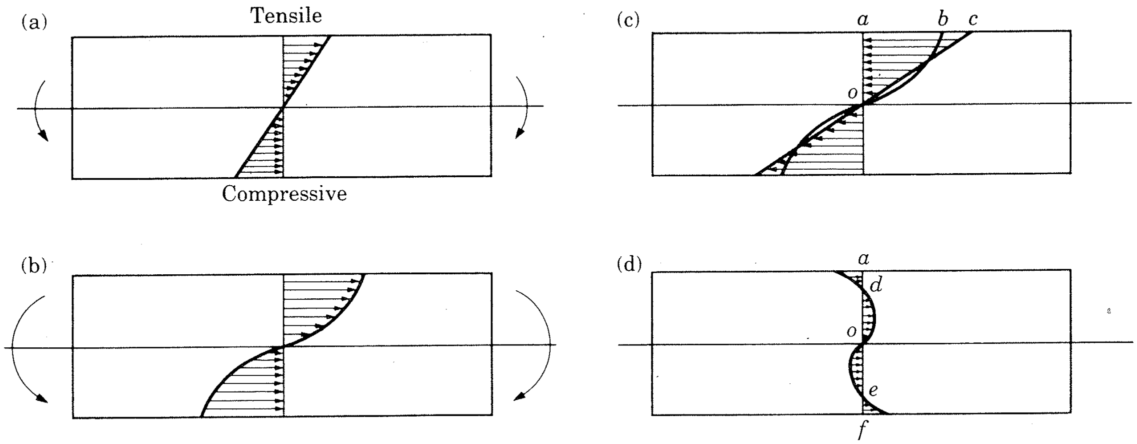


FIGURE 2.41 Residual stresses developed in bending a beam made of an elastic, strain-hardening material. Note that unloading is equivalent to applying an equal and opposite moment to the part, as shown in (b). Because of nonuniform deformation, most parts made by plastic deformation processes contain residual stresses. Note that the forces and moments due to residual stresses must be internally balanced.

The beam will then acquire a new radius of curvature in order to ensure the balance of internal forces. Another example of this effect is the drilling of round holes on surfaces with residual stresses. It may be found that, as a result of removing this material, the equilibrium of the residual stresses is disturbed and the hole becomes oval.

Such disturbances of residual stresses lead to *warping*, some simple examples of which are shown in Fig. 2.42. The equilibrium of residual stresses may also be disturbed by *relaxation* of these stresses over a period of time; this results in instability of the dimensions and shape of the component. These dimensional changes can be an important consideration for precision machinery and measuring equipment.

Residual stresses are also caused by *phase changes* in metals during or after processing due to density differences between phases (e.g., between ferrite and martensite in steels). This causes microscopic volumetric changes and results in residual stresses. This phenomenon is important in warm and hot working and in heat treatment following cold working (e.g., deformation at room temperature).

Residual stresses can also be caused by *temperature gradients* within a body, such

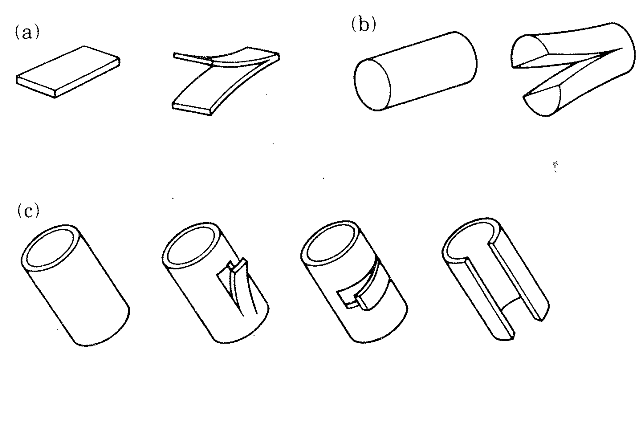


FIGURE 2.42 Distortion of parts with residual stresses after cutting or slitting. (a) Rolled sheet or plate. (b) Drawn rod. (c) Thin-walled tubing. Because of the presence of residual stresses on the surfaces of parts, a round drill may produce an oval-shaped hole because of relaxation of stresses when a portion is removed.

as during the cooling cycle of a casting, applying brakes to a railroad wheel, or in a grinding operation. The expansion and contraction due to temperature gradients are analogous to nonuniform plastic deformation.

2.7.1 EFFECTS OF RESIDUAL STRESSES

Tensile residual stresses on the surface of a part are generally considered to be undesirable because they lower the fatigue life and fracture strength of the part. A surface with tensile residual stresses can sustain lower additional tensile stresses (due to external loading) than a surface that is free of any residual stresses. This is particularly true for relatively brittle materials where fracture takes place with little or no plastic deformation.

Tensile residual stresses in manufactured products can also lead to *stress cracking* or *stress–corrosion cracking* over a period to time (see Section 3.7.1).

On the other hand, compressive residual stresses on a surface are generally desirable. In fact, in order to increase the fatigue life of components, compressive residual stresses are imparted on surfaces by techniques such as *shot peening* and *surface rolling* (see Section 4.13.2).

Residual stresses on a surface can also have an effect on the action of lubricants, as described in Section 4.11.

2.7.2 REDUCTION OF RESIDUAL STRESSES

Residual stresses may be reduced or eliminated either by *stress-relief annealing* (Section 3.5) or by further *plastic deformation.* Given sufficient time, residual stresses may also be diminished at room temperature by *relaxation.* The time required can be greatly reduced by raising the temperature of the component.

Relaxation of residual stresses by stress-relief annealing is generally accompanied by warpage of the part. Hence a "machining allowance" is commonly provided to compensate for dimensional changes during stress relieving.

The mechanism of reduction or elimination of residual stresses by plastic deformation is as follows: Assume that a piece of metal has the residual stresses shown in Fig. 2.43, namely, tensile on the outside and compressive on the inside. The part with these stresses, which are in the elastic range, is in equilibrium. Also assume that the material is elastic–perfectly plastic, as shown by the diagram in the figure.

The level of residual stresses are shown on the stress–strain diagram, both being below the yield stress Y. If a uniformly distributed tension is applied to this specimen, points σ_c and σ_t in the diagram move up on the stress–strain curve, as shown by the arrows. The maximum level that these stresses can reach is the tensile yield stress Y. With sufficiently high loading, the stress distribution becomes uniform throughout the part, as shown in Fig. 2.43(c). If the load is now removed, the stresses recover elastically and the part has no residual stresses.

Note that very little stretching is required to relieve these residual stresses. This is because the elastic portions of the stress–strain curves for metals are very steep, hence the elastic stresses can be raised to the yield stress with very little strain.

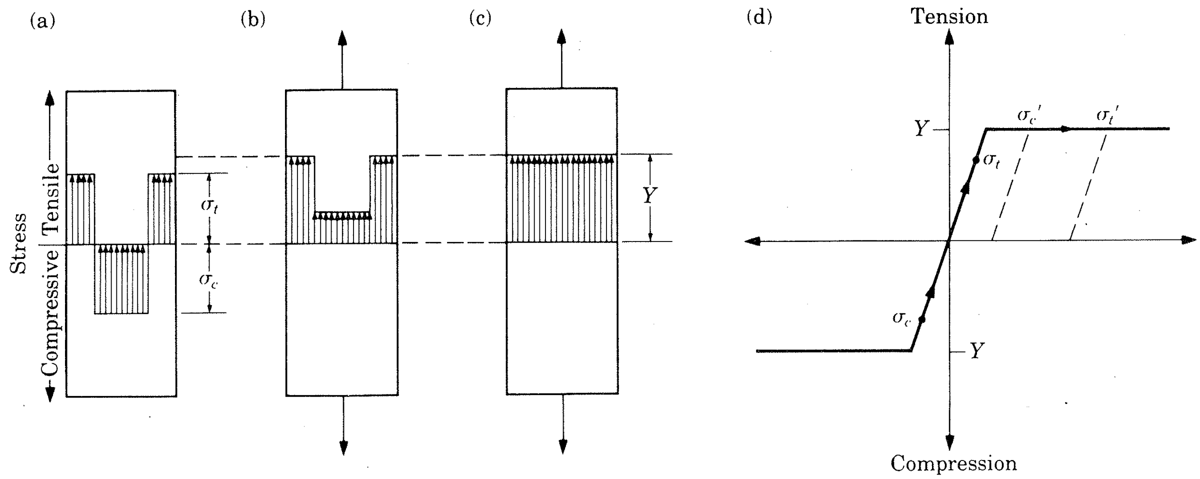


FIGURE 2.43 Elimination of residual stresses by stretching. See Section 2.7.2 for details. Residual stresses can also be reduced or eliminated by thermal treatments, such as stress relieving or annealing.

The technique for reducing or relieving residual stresses by plastic deformation, such as by stretching as described above, requires sufficient straining to establish a uniformly distributed stress in the part. It is therefore apparent that a material such as the elastic, linearly strain-hardening type, Fig. 2.8(c), can never reach this condition since the compressive stress $\sigma_{c'}$ will always lag behind $\sigma_{t'}$. If the slope of the stress–strain curve in the plastic region is small, then the difference between $\sigma_{c'}$ and $\sigma_{t'}$ will be rather small and little residual stresses will be left in the part after unloading.

● Illustrative Problem 2.6

In Fig. 2.43 assume that $\sigma_t = 140$ MPa and $\sigma_c = -140$ MPa. The material is aluminum and the length of the specimen is 0.25 m. Calculate the length to which this specimen should be stretched so that, when unloaded, it will be free of any residual stresses. Assume that the yield stress of the material is 150 MPa.

SOLUTION. The stretching should be to the extent whereby σ_c reaches the yield stress in tension, Y. Thus the total strain should be

$$\epsilon_{\text{total}} = \frac{\sigma_c}{E} + \frac{Y}{E}.$$

For aluminum, let $E = 70$ GPa. Thus,

$$\epsilon_{\text{total}} = \frac{140}{70 \times 10^3} + \frac{150}{70 \times 10^3} = 0.00414.$$

Hence the stretched length should be

$$\ln\left(\frac{\ell_f}{0.25}\right) = 0.00414$$

or

$$\ell_f = 0.2510 \text{ m.}$$

Since the strains are very small, we may use engineering strains in these calculations. Thus,

$$\frac{\ell_f - 0.25}{0.25} = 0.00414$$

or

$$\ell_f = 0.2510 \text{ m.} \quad \bullet$$

2.7.3. DETERMINATION OF RESIDUAL STRESSES

Various experimental techniques have been developed for the determination (also called measurement) of residual stresses in manufactured parts. The most common nondestructive technique is by x-ray analysis, where the interatomic spacing of a set of lattice planes is measured and is compared with that of a stress-free specimen.

The most common destructive technique involves removal of layers of material from a part and then measuring the dimensional changes, such as changes in the length or curvature of the part. These dimensional changes are very small and are generally measured with the use of strain gages. For large deflections, other measuring instruments can be used. The residual stresses are then calculated from a set of equations relating stresses to strains. The work involved is somewhat tedious and time consuming, especially in the three-dimensional case.

Another technique for determining residual stresses is to place strain gages on a surface and then drill a small hole near it, usually 0.020 in. (0.5 mm) in diameter. The strain gages measure the strains around the drilled hole. Based on these strains, the residual stresses can then be computed. This is a semidestructive method. In addition to the determinâtion of surface residual stresses by this method, a variation of this technique (called incremental drilling method) can be used to compute residual stresses in the underlying layers of a material.

2.8 YIELD CRITERIA

In most manfacturing operations involving deformation processing, the material, unlike that in a simple tension or compression test specimen, is generally subjected to *triaxial* stresses. For example, in the expansion of a thin-walled spherical shell under internal pressure, an element in the shell is subjected to equal biaxial tensile stresses (Fig. 2.44a). In drawing a rod or wire through a conical die, an element in the deformation zone is subjected to tension in its length direction and to compression on its conical surface (Fig. 2.44b). An element in the flange in deep drawing of sheet metal is subjected to a tensile radial stress and compressive stresses on its surface and in the

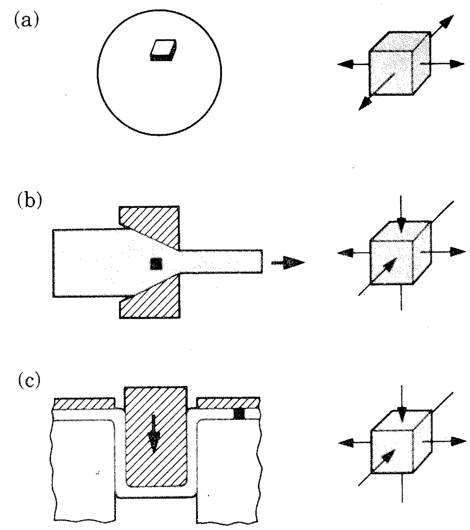


FIGURE 2.44 The state of stress in various metalworking operations. (a) Expansion of a thin-walled spherical shell under internal pressure. (b) Drawing of round rod or wire through a conical die to reduce its diameter. (c) Deep drawing of sheet metal with a punch and die to make a cup.

circumferential direction (Fig. 2.44c). As will be shown in subsequent chapters, many similar examples can be given where the material is subjected to various normal and shear stresses during processing.

In a simple tension or compression test, when the applied stress reaches the uniaxial yield stress Y, the material will deform plastically. However, if the material is subjected to a more complex state of stress there are relationships between these stresses that will predict yielding. These relationships are known as *yield criteria*, the most common ones being the maximum-shear-stress criterion and the distortion-energy criterion.

2.8.1 MAXIMUM-SHEAR-STRESS CRITERION

This criterion (also known as the *Tresca criterion*, after H. Tresca, 1864) states that yielding occurs when the maximum shear stress within an element is equal to or exceeds a critical value. As will be seen in Section 3.2.2, this critical value of the shear stress is a material property called *shear yield stress* (k). Thus, for yielding,

$$\tau_{\max} \geq k. \tag{2.32}$$

The most convenient way of determining the stresses on an element is by the use of *Mohr circles* (after O. Mohr, 1914). Some typical examples are shown in Fig. 2.45. Note that in these examples the normal stresses are *principal stresses*; in other words, they act on planes on which there are no shear stresses. (The construction of Mohr circles and the determination of principal stresses will not be discussed here as they are covered in textbooks on strength of materials.)

If the maximum shear stress, as determined from Fig. 2.45 or from appropriate equations, is equal to or exceeds k, then yielding will occur. This can be best visualized by the construction in Fig. 2.46. In order to cause yielding, the largest circle must touch the "ceiling" represented by the shear yield stress k. Note that there are many combinations of stresses (or states of stress) that can give the same maximum shear stress.

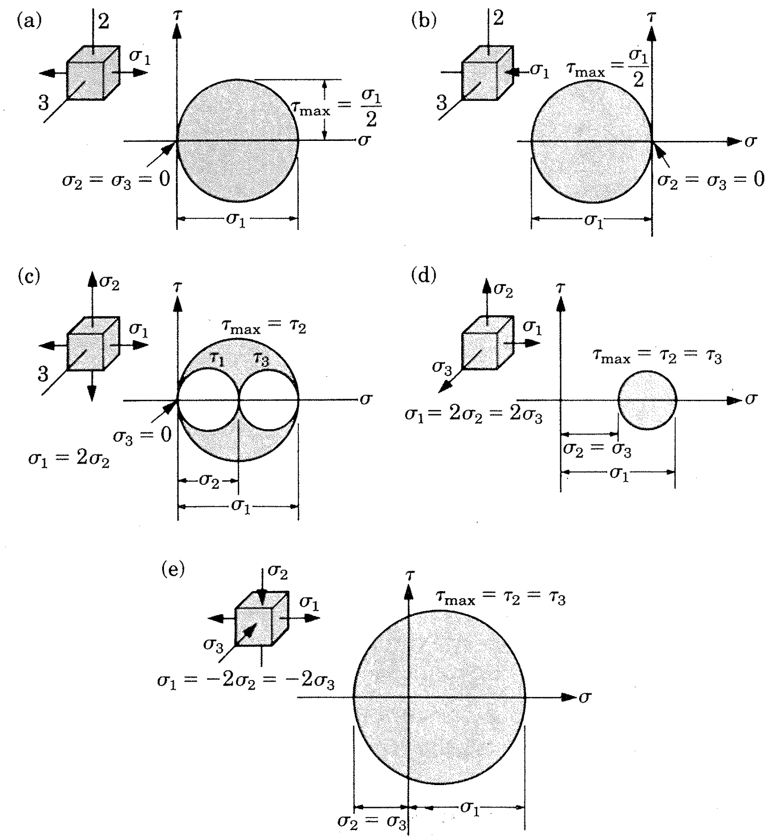


FIGURE 2.45 Mohr's circles for various states of stress. (a) Uniaxial tension. (b) Uniaxial compression. (c) Biaxial tension (plane stress). (d) Triaxial tension. (e) Biaxial compression with tension. These states of stress are encountered in most metalworking processes.

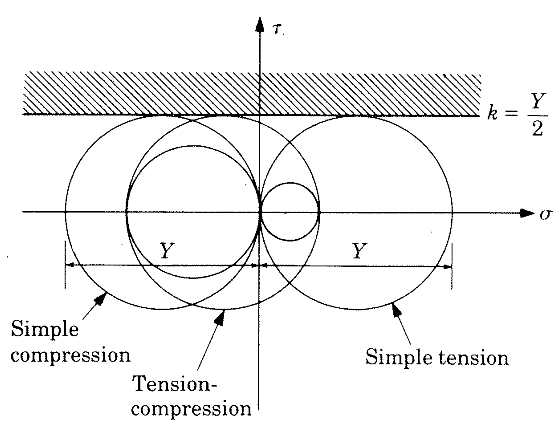


FIGURE 2.46 Three sets of Mohr's circles representing stresses that are large enough to cause yielding. Note that there is an infinite number of states of stress that produce circles of the same maximum diameter, hence yielding. For any state of stress, the radius of the largest circle must be equal to the shear yield stress k to cause yielding.

If, for some reason, we are unable to increase the stresses on the element in order to cause yielding, then the solution is simply to lower the ceiling by raising the temperature of the material. This is the basis and one major reason for hot working of materials (see Section 3.6).

From the simple tension test, we find that

$$k = \frac{Y}{2}. \tag{2.33}$$

It should be pointed out that the material is assumed to be *continuous*, *homogeneous*, and *isotropic*, i.e., it has the same properties in all directions. Also, tensile stresses are positive and compressive stresses negative, and the yield stress in tension and in compression are assumed to be equal. These are important assumptions.

The maximum-shear-stress criterion can now be written as

$$\sigma_{\max} - \sigma_{\min} = Y = 2k. \tag{2.34}$$

This means that the maximum and minimum normal stresses produce the largest circle and hence the largest shear stress. Consequently, the intermediate stress has no effect on yielding. It should be emphasized that the left-hand side of Eq. (2.34) represents the applied stresses and the right-hand side is a material property.

2.8.2 DISTORTION-ENERGY CRITERION

This criterion (also known as the *von Mises criterion*, after R. von Mises, 1913) states that yielding will occur when the relationship between the principal applied stresses and the uniaxial yield stress Y of the material is as follows:

$$(\sigma_1 - \sigma_2)^2 + (\sigma_2 - \sigma_3)^2 + (\sigma_3 - \sigma_1)^2 = 2Y^2. \tag{2.35}$$

Note that, unlike the maximum-shear-stress criterion, the intermediate principal stress is included in this equation. Here again, the left-hand side of the equation represents the applied stresses and the right-hand side a material property.

● **Illustrative Problem 2.7**

A thin-walled spherical shell is under internal pressure p. The shell is 20 in. in diameter and 0.1 in. thick. It is made of a perfectly plastic material with a yield stress of 20,000 psi. Calculate the pressure required to cause yielding of the shell according to both yield criteria discussed above.

SOLUTION. For this shell under internal pressure, the membrane stresses are given by

$$\sigma_1 = \sigma_2 = \frac{pr}{2t},$$

where $r = 10$ in. and $t = 0.1$ in. The stress in the thickness direction, σ_3, is negligible because of the high r/t ratio of the shell. Thus, according to the maximum-shear-stress criterion,

$$\sigma_{max} - \sigma_{min} = Y$$

or

$$\sigma_1 - 0 = Y$$

and

$$\sigma_2 - 0 = Y.$$

Hence, $\sigma_1 = \sigma_2 = 20{,}000$ psi.

The pressure required is then

$$p = \frac{2tY}{r} = \frac{(2)(0.1)(20{,}000)}{10} = 400 \text{ psi.}$$

According to the distortion-energy criterion,

$$(\sigma_1 - \sigma_2)^2 + (\sigma_2 - \sigma_3)^3 + (\sigma_3 - \sigma_1)^2 = 2Y^2$$

or

$$0 \quad + \quad \sigma_2^2 \quad + \quad \sigma_1^2 \quad = 2Y^2.$$

Hence $\sigma_1 = \sigma_2 = Y$.

Thus the answer is the same, or

$p = 400$ psi. ●

● **Illustrative Problem 2.8**

Explain why a correction factor has to be applied in the construction of a true stress–true strain curve, shown in Fig. P2.2, from tensile test data.

SOLUTION. The stress distribution at the neck of a specimen is shown in Fig. P2.3, where we note that in this region there is a triaxial state of stress. This is so because each element in the region has a different cross-sectional area; the smaller the area, the greater the tensile stress on the element. Hence, element 1 will contract more than element 2, and so on. However, element 1 is restrained from contracting freely by element 2, and element 2 is restrained by element 3, and so on. This restraint causes radial and circumferential tensile stresses in the necked region. This situation results in an axial tensile stress distribution as shown in Fig. P2.3.

The true uniaxial stress in tension is σ, whereas the calculated value of true stress at fracture is the average stress. Hence, a correction has to be made. A mathematical analysis by P. W. Bridgman gives the ratio of true to average stress as

$$\frac{\sigma}{\sigma_{av}} = \frac{1}{(1 + 2R/a)[\ln(1 + a/2R)]}, \tag{2.36}$$

where R is the radius of curvature of the neck and a is the radius of the specimen at the neck. Since R is not easy to measure during a test, an empirical relation has

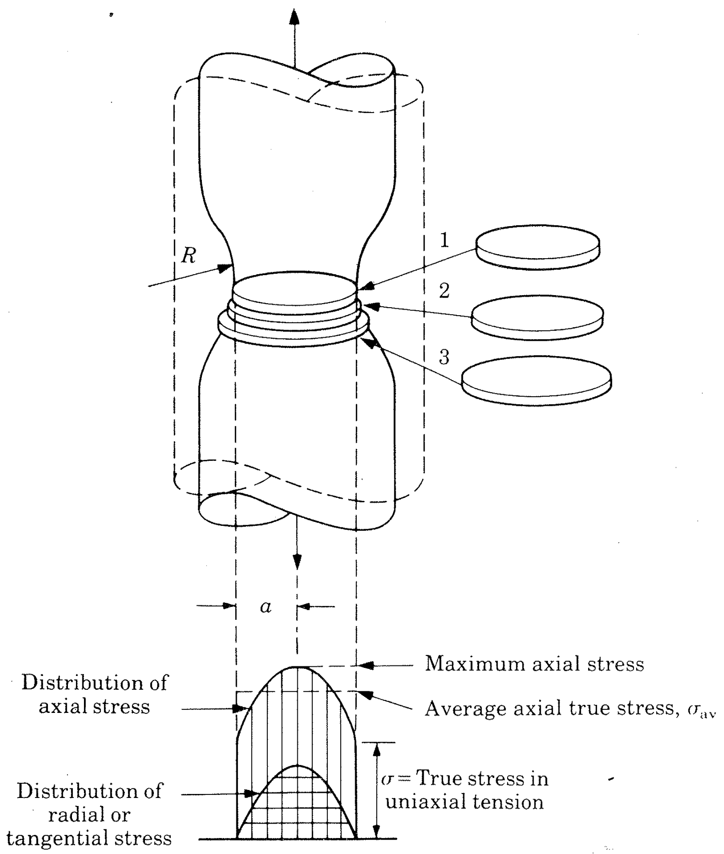


FIGURE P2.3

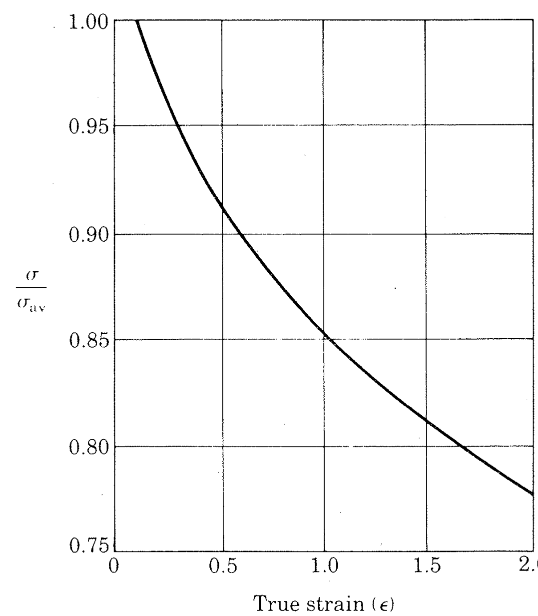


FIGURE P2.4

been established between a/R and the true strain at the neck, Fig. P2.4. The corrected true stress–true strain curve is shown in Fig. P2.2. ●

2.8.3 PLANE STRESS AND PLANE STRAIN

These two states of stress are important in the application of yield criteria. *Plane stress* is the state of stress in which one or two of the pairs of faces on an elemental cube are free of any stresses. An example is the torsion of a thin-walled tube. There are no stresses normal to the inside or outside surface of the tube; hence the state of the stress of the tube is one of plane stress. Other examples are shown in Figs. 2.44(a) and 2.47(a).

The state of stress where one of the pairs of faces on an element undergoes zero strain is known as *plane strain.* An example is shown in Fig. 2.47(b) depicting a piece of material being compressed in a die; note that one pair of faces is touching the walls (groove) of the die and cannot expand. Another example is the *plane strain compression test* shown in Fig. 2.27. Here, by proper choice of specimen dimensions, the width of the specimen is kept essentially constant. (Note that an element does not have to be physically constrained on the pair of faces for plane strain conditions to exist.) A third example is the torsion of a thin-walled tube in which the wall thickness remains constant (see Section 2.8.7).

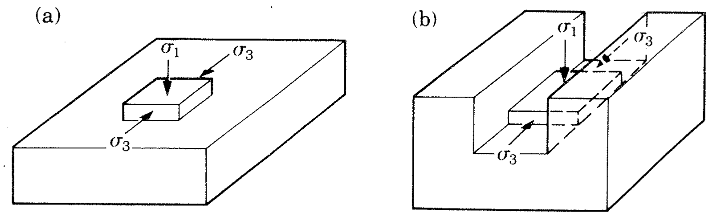


FIGURE 2.47 Two examples of states of stress. (a) Plane stress. (b) Plane strain.

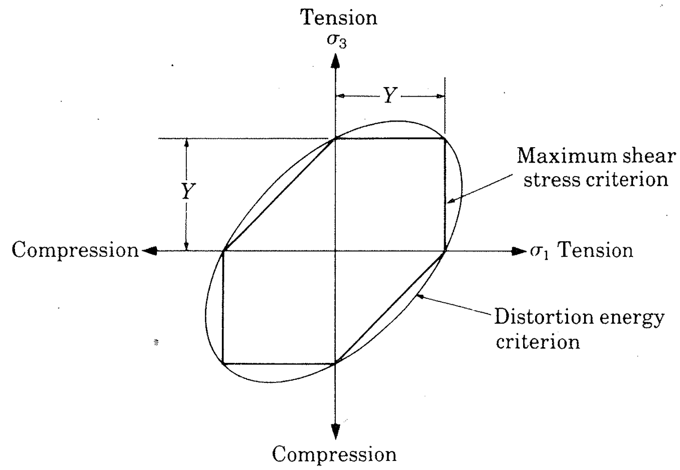


FIGURE 2.48 Plane stress diagrams for maximum shear stress and distortion energy criteria. The stresses applied should fall on or outside these curves to cause yielding. (It is assumed that the yield stress Y in tension and compression are equal, as they are for ductile but not for brittle materials.)

A review of the two yield criteria outlined above indicates that the plane stress condition can be represented by the diagram in Fig. 2.48. It can be seen that the maximum-shear-stress criterion gives an envelope of straight lines. The distortion-energy criterion for plane stress reduces to the equation

$$\sigma_1^2 + \sigma_3^2 - \sigma_1\sigma_3 = Y^2 \tag{2.37}$$

and is shown graphically in the figure. Whenever a point (with its coordinates representing the two principal stresses) falls on these boundaries, the element will yield.

Yielding under plane strain conditions requires the determination of the stress level, if any, on the faces of the element undergoing plane strain (Fig. 2.47b). This is done through the use of *generalized Hooke's law* equations shown below.

$$\begin{aligned} \epsilon_1 &= \frac{1}{E}[\sigma_1 - \nu(\sigma_2 + \sigma_3)] \\ \epsilon_2 &= \frac{1}{E}[\sigma_2 - \nu(\sigma_1 + \sigma_3)] \\ \epsilon_3 &= \frac{1}{E}[\sigma_3 - \nu(\sigma_1 + \sigma_2)] \end{aligned} \tag{2.38}$$

For plane strain, one of the strains in these equations is zero. This means that there exists a particular relationship between the principal stresses. At yielding, the Poisson's ratio ν is 0.5 (since the volume of an element undergoing plastic deformation is found to remain constant) and, therefore, all three stresses can be determined. Thus, for the case in Fig. 2.47(b),

$$\sigma_2 = \frac{\sigma_1 + \sigma_3}{2} \tag{2.39}$$

Note that σ_2 is now an intermediate stress.

For the plane strain compression of Fig. 2.47, the distortion-energy criterion (which includes the intermediate stress) reduces to

$$\sigma_1 - \sigma_3 = \frac{2}{\sqrt{3}} Y = 1.15Y = Y'.$$

● **Illustrative Problem 2.9**

A specimen in the shape of a cube 10 mm on each side is being compressed without friction in a die cavity as shown in Fig. 2.47(b), where the width of the groove is 15 mm. Assume that the material is made of an elastic, linearly strain-hardening material given by

$$\sigma = 100 + 20\epsilon \text{ MPa}.$$

Calculate the compressive force required when the height of the specimen is 3 mm, according to both yield criteria.

SOLUTION. It can be seen that, when the height of this specimen is reduced to 3 mm, the surface area A will be

$$(10)(10)(10) = (3)(A).$$

Hence,

$$A = 333.3 \text{ mm}^2.$$

Since the groove is only 15 mm wide, the specimen will touch the walls of the groove, because $(15)(15) = 225 \text{ mm}^2$, which is smaller than the final surface area required. Thus, this is both a plane strain and also a plane stress problem.

The true strain after deformation is (in absolute value)

$$|\epsilon| = \ln\left(\frac{10}{3}\right) = 1.204.$$

Hence the strength level (flow stress, see Fig. 2.51) that the material will exhibit at this strain is

$$Y_f = 100 + (20)(1.204) = 124.08 \text{ MPa}.$$

According to the maximum-shear-stress criterion, the force required is

$$P = Y_f A = (124.08)(333.3) = 41{,}350 \text{ N} = 41.35 \text{ kN}.$$

According to the distortion-energy criterion,

$$Y_f = (1.15)(124.08) = 142.69 \text{ MPa}.$$

Hence,

$$P = (142.69)(333.3) = 47{,}560 \text{ N} = 47.56 \text{ kN} \quad \bullet$$

2.8.4 EXPERIMENTAL VERIFICATION OF YIELD CRITERIA

The yield criteria described above have been tested experimentally. A suitable specimen commonly used is a thin-walled tube under internal pressure and/or torsion. Under such loading, it is possible to generate different states of plane stress. Various experiments, with a variety of ductile materials, have shown that the distortion energy criterion agrees better with the experimental data than does the maximum-shear-stress criterion (Fig. 2.49).

This would suggest that one should use the distortion-energy criterion for the analysis of metalworking processes since the latter generally makes use of ductile materials. On the other hand, the simpler maximum-shear-stress criterion can also be used, particularly by designers, since the difference between the two is negligible for most practical applications.

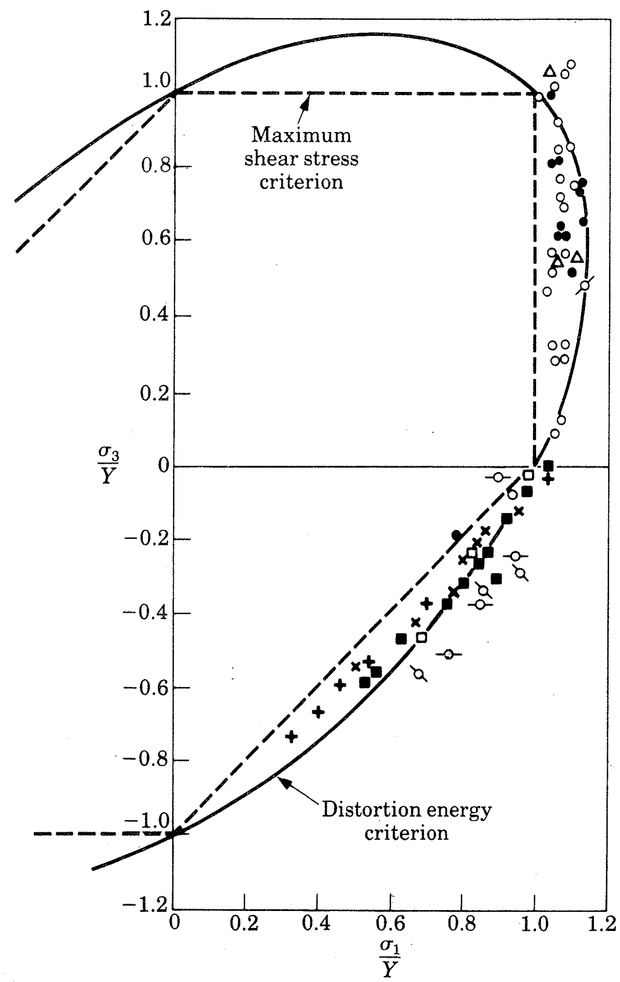


FIGURE 2.49 Verification of yield criteria by experimental data for various materials, superimposed on plane-stress diagrams. Note that the data points fall closer to the distortion-energy criterion than to the maximum-shear-stress criterion.

2.8.5 VOLUME STRAIN

By summing the three equations of the generalized Hooke's law it is seen that

$$\epsilon_1 + \epsilon_2 + \epsilon_3 = \frac{1 - 2\nu}{E}(\sigma_1 + \sigma_2 + \sigma_3), \tag{2.40}$$

where the left-hand side of the equation can be shown to be the volume strain or *dilatation* Δ. Thus,

$$\Delta = \frac{\text{vol. change}}{\text{orig. vol.}} = \frac{1 - 2\nu}{E}(\sigma_1 + \sigma_2 + \sigma_3). \tag{2.41}$$

It can now be shown that, in the plastic range, where $\nu = 0.5$, the volume change is zero. Hence, in plastic working,

$$\epsilon_1 + \epsilon_2 + \epsilon_3 = 0, \tag{2.42}$$

which is a convenient means of determining a third strain if two strains are known.

The *bulk modulus* is defined as

$$\text{Bulk modulus} = \frac{\sigma_{av}}{\Delta} = \frac{E}{3(1 - 2\nu)}, \tag{2.43}$$

where

$$\sigma_{av} = \tfrac{1}{3}(\sigma_1 + \sigma_2 + \sigma_3). \tag{2.44}$$

From Eq. (2.41) it can be seen that, in the elastic range, where $0 < \nu < 0.5$, the volume of a tension-test specimen increases and that of a compression specimen decreases.

2.8.6 EFFECTIVE STRESS AND EFFECTIVE STRAIN

A convenient means of expressing the state of stress on an element is the *effective* (*equivalent* or *representative*) *stress* $\bar{\sigma}$ and *effective strain* $\bar{\epsilon}$ given by the equations below. For the maximum-shear-stress criterion,

$$\bar{\sigma} = \sigma_1 - \sigma_3, \tag{2.45}$$

and for the distortion-energy criterion,

$$\bar{\sigma} = \frac{1}{\sqrt{2}}[(\sigma_1 - \sigma_2)^2 + (\sigma_2 - \sigma_3)^2 + (\sigma_3 - \sigma_1)^2]^{1/2}. \tag{2.46}$$

The factor $1/\sqrt{2}$ is chosen so that, for simple tension, the effective stress is equal to the uniaxial yield stress Y.

The strains are likewise related to the effective strain. For the maximum-shear-stress criterion,

$$\bar{\epsilon} = \tfrac{2}{3}(\epsilon_1 - \epsilon_3), \tag{2.47}$$

and for the distortion-energy criterion,

$$\bar{\epsilon} = \frac{\sqrt{2}}{3}[(\epsilon_1 - \epsilon_2)^2 + (\epsilon_2 - \epsilon_3)^2 + (\epsilon_3 - \epsilon_1)^2]^{1/2}. \tag{2.48}$$

Again, the factors 2/3 and $\sqrt{2}/3$ are chosen so that for simple tension the effective strain is equal to the uniaxial tensile strain. It is apparent that stress–strain curves may also be called effective stress–effective strain curves.

2.8.7 COMPARISON OF NORMAL STRESS–STRAIN AND SHEAR STRESS–STRAIN CURVES

Stress–strain curves in tension and torsion for the same material are, of course, comparable. Also, it is possible to construct one curve from the other since the material is the same. The procedure for this conversion is outlined below.

The following observations are made in regard to Fig. 2.50, showing the tension

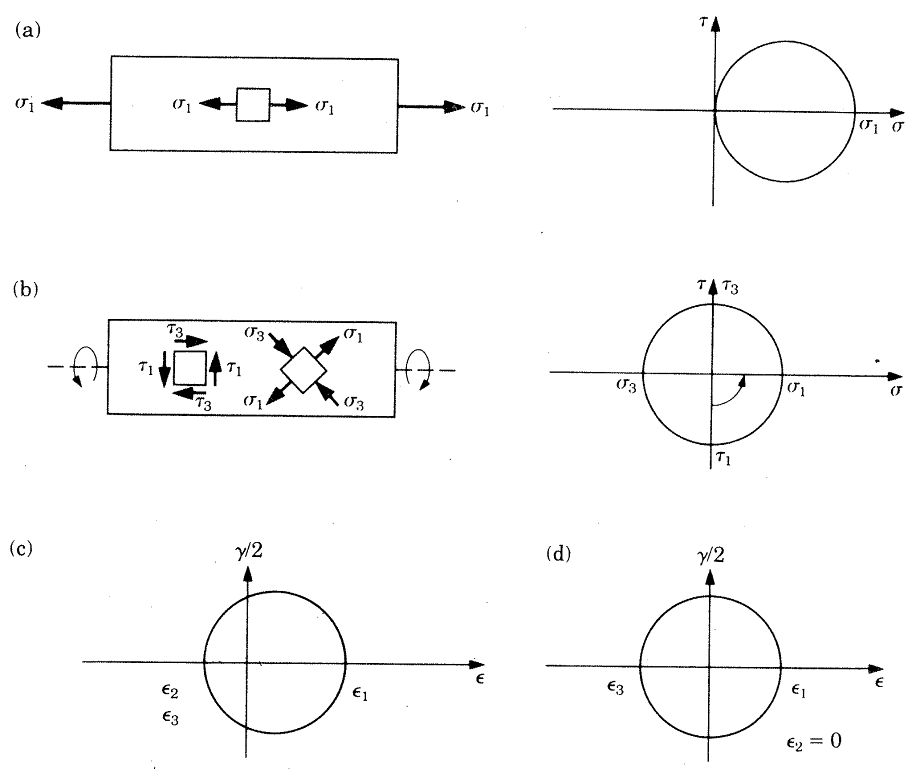


FIGURE 2.50 Mohr's circle diagrams for stress and strain in uniaxial tension, (a) and (c), and torsion, (b) and (d). Guy wires for antennas, or the spokes of a bicycle, are subjected to uniaxial tension (because they are thin and long), whereas the drive shaft of an automobile is subjected to the stresses shown in (b).

and torsional states of stress:

a. In the tension test, the uniaxial stress σ_1 is also the effective stress and the principal stress.

b. In the torsion test, the principal stresses occur on planes whose normals are at 45 degrees to the longitudinal axis; the principal stresses σ_1 and σ_3 are equal in magnitude but opposite in sign.

c. The magnitude of the principal stress in torsion is the same as the maximum shear stress.

We now have the following relationships:

$$\sigma_1 = -\sigma_3,$$

$$\sigma_2 = 0,$$

$$\sigma_1 = \tau_1.$$

Substituting these stresses in Eqs. (2.45) and (2.46) for effective stress, the following relationships are obtained: For the maximum-shear-stress criterion,

$$\bar{\sigma} = \sigma_1 - \sigma_3 = \sigma_1 + \sigma_1 = 2\sigma_1 = 2\tau_1, \tag{2.49}$$

and for the distortion-energy criterion,

$$\bar{\sigma} = \frac{1}{\sqrt{2}}[(\sigma_1 - 0)^2 + (0 + \sigma_1)^2 + (-\sigma_1 - \sigma_1)^2]^{1/2} = \sqrt{3}\sigma_1 = \sqrt{3}\tau_1. \tag{2.50}$$

With regard to strains, the following observations can be drawn from Fig. 2.50:

a. In the tension test, $\epsilon_2 = \epsilon_3 = -\dfrac{\epsilon_1}{2}$.

b. In the torsion test, $\epsilon_1 = -\epsilon_3 = \dfrac{\gamma}{2}$.

c. The strain in the thickness direction of the tube is zero, i.e., $\epsilon_2 = 0$.

Observation (c) is true because the thinning due to the principal tensile stress is countered by the thickening under the principal compressive stress of the same magnitude. Hence, $\epsilon_2 = 0$. Since σ_2 is also zero, a thin-walled tube under torsion is both a plane stress and a plane strain situation.

Substituting these strains in Eqs. (2.47) and (2.48) for effective strain, the following relationships are obtained: For the maximum-shear-stress criterion,

$$\bar{\epsilon} = \tfrac{2}{3}(\epsilon_1 - \epsilon_3) = \tfrac{2}{3}(\epsilon_1 + \epsilon_1) = \tfrac{4}{3}\epsilon_1 = \tfrac{2}{3}\gamma, \tag{2.51}$$

and for the distortion-energy criterion,

$$\bar{\epsilon} = \frac{\sqrt{2}}{3}[(\epsilon_1 - 0)^2 + (0 + \epsilon_1)^2 + (-\epsilon_1 - \epsilon_1)^2]^{1/2} = \frac{2}{\sqrt{3}}\epsilon_1 = \frac{1}{\sqrt{3}}\gamma. \tag{2.52}$$

This set of equations provides a means by which tensile-test data can be converted to torsion-test data, and vice versa.

2.9 WORK OF DEFORMATION

In this section the work required for plastic deformation of materials will be discussed. Since work is defined as the product of force and distance (collinear with force), a quantity equivalent to work per unit volume is the product of stress and strain. Since the relation between stress and strain in the plastic range depends on the particular stress–strain curve, this work is best calculated by referring to Fig. 2.51.

Note that the area under the true stress–true strain curve for any strain ϵ_1 is the *energy per unit volume u* (*specific energy*) of the material deformed. This is expressed as

$$u = \int_0^{\epsilon_1} \sigma \, d\epsilon. \tag{2.53}$$

As seen in Section 2.2.3, true stress–true strain curves can be represented by the expression

$$\sigma = K\epsilon^n.$$

Hence, Eq. (2.53) can be written as

$$u = K \int_0^{\epsilon_1} \epsilon^n \, d\epsilon$$

or

$$u = \frac{K\epsilon_1^{n+1}}{n+1} = \overline{Y}\epsilon_1, \tag{2.54}$$

where $\overline{Y}$ is the average flow stress of the material.

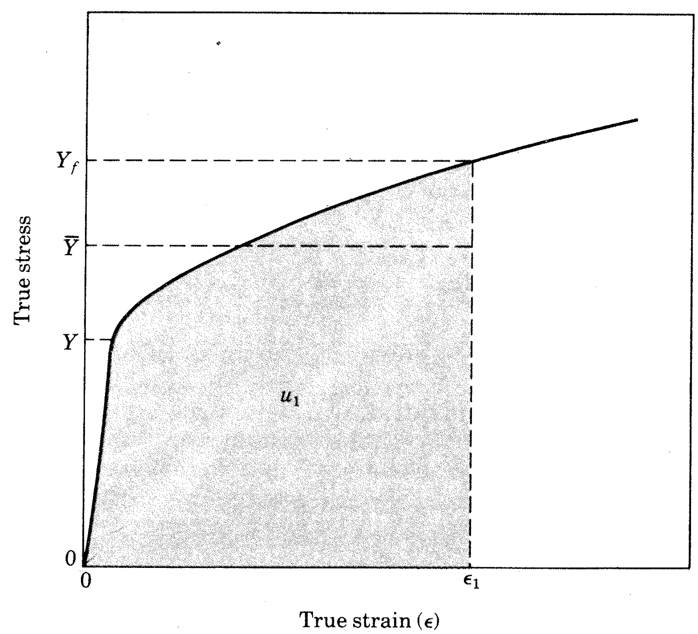


FIGURE 2.51 Schematic illustration of true stress-true strain curve showing yield stress Y, average flow stress $\overline{Y}$, specific energy u_1, and flow stress Y_f. Flow stress is defined as the true stress required to continue plastic deformation at a particular true strain. Thus, for strain-hardening materials there is a flow stress for each strain.

This energy represents the work dissipated in uniaxial deformation. For a more general condition, where the workpiece is subjected to triaxial stresses and strains, the effective stress and effective strains can be used. The energy per unit volume is then

$$u = \int_0^{\bar{\varepsilon}} \bar{\sigma}\, d\bar{\varepsilon}. \tag{2.55}$$

To obtain the work expended, we multiply u by the volume of the material deformed. Thus,

$$\text{Work} = (u)(\text{volume}). \tag{2.56}$$

The energy represented by Eq. (2.54) is the minimum energy or the *ideal* energy required for uniform (homogeneous) deformation. The energy required for actual deformation involves two additional factors. One is the energy required to overcome friction at the die–workpiece interfaces. The other is the *redundant work* of deformation, which is described as follows:

In Fig. 2.52(a), a block of material is being deformed into shape, such as by forging, extrusion, or drawing through a die as described in Chapter 6. As shown in sketch (b) this deformation is uniform, or homogeneous. In reality, however, the material more often than not deforms as in sketch (c) due to the effects of friction and die geometry. The difference between (b) and (c) is that (c) has undergone additional shearing along horizontal planes.

This shearing requires expenditure of energy since additional plastic work has to be done in subjecting the various layers to shear strains. This is known as *redundant work*; the word redundant is due to the fact that this work does not contribute to the shape change of the material. [Note that (b) and (c) have the same overall shape and dimensions.]

The total specific energy required can now be written as

$$u_{\text{total}} = u_{\text{ideal}} + u_{\text{friction}} + u_{\text{redundant}}. \tag{2.57}$$

The efficiency of a process is defined as

$$\eta = \frac{u_{\text{ideal}}}{u_{\text{total}}}. \tag{2.58}$$

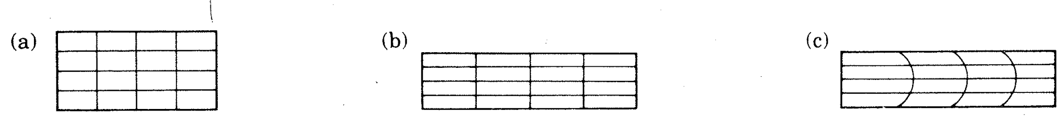


FIGURE 2.52 Deformation of grid patterns in a workpiece. (a) Original pattern. (b) After ideal deformation. (c) After inhomogeneous deformation, requiring redundant work of deformation. Note that (c) is basically (b) with additional shearing, especially at the outer layers. Thus, part (c) requires greater work of deformation than part (b).

The magnitude of this efficiency varies widely, depending on the particular process, frictional conditions, die geometry, and other process parameters. Typical values are estimated to be 30 to 60% for extrusion and 75 to 95% for rolling.

- **Illustrative Problem 2.10**

A thin-walled spherical shell, made of a perfectly plastic material of yield stress Y, original radius r_o and thickness t_o, is being expanded by internal pressure. Calculate the work done in expanding this shell to a radius of r_f. If the diameter expands at a constant rate, what changes take place in the power consumed as the radius increases?

SOLUTION. The membrane stresses are given by

$$\sigma_1 = \sigma_2 = Y$$

(from Illustrative Problem 2.7), where r and t are instantaneous dimensions. The true strains in the membrane are given by

$$\epsilon_1 = \epsilon_2 = \ln\left(\frac{2\pi r_f}{2\pi r_o}\right) = \ln\left(\frac{r_f}{r_o}\right).$$

Because an element in this shell is subjected to equal biaxial stretching, the specific energy is

$$u = \int_o^{\epsilon_1} \sigma_1\, d\epsilon_1 + \int_o^{\epsilon_2} \sigma_2\, d\epsilon_2 = 2\sigma_1\epsilon_1 = 2Y \ln\left(\frac{r_f}{r_o}\right).$$

Since the volume of the shell material is $4\pi r_o^2 t_o$, the work done is

$$W = (u)(\text{volume}) = 8\pi Y r_o^2 t_o \ln\left(\frac{r_f}{r_o}\right)$$

The specific energy can also be calculated from the effective stresses and strains. Thus, according to the distortion-energy criterion,

$$\bar{\sigma} = \frac{1}{\sqrt{2}}[(o)^2 + (\sigma_2)^2 + (-\sigma_1)^2]^{1/2} = \sigma_1 = \sigma_2$$

and

$$\bar{\epsilon} = \frac{\sqrt{2}}{3}[(o)^2 + (\epsilon_2 + 2\epsilon_2)^2 + (-2\epsilon_2 - \epsilon_2)^2]^{1/2} = 2\epsilon_2 = 2\epsilon_1.$$

(The thickness strain $\epsilon_3 = -2\epsilon_2 = -2\epsilon_1$ because of volume constancy in plastic deformation, where $\epsilon_1 + \epsilon_2 + \epsilon_3 = 0$.) Hence,

$$u = \int_o^{\bar{\epsilon}} \bar{\sigma}\, d\bar{\epsilon} = \int_o^{2\epsilon_1} \sigma_1\, d\epsilon_1 = 2\sigma_1\epsilon_1.$$

Thus the answer is the same.

Power is defined as

$$\text{Power} = \frac{dW}{dt}.$$

The expression for work can be written as

$$W = K \ln\left(\frac{r}{r_o}\right) = K(\ln r - \ln r_o),$$

since all other factors in the expression are constant. Thus,

$$\text{Power} = \frac{K}{r}\frac{dr}{dt}.$$

Because the shell is expanding at a constant rate, dr/dt = constant. Hence the power is related to the instantaneous radius r by

$$\text{Power} \propto \frac{1}{r}. \quad \bullet$$

2.9.1 WORK AND HEAT

Almost all the mechanical work of deformation in plastic working is converted into *heat*. This conversion of work to heat is not 100% because a small portion of this energy is stored within the deformed material as elastic energy. This is known as *stored energy* (discussed in Section 3.5). Stored energy is generally 5 to 10% of the total energy input. However, it may be as high as 30% in some alloys.

In a simple frictionless process and assuming that work is completely converted into heat, the temperature rise is given by

$$\Delta T = \frac{u_{\text{total}}}{\rho c}, \tag{2.59}$$

where u_{total} is the specific energy from Eq. (2.57), 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 smaller values of specific heat.

The theoretical temperature rise for a true strain of 1 (such as a 27-mm-high specimen compressed down to 10 mm) has been calculated to be as follows:

Aluminum	165°F	(75°C)
Copper	285	(140)
Low-carbon steel	535	(280)
Titanium	1060	(570)

The temperature rise given by Eq. (2.59) is for an ideal situation, where there is no heat loss. In actual operations heat is lost to the environment, to the tools and dies, and to any lubricants or coolants used. If the process is performed very rapidly, these losses are relatively small.

Under extreme conditions, an adiabatic process is approached, with very high temperature rise, leading to *incipient melting*. This rise in temperature can be calculated provided the stress–strain curve used is at the appropriate strain rate level. On the other hand, if the process is carried out slowly, the actual temperature rise will be a small portion of the calculated value. It should also be pointed out that properties such as specific heat and thermal conductivity depend on temperature and this should be taken into account in the calculations.

● **Illustrative Problem 2.11**

Calculate the total work done for the specimen in Illustrative Problem 2.2. Calculate the specific energy for an element in the necked area and the theoretical temperature rise.

SOLUTION. The total work done on the specimen can be obtained from the area under the curve in Fig. P2.1. This area can be determined graphically and is found to be

$$\text{Work} = 3950 \text{ in.} \cdot \text{lb.}$$

The specific energy of an element in the necked area, i.e., at fracture, is the area under the true stress–true strain curve (corrected) in Fig. P2.2, namely,

$$u = \int_{o}^{1.253} \sigma \, d\epsilon = 155{,}000 \text{ in.} \cdot \text{lb/in}^3.$$

The theoretical temperature rise, i.e., adiabatic, is obtained from Eq. (2.59), where, for stainless steel, $\rho = 0.29$ lb/in^3 and $c = 0.12$ Btu/lb $\cdot$ °F. Thus,

$$\Delta T = \frac{155{,}000}{(0.29)(0.12)(778)(12)} = 477°\text{F} = 247°\text{C}.$$

Note that in these calculations only the work of plastic deformation is considered since no friction or redundant work is involved in a simple tension test. Also, the actual temperature rise will be lower because of the heat loss from the necked zone to the rest of the specimen and to the environment. ●

2.10 METHODS OF ANALYSIS OF METALWORKING PROCESSES

The treatments thus far in this chapter pertain generally to simple situations of tension, compression, and torsion where the state of stress can easily be defined. In actual metalworking processes, however, the material is generally subjected to a complex state of stress along with the additional component of friction at the die–workpiece interfaces. Furthermore, these stresses and friction generally vary along the interfaces.

The accurate calculation or prediction of loads, forces, stresses, and temperatures is essential not only for the proper design of equipment but also for studying the behavior of the material during processing.

In this section, the major methods of analysis used for deformation processing of materials will be reviewed briefly. The advantages and limitations of each technique are described individually.

2.10.1 SLAB METHOD OF ANALYSIS

This is one of the simpler methods of analyzing the stresses and loads in plastic deformation of materials. It requires the selection of an element in the workpiece and identifying all the normal and frictional forces acting on this element.

As an example, let us take the case of a simple compression process with friction (Fig. 2.53a), which is the basic deformation mode in forging. The purpose of this analysis is to determine the die pressure distribution from which one can then calculate the load required.

As the flat dies compress the part, it is reduced in thickness and, since the volume remains constant, the part expands laterally. This relative movement at the die–workpiece interfaces causes frictional forces acting in opposition to the movement of the piece. These frictional forces are shown by the horizontal arrows in Fig. 2.53(a). For simplicity, let us also assume that the deformation is in plane strain, i.e., the workpiece is not free to flow in the direction perpendicular to this page.

Let us now take an element, as shown in the figure, and indicate all the forces acting on it (Fig. 2.53b). Note the correct direction of the frictional forces. Also note the difference in the horizontal forces acting on the sides of the element; this difference is due to the presence of frictional forces on the element. It is assumed that the lateral stress distribution σ_x is uniform along the height h.

The next step in this analysis is to balance the horizontal forces on this element since the element must be in static equilibrium. Thus,

$$(\sigma_x + d\sigma_x)h + 2\mu\sigma_y\,dx - \sigma_x h = 0$$

or

$$d\sigma_x + \frac{2\mu\sigma_y}{h}\,dx = 0.$$

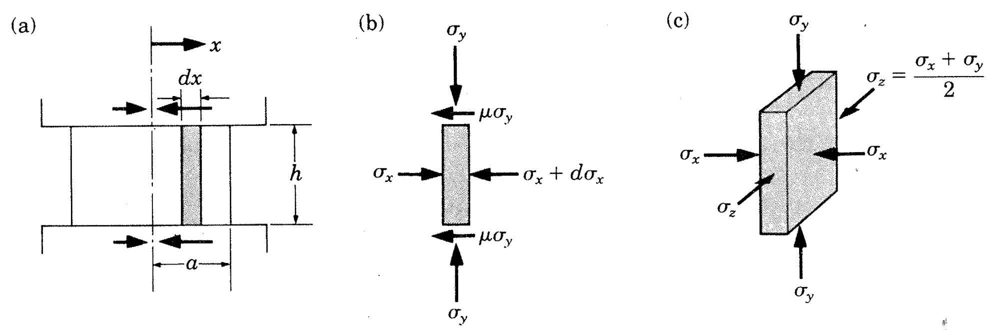


FIGURE 2.53 Stresses on an element in plane-strain compression (forging) between flat dies. The stress σ_x is assumed to be uniformly distributed along the height h of the element. Identifying the stresses on an element (slab) is the first step in the slab method of analysis for metalworking processes, described in Section 2.10.1, and used extensively in Chapter 6.

Note that we have one equation but two unknowns. The second equation is obtained from the yield criteria as follows. As shown in Fig. 2.53(c), this element is subjected to triaxial compression (compare this element with Fig. 2.47b). Using the distortion-energy criterion for plane strain, the following expression is obtained:

$$\sigma_y - \sigma_x = \frac{2}{\sqrt{3}} Y = Y'. \tag{2.60}$$

Thus,

$$d\sigma_y = d\sigma_x.$$

Note that σ_y and σ_x are assumed to be principal stresses. In the strictest sense, σ_y cannot be a principal stress because a shear stress is also acting on the same plane. However, this assumption is acceptable for low values of the coefficient of friction μ and is the standard practice in this method of analysis.

There are now two equations that can be solved by noting that

$$\frac{d\sigma_y}{\sigma_y} = -\frac{2\mu}{h}\,dx$$

or

$$\sigma_y = Ce^{-2\mu x/h}.$$

The boundary conditions are such that at $x = a$, $\sigma_x = 0$ and thus $\sigma_y = Y'$ at the edges of the specimen. (Since all stresses are compressive, we may ignore negative signs for stresses. This is traditional in such analysis. See also Chapter 6.) Hence the value of C becomes

$$C = Y'e^{2\mu a/h},$$

and thus

$$p = \sigma_y = Y'e^{2\mu(a-x)/h}. \tag{2.61}$$

Also,

$$\sigma_x = \sigma_y - Y' = Y'[e^{2\mu(a-x)/h} - 1]. \tag{2.62}$$

Equation (2.61) is plotted qualitatively in Fig. 2.54 in dimensionless form. Note

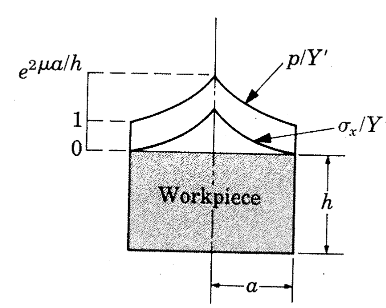


FIGURE 2.54 Distribution of die pressure p in plane-strain compression with sliding friction. Note that the pressure at the left and right boundaries is equal to the yield stress in plane strain, Y'. Sliding friction means that the frictional stress is directly proportional to the normal stress. See Section 4.6.

that the pressure increases exponentially toward the center of the part and also that it increases with the a/h ratio. Further coverage of these formulas is given in Section 6.3.1 on forging.

Because of its simplicity, this method of analysis has been used widely for almost all metalworking processes and is quite useful in estimating forces in processing. Other examples are given in Chapter 6.

2.10.2 SLIP-LINE ANALYSIS

This method has been applied generally to plane-strain conditions. The deforming body is assumed to be rigid, perfectly plastic, and isotropic. The technique consists of the construction of a family of straight or curvilinear lines that intersect each other orthogonally (see also Fig. 2.39). These lines, known as a *slip-line field*, correspond to the directions of yield stress of the material in shear, k.

The network of slip-lines, whose construction depends largely on intuition and experience and are postulated a priori, must satisfy certain conditions. These are: static equilibrium of forces, yield criteria, and boundary conditions. The slip-line fields must be checked for equilibrium conditions so that, for instance, the slip-lines meet at a free surface at a 45-degree angle. Also, they must be compatible with the velocity field. This means that the movement of the material must be such that mass continuity is maintained.

A simple two-dimensional example using this method is given in Fig. 2.55, simulating the hardness test with a flat rectangular indenter. In this model, the material is assumed to deform along shear planes as shown. The movements of the triangular blocks of material should be such that they can slide along each other, but

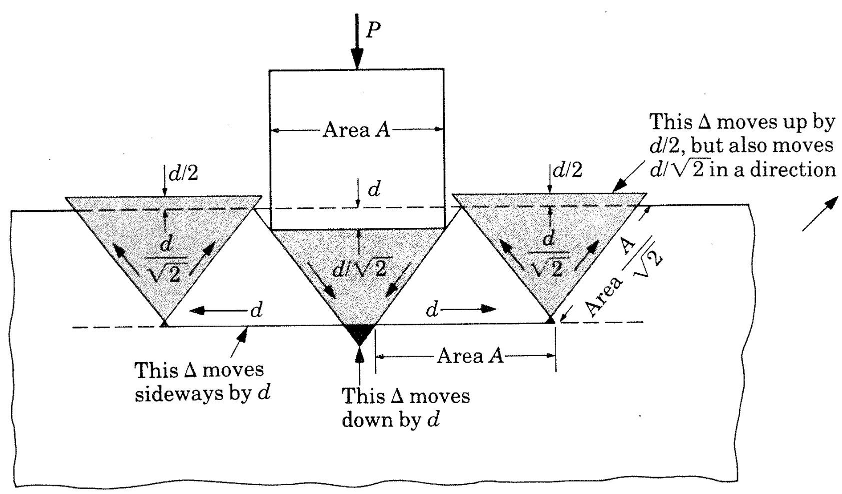


FIGURE 2.55 A simple two-dimensional model of slip-line analysis for indentation of a solid body with a flat rectangular punch. This model is similar to stacking of triangular wooden blocks. *Source*: Adapted from M. A. Ashby and D. R. H. Jones, *Engineering Materials*, Elmsford, N.Y.: Pergamon, 1980, p. 106.

cannot separate from each other. Thus the velocity normal to a plane of shear must be the same for points on either side of the plane. This is a simplified model. Instead of three triangular blocks, the deformation could consist of a large number of smaller blocks.

We can now determine the hardness of this material (P/A) as follows. The work done by the indenter is the product of P and d (assuming that P is not a function of the magnitude of d). This external work must be equal to the internal work done in shearing the blocks against each other. Letting k be the shear yield stress and noting that each triangular block has two sides, the following relation is obtained:

$$Pd = 2k\left(\frac{A}{\sqrt{2}}\,d\sqrt{2} + Ad + 2\frac{A}{\sqrt{2}}\frac{d}{\sqrt{2}}\right)$$

or

$$P = 6Ak.$$

Hence,

$$\text{Hardness} = \frac{P}{A} = 6k = 3Y. \tag{2.63}$$

Even with this very simplified model, one is able to obtain an expression for hardness that is in good agreement with Eq. (2.29).

Because a number of slip-line fields that satisfy all the conditions can be constructed, this method cannot give a unique solution for a particular deformation process. A solution must be verified by checking it with experimental data. (Another example of slip-line analysis is given in Fig. 8.11 in regard to chip formation in metal cutting.)

The slip-line method has been used successfully in the analysis of a number of metalworking processes (forging, rolling, extrusion) to predict stresses, loads, directions of material flow, and temperature variations within a workpiece. The frictional conditions at the die–workpiece interfaces can also be included in the analysis. Some success has also been obtained in studying axisymmetric cases (compression of a cylinder), strain-hardening effects, and anisotropic materials.

Many details are involved in the application of this technique. Therefore further treatment of this topic is beyond the scope of this text.

2.10.3 UPPER-BOUND TECHNIQUE

In this method, the overall deformation zone is divided into a number of smaller zones within which the velocity of a particle is continuous. The particle velocities in adjacent zones may be different. However, just as in slip-line analysis, at the boundaries between the zones, or between a zone and the die surface, all movement must be such that discontinuities in velocity occur only in the tangential direction.

This technique, which has been used successfully for axisymmetric cases, such as extrusion and rod drawing, is lengthy and gives results similar to those in slip-line analysis. In this analysis, the total power consumed in an operation is the sum of the following:

a. *Ideal* power of deformation,

b. Power consumed in *shearing* the material along velocity discontinuities, and

c. Power required in overcoming *friction* at the die–workpiece interfaces.

In the final analysis, a velocity field that minimizes the total calculated power is taken as the actual one and is subsequently compared with experimental data.

Theorems have been developed to obtain lower- and upper-bound solutions for the required loads. The lower-bound solution underestimates the load and hence is of no practical value. The upper-bound solutions, on the other hand, overestimate the load and are therefore of practical interest in metalworking operations.

2.10.4 VISIOPLASTICITY

This is an experimental technique developed to determine the strain rates and the stress distributions in the deformation zone. It consists of placing a grid pattern on a flat surface (or at an interface, such as the median plane in a solid cylinder) and observing the distortion of the grid after subjecting the specimen to an incremental (small) deformation. The process is then repeated a number of times.

From the distortions, the strain rates are calculated (since the strain and the time of incremental deformation are known) and then, with the use of plasticity equations, the stresses are calculated.

The visioplasticity technique can be used for plane strain and for axisymmetric cases. It is also used to validate predictions of stress distributions obtained from other analytical or numerical methods. Accuracy depends on the precision with which the grid patterns are placed and measured after each incremental deformation. Small errors in measurement can lead to significant errors in the calculation of stresses.

2.10.5 FINITE-ELEMENT METHOD

In this method, the deformation zone in an elastic–plastic body is divided into a number of elements interconnected at a finite number of nodal points. The actual velocity distribution is then approximated for each element. A set of simultaneous equations is then developed representing unknown velocity vectors. From the solution of these equations actual velocity distributions and the stresses are calculated.

This technique can incorporate friction conditions at the die–workpiece interfaces and actual properties of the workpiece material. It has been applied to relatively complex geometries in bulk deformation and in sheet-forming problems. Accuracy is influenced by the number and shape of the finite elements, the deformation increment, and the methods of calculation. To ensure accuracy, complex problems require extensive computations. The finite-element method gives a detailed outline of the actual stresses and strain distributions throughout the workpiece.

SUMMARY

1. The mechanical behavior of materials in manufacturing processes is related to their strength, ductility, elasticity, hardness, and the energy required for plastic deformation.

2. The behavior of materials depends on the particular material and on a number of other variables, such as temperature, strain rate, and state of stress.

3. Tests have been developed to obtain the relation between stress and strain as a function of the above variables. These relationships are important in assessing the behavior of a particular material in a manufacturing process, especially in regard to forces required and the capability of the material to undergo the desired deformation without failure.

4. Two important parameters are the strain-hardening capability (indicated by the strain-hardening exponent n) and the strain-rate sensitivity of the material (indicated by the exponent m).

5. How a material is subjected to a shape change is also important. This requires the study of the deformation-zone geometry. The manner in which a material is subjected to plastic deformation is important in determining the nature and extent of residual stresses in the part after all external forces are removed. These stresses are important in subsequent processing of the part or during its service life.

6. Since in actual metalworking operations the material is generally subjected to three-dimensional stresses, yield criteria have been developed to establish relationships between the uniaxial yield stress of the material (generally obtained from a tension test) and the stresses applied.

7. Various methods of analysis are available to calculate the stresses, forces, and energies required in deformation processing of materials. These analyses are important not only for the selection of appropriate equipment for metalworking, but also in their design.

BIBLIOGRAPHY

General Introductory Texts

Alexander JM, Brewer RC. *Manufacturing Properties of Materials.* London: Van Nostrand, London, 1963.

Ashby MF, Jones DRH. *Engineering Materials: An Introduction to their Properties and Applications.* New York: Pergamon, 1980.

Biggs WD. *The Mechanical Behavior of Engineering Materials.* New York: Pergamon, 1965.

Cook NH. *Manufacturing Analysis.* Reading, Mass.: Addison–Wesley, 1966.

Cottrell AH. *The Mechanical Behavior of Matter.* New York: Wiley, 1964.

Crane FAA. *Mechanical Working of Metals.* New York: Macmillan, 1964.

Davis HE, Troxell GE, Hauck GFW. *The Testing of Engineering Materials*, 4th ed. New York: McGraw-Hill, 1982.

Dieter GE. *Mechanical Metallurgy*, 2d ed. New York: McGraw-Hill, 1976.

Honeycombe RWK. *Plastic Deformation of Metals*. London: Edward Arnold, 1968.

Marin J. *Mechanical Behavior of Engineering Materials*. Englewood Cliffs, N.J.: Prentice-Hall, 1962.

McClintock FA, Argon AS, eds. *Mechanical Behavior of Materials*. Reading, Mass.: Addison–Wesley, 1966.

Moss JB. *Properties of Engineering Materials*. London: Butterworths, 1971.

Mott BW. *Micro-Indentation Hardness Testing*. London: Butterworths, 1956.

O'Neill H. *Hardness Measurement of Metals and Alloys*. London: Chapman and Hall, 1967.

Parkins RN. *Mechanical Treatment of Metals*. London: Allen and Unwin, 1968.

Polakowski NH, Ripling EJ. *Strength and Structure of Engineering Materials*. Englewood Cliffs, N.J.: Prentice-Hall, 1964.

Pugh HLD. *Mechanical Behaviour of Materials under Pressure*. New York: Elsevier, 1971.

Schey JA. *Introduction to Manufacturing Processes*, New York: McGraw-Hill, 1977.

Suh NP, Turner APL. *Elements of the Mechanical Behavior of Solids*. New York: McGraw-Hill, 1975.

Tabor D. *The Hardness of Metals*. New York: Oxford, 1951.

Advanced Texts

Avitzur B. *Handbook of Metal-Forming Processes*. New York: Wiley, 1983.

Avitzur B. *Metal Forming: Processes and Analysis*. New York: McGraw-Hill, 1968.

Backofen WA. *Deformation Processing*. Reading, Mass.: Addison–Wesley, 1972.

Calladine CR. *Engineering Plasticity*. New York: Pergamon, 1969.

Ford H, Alexander JM. *Advanced Mechanics of Materials*, 2d ed. New York: Halsted, 1977.

Hill R. *The Mathematical Theory of Plasticity*. New York: Oxford, 1950.

Hoffman O, Sachs G. *Introduction to the Theory of Plasticity for Engineers*. New York: McGraw-Hill, 1953.

Johnson W, Mellor PB. *Engineering Plasticity*. New York: Van Nostrand, 1973.

Mendelson A. *Plasticity: Theory and Application*, New York: Macmillan, 1968.

Nadai A. *Theory of Flow and Fracture of Solids*, 2d ed. New York: McGraw-Hill, vol. I, 1950, vol. II, 1963.

Prager W, Hodge PG. *Theory of Perfectly Plastic Solids*. New York: Wiley, 1951.

Slater RA. *Engineering Plasticity: Theory and its Application to Metal Forming Processes*, New York: Halsted, 1974.

Thomsen EG, Yang CT, Kobayashi S. *Mechanics of Plastic Deformation in Metal Processing*. New York: Macmillan, 1964.

PROBLEMS

2.1. Explain why an offset method is used to determine the yield stress on a stress–strain curve in tension. Would it be necessary to use this method for a highly strain-hardened metal?

2.2. Using the same scale for stress, it can be seen that the tensile engineering stress–strain curve for a material is lower than its true stress–true strain curve. Is this also true for a compression test on the same material?

2.3. Assume that a material, with a uniaxial yield stress Y, yields under a stress system of principal stresses $\sigma_1, \sigma_2, \sigma_3$, where $\sigma_1 > \sigma_2 > \sigma_3$. Show that the superposition of a hydrostatic stress p on this system (such as placing the specimen in a pressurized chamber) does not affect yielding. In other words, the material will still yield according to the yield criteria.

2.4. A cylindrical specimen of a brittle material 1 in. high and 1 in. in diameter is subjected to compression. It is found that fracture takes place at an angle of 45 degrees under a load of 30,000 lb. Calculate the shear stress and the normal stress on the fracture surface.

2.5. A paper clip is made of wire 1 mm in diameter. If the original material from which the wire is made is a rod 10 mm in diameter, calculate the longitudinal and lateral (diametral) strains that the wire has undergone during processing.

2.6. A thin-walled tube is subjected to tension in the elastic range. Using the generalized Hooke's law equations, show that both the diameter and the thickness of this tube decrease when it is under tension.

2.7. Calculate the work involved in frictionless compression of a solid cylinder 30 mm high and 20 mm in diameter to a reduction of 50% for the following materials in Table 2.4: 6061-T6 aluminum, annealed copper, annealed 4340 steel, annealed 304 stainless steel.

2.8. In Illustrative Problem 2.2, calculate the toughness of the material by the area under the stress-strain curve. Check your answer by first determining the K and n values for this material and then obtaining the toughness by integration.

2.9. Explain how the modulus of resilience of a strain-hardening material changes as it is cold-worked.

2.10. What is the volume of a 50-mm diameter solid steel sphere when subjected to a hydrostatic pressure of 2 GPa?

2.11. A rod is 10 in. long. It is stretched in two steps, first to a length of 12 in. and then to 17 in. Show that the true strains are additive, whereas engineering strains are not.

2.12. If you remove the layer of material *ad* from the specimen in Fig. 2.41, such as by machining or grinding, which way will the specimen curve? (*Hint*: Assume that the bar in sketch (d) is composed of four horizontal springs held at the ends. Thus, from the top down: compression, tension, compression, tension springs.)

2.13. Explain why deformation rate and strain rate are not equivalent.

2.14. In Fig. 2.7, identify two materials that have the lowest and highest uniform elongation, respectively. Calculate these quantities as percentages.

2.15. A material has a strength coefficient $K = 100{,}000$ psi and $n = 0.2$. Assuming that a tensile-test specimen made from this material begins to neck at a true strain of 0.2, show that the ultimate tensile strength (engineering) of this material is 59,340 psi.

2.16. Modify the curves in Fig. 2.8 to indicate the effects of temperature on material properties.

2.17. Is it possible to completely remove residual stresses in a piece of material, by the technique described in Fig. 2.43, if the material is elastic, linearly strain hardening?

2.18. Using the generalized Hooke's law equations (Eq. (2.38)), show that a thin-walled tube does not undergo any thickness change when subjected to torsion.

Exhibit D

This realtime draft is unedited and uncertified and may contain untranslated stenographic symbols, an occasional reporter's note, a misspelled proper name and/or nonsensical word combinations.

This is an unedited version of the deposition transcript and should not be used in place of a certified copy.

This document should not be duplicated or sold to other persons or businesses. This document is not to be relied upon in whole or in part as the official transcript.

This uncertified realtime rough draft version has not been reviewed or edited by the certified shorthand reporter for accuracy. This unedited transcript is computer generated and random translations by the computer may be erroneous or different than that which will appear on the final certified transcript.

Due to the need to correct entries prior to certification, the use of this realtime draft is only for the purpose of augmenting counsel's notes and cannot be used to cite in any court proceeding or be distributed to any other parties.

Acceptance of this realtime draft is an automatic final copy order.

EXAMINATION

BY MR. GABRIC:

Q. Good morning, Mr. Meyst.

A. Good morning.

Q. Nice to see you again, sir.

A. Ditto.

Q. You're here to give a deposition in connection with an enter party review with respect to the '548 patent, correct?

A. Yes.

Q. All right. And the '548 patent is in the same family as the '742 patent?

A. Yes, it is.

Q. Okay. And you gave a deposition with respect to an IPR involving the '742 patent about a month ago, right?

A. Sounds about right.

Q. Give or take. Beginning of June.

A. Pardon?

Q. It was the beginning of June?

A. Okay.

Q. Sometime.

And you -- you recall you testified in that deposition about a reference we referred to as Hon '043 correct?
A. Yes.
Q. Okay. And the Hon '043 reference, that same reference is at issue in this IPR, correct?
A. That's correct.
Q. Okay. And you would agree with me, what you said in this -- that prior deposition in the '742 IPR regarding how Hon '043 operates Hon '043 operates for the -- in the same way with respect to this IPR as well, correct?
A. I believe so. It's the same patent.
Q. Fair enough. And why don't I just go ahead and mark it. We'll mark as Exhibit 1028 in this IPR deposition you gave in the '742 IPR and it's dated June 7, 2017. Ask you if you can identify whether that's your deposition transcript in the '742 IPR?
A. I believe it is my deposition transcript.
Q. All right.
(Exhibit ## was previously marked for identification and is attached hereto.)
BY MR. GABRIC:
Q. Now?

MR. HAMILTON: I just want to put on the record Mr. Meyst made some corrections to this transcript that don't appear to be included.

MR. GABRIC: Oh, I wasn't even aware of that. I apologize?

MR. HAMILTON: No, that's all right. I don't think they're of any substance but I just want to make sure that's on the record.

MR. GABRIC: Appreciate it. I'm always the last to learn things, Joe.

BY MR. GABRIC:

Q. Now, I'm going to show you what's been marked as Exhibit 1001 in this IPR. The '548 IPR. And for the record, this is the '548 patent. Correct?

A. Yes.

Q. Yes? All right. And the '548 patent has claims at the end, claims 1 through 14, correct?

A. You're correct.

Q. Okay. And claim 1 through 14 are directed to the embodiment of the '548 patent depicted in figure 17 and 18, correct?

MR. HAMILTON: Objection. Form.

THE WITNESS: We'll they're described I think in those embodiments. That doesn't mean it's

limited to those embodiments however.
BY MR. GABRIC:
Q. But you would agree with me figure 17 and 18 fall within the scope of claims 1 through 14?
MR. HAMILTON: Objection. Form.
THE WITNESS: Figure 17 and 18 contain elements described in the claims, yes.
BY MR. GABRIC:
Q. Okay. And -- and figure 17 and 18 in the '548 patent those are the same figures 17 and 18 found in the '742 patent? And I'm happy to give you that patent if you want it?
MR. HAMILTON: Objection. Form.
MR. GABRIC: Let me show you what's been marked as Exhibit 2002 in this IPR. And it's the '742 patent.
MR. HAMILTON: Thank you.
MR. GABRIC: You're welcome.
(Exhibit ## was previously marked for identification and is attached hereto.)
BY MR. GABRIC:
Q. Do you recognize that patent?
A. I have seen it before, yes.
Q. Fair enough. And figure 17 and 18 of the

'742 patent those are the same figures that are figures 17 and 18 in the '548 patent, correct?
A. I believe so, yes.
Q. All right. And at your deposition in the '742 IPR you testified about figure 17 and 18, correct?
MR. HAMILTON: Objection. Form.
THE WITNESS: I believe I did, yes. I'm just noticing there are some little differences between the figures. Because all the numbers aren't called out.
BY MR. GABRIC:
Q. I see. So the '548 patent has some references numerals that are not included in the corresponding figure of the '742 patent?
A. That's correct.
Q. Other than that, do you see any other differences?
A. I haven't studied it but I believe they're the same specification.
Q. Okay. And the devices depicted -- strike that.
Figure 17 and 18 of the '742 patent, those depict an atomizer, correct?
A. These are components of an atomizer.

Q. And that's also true with respect to figure 17 and 18 of the '548 patent?

MR. HAMILTON: Objection. Form.

THE WITNESS: Yes.

BY MR. GABRIC:

Q. Okay. And do the components illustrated in figure 17 and 18 of the '742 patent and figure 17 and 18 of the '548 patent, do they work in the same way?

A. I believe so, yes.

Q. Now, you've been deposed quite a few times over your career, correct?

A. That's correct.

Q. All right. And you generally understand the ground rules, correct?

A. I believe so.

Q. All right. And -- and I just want to remind you of a few things that are fairly important.

In these proceedings it's important that you don't discuss your testimony with counsel during breaks. You understand that?

A. I do.

Q. And I'll try to do a better job than I did last time of not talking over you and you should try not to talk over me. Let's let each other get our questions and answers out respectively okay?

A. I understand the process, yes.

Q. All right. Any reason you can't testify truthfully and accurately today?

A. No.

Q. All right. Now, I want to ask you a few questions about figures 18 -- Figure 18 of the '548 patent. If you want to get that in front of you.

A. Okay.

Q. Could you describe for me how these components -- what role they play in atomizing liquid?

A. Yes.

Q. Could you do that, please?

A. I will.

Q. Thank you.

A. Well, we could start with the item marked 83 which is the heating wire. And that is the device where an electricity can put energy in to effect change and phase of the materials. So it's the energy input.

81 is a porous component that the wire heating element is wrapped around and that trance supports liquid to the wire for atomization. And

what is labeled as 821 which is a hole, that is the location where at any air stream passes through the hole and then onto the -- the heating wire that's wrapped around the porous component. And 82 -- let me check something here. Just bear with me for a moment.

Q. Yeah, sure. Take your time.

A. I'm just looking for in the specification where it identifies the -- what they're calling item number 82. That's interesting. Because in Figure 6 of the patent, it says 82 is the electric heating rod. And that's obviously not correct.

Q. Yeah. You're looking for 82? I'm just trying to help you out. Column 5, line 65.

A. It's also called out in other places where 82 is identified as another component. The frame. I thought so. Okay. 82 is the frame. And the frame provides structural integrity for this.

Q. What do you mean structural integrity"?

A. Well, it's a rigid component.

Q. What's a rigid component?

A. 82 is a rigid component.

Q. Okay. And what -- what does it provide structural integrity for?

A. Well, we should probably look and see what the specification says.

Q. Probably a good starting point.

A. Well, the frame does a number of things as it says in column 5 of the '548 patent. The porous component is set on the frame. So it supports and specifically contains that porous component so that it can't move relative to the other components. It also creates the space for atomization which doesn't have a number. And the hole through it where air can come through, it also creates that passage way.

Q. So how does the frame contain the porous component?

MR. HAMILTON: Objection. Form.

THE WITNESS: I don't know if I said contain much it's set upon the element of the porous component where the wire is wrapped around passes through a hole. Actually two holes in the side of the -- the frame. And because it's in there, it can not go up, down, sideways, whatever it's totally contained and constrained in that location.

BY MR. GABRIC:

Q. So when you say "it's contained in that location, the porous component, because of these holes in the cavity wall cannot move at all?

MR. HAMILTON: Objection. Form.

THE WITNESS: Well if you were to crush this with pliers it could move. Certainly. But the purpose of that is to keep the heating element and the wire aligned with the hole so that the air coming through hits on that. There's likely other parts of this because it is the atomizer component of figure 19 and so in figure 19 we don't see any detail but we know that the atomizer shown in Figure 18 is in 19. And it doesn't show how it connects with the porous bulge but the -- the frame keeps the porous component oriented around the atomization cast, lines it up with the hole. Could you probably pull it out of there and destructively remove it. But it -- it can't go anywhere. There's no space for it to move in any direction.

BY MR. GABRIC:

Q. Why -- why does the porous component require the frame to keep it in place?

A. Why does the porous component require the frame?

Q. Right.

A. Well, if there wasn't a frame, what would the porous component be doing? Would it be hanging out waiting? It's part of the device so it's part of the structure.

Q. What if those -- and these are basically two cylinders we're -- well, strike that.

Figure 18 you have a portion of the frame inside of the porous component. You see that?

A. I guess you could say it's inside, yeah.

Q. Okay. Is that portion of the frame keeping the porous component in place?

A. Which portion.

Q. The portion that's inside the porous component.

A. Well, it's all keeping it. You can't break a piece of it away. I mean, the whole thing is providing a three-dimensional structure that meets up with he the form factor of the porous component such that it holds it where it needs to be to effect its function.

Q. I'm curious though from the perspective of one skilled in the art where is the internal portion of that frame, the portion that's internal to the porous component, why is it there? Why is it necessary?

MR. HAMILTON: Objection. Form.

THE WITNESS: When you say "the internal component, are you speaking the cylinder with the two holes in it that is part of item 82?

BY MR. GABRIC:

Q. Correct.

A. Well, we could go back to the specification.

Well, the porous component is not a rigid -- necessarily a rigid material. It could be. It could be -- have a wide range of properties. However, the purpose of having the frame with the porous component set on it and "set on" has a specific meaning as I'm sure you're aware so it is a specific location where it needs to be. It's not sitting on top of it. It is set opinion on that and integrated into it by passing through the hole and having pieces on both sides of the holes so it's a way of mounting it.

Q. What -- what does a rigidity of the porous component have to do with why you would have this portion of frame positioned internal to the porous component?

MR. HAMILTON: Objection. Form.

THE WITNESS: I just mentioned that it -- it could be based on testimony, it could be a very soft, pliable material. It needs to be supported in that function, in that position, in that location and that conviction to work. So it's part of the

design.
BY MR. GABRIC:
     Q.   And this patent, the '548 patent doesn't contemplate a frame that omits the portion of the frame that's internal to the porous component, correct?
          MR. HAMILTON: Objection. Form.
          THE WITNESS: I didn't understand your question.
BY MR. GABRIC:
     Q.   Do you -- do you see any discussion in the '548 patent where the patentee contemplated with respect to the components shown in Figure 18 that you could omit the portion of the frame that's located internal to the porous component?
     A.   Never thought of that. But it may be in there. I guess I would have to go back and read through the patent. But I don't believe that's the case. Would you like me to --
     Q.   No, no, that's okay. It's fine. It's Friday. We all want to get out as a reasonable hour.
          Now there's a -- is there a contacting fit between the frame 82 and the porous component 81?
          MR. HAMILTON: Objection. Form.

THE WITNESS: Where do you mean by "contacting"?

BY MR. GABRIC:

Q. Do they touch each other?

A. Well, they would have to touch each of other at some point.

Q. Okay. And do they touch -- where do they touch each other according to Figure 18?

A. Well, anyplace that the cross hatched item is in contact with item 82 is where they touch each other or 81 and 82 are sharing a common border.

Q. Now, under normal use of the device illustrated in figure 18 by the user, say they're smoking it and they're walking or trotting, would that porous component move whatsoever at all relative to the frame?

MR. HAMILTON: Objection. Form.

THE WITNESS: Well, as it's shown here, it would have no room to move anywhere.

BY MR. GABRIC:

Q. And what kind of fit would be necessary to quote have no room for it to move?

A. Well, it could be line to line or it could be an interference fit.

Q. What's an interference fit?

A. That would be where one part is larger than another part that it's put inside.

Q. And --

A. Did I say that right? I'm not -- so interference fit would be if the internal item is slightly larger than the external item that it's going into.

Q. So, for example, with respect to Figure 18, an interference fit would be where the diameter of the internal -- of the portion of the frame internal to the porous component is slightly larger than the diameter of the porus component so when they're stuck together you get an interfering fit?

MR. HAMILTON: Objection. Form.

BY MR. GABRIC:

Q. Do you want me to try that again? It was a lot?

A. Yeah, well, I think we all know what an interference fit is so I don't know what your point.

Q. I'm actually a chemist so I'm actually not pretending that I'm clueless I actually am on this stuff so could you explain to me what an interfering fit would be between the portion of the porus component 81 and the -- I'm sorry. Strike that.

Could you explain to me what an interfering fit would be with respect to the portion of the frame 82 that's internal to the porus component 81?

MR. HAMILTON: Objection. Form.

THE WITNESS: Well, there's a couple things here that could be considered. It doesn't all have to be an interference fit. For instance, the distance that the porous component which is number 81, the component part that has the wire wrapped around it, so if we look at it looks like a capital I. Okay?

Q. Uh-huh.

A. The vertical part between the top and the bottom, if that was just slightly less than the outside diameter where the holes pass through, that would cause the cylinder portion to be touching and against the -- the frame.

So, I mean, there's a number of ways. Basically one part the two parts don't have slot between them, there's no space with an interference fit. So one part -- it could be line to line which is where the dimensions are identical. Or it could be where one part is slightly bigger than the other.

Q. And so for there to be an interference fit with respect to the porus component 81 and frame 82, for example, the vertical portion of the I, the --

the external diameter of that vertical portion would be slightly greater than the diameter of the holes of the frame? That would be an example of a interfering fit?

MR. HAMILTON: Hold on a minute.

Objection. Form.

THE WITNESS: I don't think you said that correctly. But why don't you read me back when you said.

BY MR. GABRIC:

Q. For there to be an interference fit with respect to the porus component 81 and the frame 82, for example, the vertical portion of the eye, the external diameter of that vertical portion would be slightly greater than the diameter of the holes of the frame. That would be an example of an interfering fit?

MR. HAMILTON: Objection. Form.

THE WITNESS: The -- the -- the eye portion doesn't really have a diameter. It goes across the hole. So it -- the -- that -- there's really not a clear definition of where the diameter of that would -- I guess I'm misunderstanding your question.

BY MR. GABRIC:

Q. Okay. Let's back up. I'm just going to ask you to describe it.

And I want to focus on a different part of the porus component. The porus component 81 that wraps around the portion of the frame that's internal to the porous component, you see that? The horizontal portion of the frame?

A. Yes.

Q. Okay. What would be the relative diameters of the external portion of the frame at that location and the diameter of the -- internal diameter of the porous component 81 for there to be an interfering fit?

A. Well, it could either be the diameter of the frame 82 is slightly bigger than the internal diameter of 81 or 81 could be slightly smaller than the diameter of the rigid part. So --

Q. And does an interfering fit result in a friction fit?

MR. HAMILTON: Objection. Form.

THE WITNESS: A friction fit implies that there's forces involved. I mean, it could be put together, for instance, you can put a bearing together by heating it up and it actually expands, you put it on a heart and it will cool down and be tight on it. So that could be considered friction, I guess, it is pushing on it. I wouldn't call an interference fit necessarily a friction fit. It depends on how you put it together, I guess.

BY MR. GABRIC:

Q. If Figure 18 has an interfering fit the location between the porous component 81 and the frame that's the portion of the frame that's internal to the porus component, would that mean there's a friction fit at that location?

MR. HAMILTON: Objection. Form.

THE WITNESS: Well, we don't know about the -- the porus component. It may be somewhat elastomeric so, as you put it, in there it may be smaller than the outside diameter of the frame but once it's in there it could stretch out and just be a line to line contact.

BY MR. GABRIC:

Q. What would one of ordinary skill in the art think look at figure 18? Would it be an interfering fit or not?

MR. HAMILTON: Objection. Form.

THE WITNESS: I think a person skilled in the cart would look at this and say it doesn't make in he difference.

BY MR. GABRIC:

Q. Why not?
A. It doesn't change the function of the part. It doesn't alter what it is intended to do.
Q. And one skilled in the art looking at figure 18 would they understand there to be a friction fit between the porous component 81 and the frame 82?
MR. HAMILTON: Objection. Form.
THE WITNESS: Not necessarily. It could be line to line.
BY MR. GABRIC:
Q. What forces would one of ordinary skill in the art anticipate the device of the '548 patent would undergo in normal operation?
MR. HAMILTON: Objection. Form.
THE WITNESS: Are you speaking of just these components.
BY MR. GABRIC:
Q. Yeah, let's focus on figure 18. What forces of any would one of ordinary skill in the art expect these pieces to be subject to during normal operation of this device?
A. Gravity.
Q. Any other forces?
A. And a force acting on it.

Well, if you have an air stream passing through it the air would have some momentum, have some energy and as it midst the porus component with the wire that could result in some forces. Depending on how it's put together there may be some forces between 81 and 82 but not necessarily.

Q. What --

A. You can't tell by looking at it. If it's a line to line, there's no forces, if they're just touching.

Q. And if it's an interfering fit what forces would exist?

MR. HAMILTON: Objection. Form.

THE WITNESS: Well depending on where the interference fit is. For instance, another place where you could have an interference fit is the hole that the wire wound porus component passes through, that hole could be slightly smaller in diameter than the outside diameter of the porus component. However, the porus component, if it is somewhat elastic, once it's inside there, things are in equilibrium. It's all being held. So if you were to drop the device, it could be impact, I guess. If you were to go into space you could have no gravity on it. And then different. But --

BY MR. GABRIC:

Q. So one of ordinary skill in the art wanted to make the device or the components in figure 18 is there any reason why they would opt for an interfering fit versus a contact fit or vice versa?

A. I think it would really be up to them of what they wanted to do.

Q. Well, and what would drive those -- that decision? What factors?

A. Personal choice.

Q. What do you mean by that?

A. I chose to do it because I chose to do it.

Q. Well --

A. That's what I mean by that.

Q. Well, why would one choose an interfering fit over a contact fit?

A. Well, line to line would be maybe a little more difficult to maintain in terms of consistency of the parts. But like I said before, even if it wasn't an interference fit, it's not going to change the function of the device. It's going to work whether the porus component is slightly loose, line to line, or has an interference fit.

Q. Now, I think you said something to the effect that the portion of the porus component with wire on it is aligned with the through hole 821 correct?

MR. HAMILTON: You can answer.

THE WITNESS: The hole 821 is in the center of the frame.

BY MR. GABRIC:

Q. Right.

A. And the porus component with the wire is through the major diameter. And so those two are aligned. Yes. I think I said in my report, I'm not sure which one, that if you were to take a cylinder and pass it through hole 821, it would come in and basically contact the wire -- the wire wound component.

Q. And if I'm one skilled in the art and I want to make sure I maintain that alignment, would I opt for a interfering fit between the porus component and the holes in the cavity wall or a contact fit?

A. A contact fit? I think it could be others. Small changes there aren't going to make a huge difference.

Q. What if I didn't want any changes there? I absolutely don't want that thing moving relative to the through hole? What type of fit would I prefer?

MR. HAMILTON: Objection. Form.
THE WITNESS: Well line to line would do it.
BY MR. GABRIC:
Q. Would do it, as well as an interfering fit?
A. Yeah, an interference fit there's a potential that that could distort one of the parts though. And, actually, change it. If you have a hard and a soft part and I try to put them together one of the parts could actually deform. So I would put it together line to line. Like I said, small changes, little tiny dimensional differences aren't going to make probably even a measurable impact on the performance of it.
Q. So the deformation that you're referring to is deformation of the porus component?
A. Yes.
Q. And the deformation would take place where?
A. I'm sorry?
Q. Where would that deformation take place that you referred to?
A. Well, it depends on where the interference is. The eye section or the vertical section of the T, if it's larger diameter than the hole that it is passing through it would compress it a little bit.

Q. And would that be a problem?

A. No, what I said before is it doesn't make any difference if it does have a little bit of motion or not. It's not going to appreciably affect the overall performance of the device.

Q. And my question is, if I'm one skilled in the art and I don't want any relative motion of the porus component that's lined with the through hole would an interfering fit accomplish that better than a contact fit?

MR. HAMILTON: Objection. Form.

THE WITNESS: No.

BY MR. GABRIC:

Q. Why not?

A. A contact fit isn't going to allow any motion either. If it's a line to line, there's no place for it to go. So it's held in exactly the position that it's put in there.

Q. Well, what about if the porus component is subject deformation even if there's a contact fit can't it move when it compresses under force?

A. What --

Q. At the line to line fit?

A. That was a compound question. I don't understand what you were asking.

Q. So you're assuming that if you have a contact fit, you can have no movement of the porus component relative to the frame?

MR. HAMILTON: Objection. Form.

THE WITNESS: If it's a line to line which -- I'm not familiar -- the contact fit, I'm not familiar with using that term.

But if it is a line to line and there are -- if the inside diameter of the porus component is the same diameter as the outside diameter of the frame that it's mounted on, it has no place to go. And it's not going deform and it's not going to certainly change where the wire is relative to the hole.

BY MR. GABRIC:

Q. How does a line to line fit prevent relative movement of the porus component to the frame?

A. There's no place for it to go. If there's line it line there's no gaps, place, no spaces. It's already touching.

Q. Now, the portion of the porus component 81 that encompasses the internal portion of the frame, the portion of the frame that's internal to the porus component, is that portion held in place by the frame?

MR. HAMILTON: Objection. Form.

BY MR. GABRIC:

Q. I'll try to clarify.

A. I was going to ask you for that because --

Q. Yeah. We'll go back to your I analogy.

And this question is focused on the -- the -- the cross hatches of the I the cylindrical portion that encompasses the frame are you aware with me?

A. I believe so.

Q. All right. Is that portion of the porus component held in place by the frame?

A. Yes.

Q. How is that held in place by the frame?

A. Well, the frame has -- the internal diameters, the internal diameter of the porus component is contacting the external diameter of the frame and, again, there's to place for it to go so it's -- and the vertical section of the I passes through it and keeps it from moving in any direction for that matter.

Q. What keeps the frame from moving relative to the porus component?

MR. HAMILTON: Objection. Form.

THE WITNESS: I think the two if they do move would move together. Because they're connected. And if doesn't really show in the patent these -- we've been through this before. These are schematic representations or block diagrams. These are not engineering drawings. And so you can't interpret these as if everything that is shown is actually as it would be built.

BY MR. GABRIC:

Q. Does the porus component in 81 hold the frame in place, 82?

A. Not in figure 18, no.

Q. Why not?

A. Well, they're just floating in space in this picture. So, I mean, there's nothing attached. They are not sitting on anything, they're not connected to anything in figure 18.

Q. Does the porus component in 81 prevent the frame 82 from moving relative to the porus component?

MR. HAMILTON: Objection. Form.

THE WITNESS: I think the best answer to that is its the combination of the two designs that lock in all of the mating surfaces that keep either of the parts -- if one moves, the other moves.

BY MR. GABRIC:

Q. What do you mean by "it's the combination of the two designs that lock in all of the mating surfaces that keep either of the parts in one moves the other moves? I'm not sure what you mean by that?

MR. HAMILTON: Objection. Form.

THE WITNESS: Well, they're all locked together. The two parts are connected and the way they are connected, one part will not come apart from the other so, again, these are just floating in space in this particular illustration. So there's really, no force, I guess gravity would exist but we don't know what orientation it is. They're just two parts that are connected. You move one, you move the other.

BY MR. GABRIC:

Q. So is the frame 82 hold the porus component in place? What do you mean holds it in place? In place relative to the frame?

A. I think what I mean and I'll say this again, try to get some clarity, the two parts are locked together by the design. The porus component has a section that goes through holes and keeps it from moving axially. It's on the diameter. The inside diameter of the porus component mates with the outside diameter and so there's no -- no relative motion between these two parts.

Q. And so with respect to relative movement of these two parts, is it fair to say that they each hold each other in place, they cooperate?

MR. HAMILTON: Objection. Form.

THE WITNESS: No, I wouldn't say that.

BY MR. GABRIC:

Q. Why not?

A. Because the purpose of the porus component is to effect aerosolization. The purpose of the frame is what I described at the beginning of the deposition. And they have different functions.

Q. But for all practical matters isn't the porus component holding the frame in place with respect to relevant movement of those two components?

MR. HAMILTON: Objection. Form.

THE WITNESS: The porus component is locked into the frame. And it is not intended to support or hold anything. That's not it's function.

BY MR. GABRIC:

Q. What's the basis for that? Where does the patent say that's not it's function?

A. Well, back to the patent I guess huh? The frame has a function of creating atomization cavity of holding the porus component relative to the frame for the alignment. The porus component is not a structural component that is intended to provide a structure or a way of holding things.

Q. So what -- what portion of this porus component is set on the frame in Figure 18? Is it the entire porus component? Portions of it?

A. The entire porus component is set upon the frame. Yes.

Q. So even the porus component that's in the internal of the frame, that portion of the porus component is set on the frame in your opinion?

A. The internal portion? What was that?

Q. The vertical, the I we've been talking about the I the vertical portion that's internal to the frame 82, in your opinion, that portion of the porus component is set on the frame?

MR. HAMILTON: Objection. Form.

THE WITNESS: I wouldn't take that vertical section and start parsing it into little pieces and say is this set on the frame or is this not set on the frame. It spans the gap there. The purpose of the frame is to ensure that that space is there.

That little portion where the wire wound around the porus component is not touching the frame. But the entire porus component is set upon the frame to keep it in a particular location such that it will function as intended.

BY MR. GABRIC:

Q. So just so I understand your opinion, so what portion of the porus component is set on the frame?

MR. HAMILTON: Objection. Form.

THE WITNESS: I think you can't distinguish part of it that's set and part that's not. The entire porus component is set on the frame. And it doesn't mean that it needs to touch it and the fact that that second -- that center section is not touching the frame doesn't mean that it's not set on the frame. I think you're confusing "set on" with "sitting on."

BY MR. GABRIC:

Q. Now, you see the portion of the porus component with the wire wrapped on it in figure 18, the heating wire wrapped around the portion of the porus component?

A. You said 13?

Q. I'm sorry, I meant to say 18. Figure 18.

A. Okay. Yeah.
Q. I'm on figure 18 right now?
A. Okay.
Q. You see the portion of the wire wrapped around the porus component there, right?
A. Yes, the wire wrapped around the vertical section of the I as we've referred to it.
Q. All right. And I take it that wire does something to the liquid that's in that porus component?
MR. HAMILTON: That's it? Objection. Form.
MR. GABRIC: We're on the edges of our seat.
MR. HAMILTON: Yes, yes.
THE WITNESS: The heating wire atomizes the liquid and produces the atomization particles.
BY MR. GABRIC:
Q. What do you mean by "atomizes the liquid?
A. Goes through a phase change.
Q. What is the phase change that occurs?
A. It goes from a liquid to a vapor or an aerosol.
Q. Do you use vapor and aerosol interchangeably?

A. I believe in these proceedings it has been used interchangeably.

Q. And what in fact happened? If I understand you correctly the heat from the wire will vaporize the liquid that's in the porus component?

A. It vaporizes the liquid that's touching the wire. Whether it vaporizes all of the liquid or not would be a function of how long the wire is on and if the porus component is going to continually replace the liquid because it is -- like I said before, there are un-shown portions here and it is in contact with the liquid supply.

Q. And what happens after the liquid is vaporized? Does the vapor coalesce to form droplets? What happens?

MR. HAMILTON: Objection. Form.

THE WITNESS: They could as they cool. I don't know if coalesce is the right word. But --

BY MR. GABRIC:

Q. Probably isn't?

A. What happens afterwards you produce this, there's a dynamic situation. There's air flow through here so the small particles, the droplets, the aerosol particles are swept away in stream and they would exit through the mouthpiece B 1 eventually.

Q. Now, you're familiar with the Whitmore reference?

A. I am.

Q. And that's the wire wrapped wick right?

MR. HAMILTON: Objection. Form.

BY MR. GABRIC:

Q. A wire wrapped wick?

A. I'm familiar with Whittemore very much.

Q. I'll give it to you. I don't have to -- show you what's been marked as Exhibit 1004 in these proceedings.

MR. HAMILTON: Thank you.

MR. GABRIC: You're welcome.

BY MR. GABRIC:

Q. And my question is, with respect to Whittemore's wire wrapped wick, does that configuration atomize liquid any differently than the wire wrapped porus component of figure 18?

MR. HAMILTON: Objection. Form.

THE WITNESS: Does it atomize differently?

BY MR. GABRIC:

Q. Right. Does it function differently than the wire wrapped porus component of figure 18?

THE WITNESS: Well, they mouth have a

porous component that contains liquid that's being provided through capillary action and the wire is wrapped around and intermittently touches it so there is contact between and so they're very similar yes.

BY MR. GABRIC:

Q. Now in the May 2006 time frame, in one of ordinary skill in the art have understood that the wick, for example, disclosed in Whitmore could be made from fiber material?

A. I would think so, yes.

Q. Now, I just have a couple more questions and we can take a break. We've been going almost an hour. Are you okay? Do you want to take a break now are you all right?

A. I'm terrible.

Q. You seem fine?

A. This is grueling.

Q. Isn't it?

A. Oh, terrible. I'm fine.

Q. Am I going too easy on you today? You're disappointed?

A. You know, whatever you do is fine with me.

Q. All right.

A. Because it doesn't matter.

Q. That's the attitude.
A. You've got to live with yourself that's all I know.
MR. GABRIC: Nothing's gonna happen don't worry.
THE WITNESS: So we don't need Whitmore.
BY MR. GABRIC:
Q. I'm probably done with Whitmore.
A. Okay.
Q. Show you what's Exhibit 2030 in this proceeding. This is one of the declarations you provided.
A. Okay.
Q. And I have a question about paragraph 12. Give you a chance to read it. Its short?
MR. HAMILTON: Do you want to clarify what this is.
MR. GABRIC: I thought I said. Did I not say.
MR. HAMILTON: You said one of his declarations.
MR. GABRIC: I'm sorry. This is Exhibit 2030. It's one of the declarations submitted in this matter. It's dated May 25, 2017.
BY MR. GABRIC:

Q. Okay?
A. Yes, sir.
Q. All right. You read paragraph 12?
A. I have it queued up.
Q. Fair enough.
And correct me if I'm wrong, but when I read paragraph 12, you seem to be saying that this wire wrapped porus component disclosed in figure 18 of the '548 patent plays some role in improving the aerosol effects and/or the atomizing efficiency.
Is that what you're trying to convey in paragraph 12?
MR. HAMILTON: Objection. Form.
THE WITNESS: I'm not quite sure.
BY MR. GABRIC:
Q. Let me ask it a different way.
In paragraph 12 you're talking about the atomizer -- the '548 patent's atomizer.
Do you see that?
A. Yes, I do.
Q. All right. And you talk about some of the design elements of the atomizer of the '548 patent.
Do you see that?
A. I do.
Q. And -- and the once you calling out are the

one through hole which I'm not going to ask about right now, and the other one is the heating wire wound on the porus component in the air flow path.

Do you see that?

A. I do.

Q. What, if anything -- what, if any, role does the heating wire wound on the porus component play in improving the aerosol effects or atomizing efficiency of the atomizer's Figure 548 patent?

MR. HAMILTON: Objection. Form.

THE WITNESS: I'm comparing and contrasting this to the design in the '043 Hon patent.

BY MR. GABRIC:

Q. Okay.

A. Okay? And I believe that this design is a -- can be an improvement. I mean, any design could go backwards if it's not properly executed. So you may think it's an improvement but in fact it may not be. Depending on how you actually as I said, execute the design. The devil is in the details so what I was saying is that these elements help to improve the aerosol efficiency. Because there's direct contact. Because there's a good feed of liquid to the wire. It's a consistent feed through capillary action. You were just asking about the wire and the wick right? So you have -- you have a continuous supply of liquid to the heating element which causes the vaporization on power supplied.

MR. GABRIC: We can take a break.

THE VIDEOGRAPHER: Off the record. The time is 10:08.

(Recess taken.)

THE VIDEOGRAPHER: We're back on the record at 10:23.

BY MR. GABRIC:

Q. Welcome back Mr. Meyst. Did you discuss your testimony at all during the break?

A. Did not.

Q. I didn't think so. Thank you.

Let me show you what's been mark as Exhibit 1003 in this proceeding. It's the Hon '043 reference. I take it that looks familiar to you?

A. I've seen it once before.

Q. For the record, you're being somewhat facetious.

A. Many times.

Q. Now, turn to Figure 6 of Hon '043.

A. Okay.

Q. And focus your attention to cavity wall 25.

Do you see that?
A. I do.
Q. Okay. One of ordinary skill in the art reading this reference in the 2006 time frame, would they understand whether the cavity wall could be made from both air permeable or air impermeable material?
A. I believe that's the case, yes.
Q. Now Hon --
A. Excuse me.
Q. Yeah?
A. Maybe I should go back and just check on what they say. I know they mentioned materials but I don't know if it identifies whether they're air permeable or not.
Q. Yeah, let me see if I can help you. I think they disclose materials. I'm not sure they say anything about air permeability. Let's see if we can find it. I know it's least on page 9, about six lines from the bottom. I don't know if this is what you're looking for. But it says the atomization cavity wall could be made of aluminum oxide or ceramic.
Do you see that?
A. I remember that. I'm looking for it. Yes,

there it is.
Q. Got it?
A. Right.
Q. Okay. And my question, let me reask it would one skilled in the art in the 2006 time frame understand that aluminum oxide or ceramic they could be air permeable or air impermeable?
MR. HAMILTON: Objection. Form.
THE WITNESS: The patent doesn't say that but those two materials could have wide-ranging properties.
BY MR. GABRIC:
Q. Okay. Could they be air impermeable?
A. Impermeable.
Q. Impermeable, yes.
A. I'm not sure. I think they could be, though.
Q. Now, back to Figure 6. I just have a few questions.
You see there's ejection holes 24 referenced in Figure 6?
A. Yes, I do.
Q. And Hon never calls these ejection holes 24 atomizers does he?
MR. HAMILTON: Objection. Form.

THE WITNESS: What he says is that these holes are high speed ejection streams that eject droplets into the atomization cavity and converting a liquid to a droplet is atomization process. The patent doesn't call it saying but one skilled in the art would look at that and know that that is in fact an atomizer.

BY MR. GABRIC:

Q. And Hon never discloses an embodiment that doesn't include at least the piece oh electrical element 23, the heating element 26, or the second piezo electric element?

MR. HAMILTON: Objection. Form.

MR. GABRIC: Let me finish. 35. Now, you may object.

MR. HAMILTON: Objection. Form.

THE WITNESS: Okay. I was caught in the middle what's the question again.

BY MR. GABRIC:

Q. Yeah Hon never discloses an embodiment and the atomizer omits or -- strike that.

Hon doesn't disclose an embodiment where there isn't at least one of heating element 26 piezo electric element 23 or piezo electric element 35, correct?

A. I don't believe there's an embodiment shown but that as we've discussed previously doesn't mean that it can't be.

Q. And -- and I think -- now, the -- the -- the heating element 26, that requires electrical power, right?

A. Yeah.

Q. Hon discloses a battery to provide that power, correct?

A. It's a heating wire so, yes, it heats when electricity is put through it.

Q. Okay. And the piezo electric elements 23 and 35, they also require electric power, correct?

A. Yes.

Q. All right. Does any other component in Hon '043 require electric power?

A. Well, yes.

Q. What else?

A. I'd have to go back and look but I believe it does have a light indicator. There's a circuit board in here. There's a sensor. There are other electronic components.

Q. Now, Hon '043, he discloses -- and I'm at the top page 11, the first full sentence after the atomization, the large diameter droplet stick to the wall under the action of eddy flow and are reabsorbed by the porous body 27 et cetera, et cetera it goes on.

I want to talk about the clause the large diameter droplet stick to the wall under the action of eddy flow.

Do you see that?

A. I do.

Q. What would one of ordinary skill in the art understand Hon to be teaching here? How do these droplets stick to the wall under the action of eddy flow?

MR. HAMILTON: Objection. Form.

THE WITNESS: I don't believe that the patent goes into any additional detail, other than to describe -- as shown in figure 7, I believe, 29 are the overflow holes, let me see what they call those.

What was the citation you were just looking at the top of page 11.

BY MR. GABRIC:

Q. It's in the third line, bridging second and 13 line the large diameter droplet stick to the wall under the action of eddy flow.

Do you see that?

A. I do.
Q. I'm curious of how one of ordinary skill in the art would interpret that statement back in 2006? What is Hon teaching here?
A. Well, I don't think there's a lot of interpretation required. The large droplets will hit the inside surface of the atomization chamber and they are provided exit holes 29 which communicate directly through the atomizer to item number 27 which is the porous body. So the purpose of that is to eliminate I think they call it wet mouth. You want to get rid of the large droplets have those removed from the air stream so they don't go the patient's -- or the user's mouth.
Q. What is eddy flow?
A. It's flow that -- eddys are circulating flows and flows that move around somewhat randomly.
Q. And what is causing eddy flow in the atomizer of Figure 6?
A. The air flow through the device.
Q. Why is the air flow through the device have eddy flows?
A. Well, because it's a complex structure and as you put air through it will move depending on what the restrictions are in the confinement space.

Q. Now, the heating element 26 does that play any role in causing eddy flows to occur?
A. I don't believe it discusses it in here, but I would say, yes. It interrupts the flow. It stops the laminar direct flow in and causes it to mix up and be a very complex air pattern.
Q. What role does the eddy flow play in causing the large droplets to stick to the cavity wall?
A. Gets them in proximity to the wall so they hit it.
Q. What else in Figure 6 would cause eddy flows other than the heating element 26?
A. What else is causing what?
Q. Eddy flows, other than heating element 26?
MR. HAMILTON: Objection. Form.
THE WITNESS: Well, we have the material coming in, the droplets are coming in through the high speed ejection holes 24. And the heating element could be a plate or a sheet. It doesn't have to be a wire as shown here. So the inside of the atomizer would -- could have physical components, elements, that are not porous, that cause the air to go around them or pass them or abut them and bounce off. So there's a whole host of

things in here that could do that.

BY MR. GABRIC:

Q. Now, it's my understanding that it's your opinion that the porous body 27, and I'm referring to Figure 6, holds the cavity wall 25 in place, right?

A. Well, the wall 25 is contained within so it is totally encapsulated inside the -- the porous body, yes.

Q. So does the porous body hold the cavity wall in place?

MR. HAMILTON: Objection. Form.

THE WITNESS: Yes, it does. It holds it.

BY MR. GABRIC:

Q. How does it do that?

A. How does it hold --

Q. How does it hold it in place?

A. It completely encapsulates it. It surrounds it on all surfaces. So the item 25 doesn't touch anything else. The only thing it's touching is the porous body.

Q. And can the porous body -- I'm sorry, can the cavity wall move relative to the porous body?

A. Not on its own, no.

Q. What do you mean by "not on its own?

A. Well, it doesn't have any mode of force?
Q. In the normal use of this device, Hon '043, can the cavity wall move relative to the porous body?
A. Well, in an absolute measure maybe angstroms it's possible but the way it's shown here it is totally encapsulated so it has no place to go.
Q. Now, the -- there's an -- a piezo electric element 23.
Do you see that in Figure 6?
A. Yes.
Q. What's a piezo electric element?
A. It's an electro mechanical device that when power is put to it, it will vibrate.
Q. And does the piezo electric element 23 cause this atomizer in Figure 6 to vibrate?
A. I'm sorry.
Q. Does it cause the atomizer to vibrate when its turned on?
A. It could cause item 27 the porous body to vibrate some. It depends on how powerful the piezo electric device is and what the constituents are if the properties of the porous body.
Q. Now, one skilled in the art looking at this, would they have an understanding of whether

the cavity wall would rattle at all inside the porous body when the piezo electric element is turned on?

A. I've never heard the word "rattle other than in Dr. Sturges, I think, his report and now, you just used it. But no these parts don't rattle.

Q. Why not?

A. Well, I don't believe they're materials that if they were next to each other would emit any noise. I don't think that's the kind of properties they are. But it's not spoken of in the patent and I haven't really researched it. So I guess it's possible they could be used as a rattle. But I don't know.

Q. Okay. Maybe I shouldn't use rattle because I think we use the term differently.

When the ultrasonic -- I'm sorry, the piezo electric element 23 is turned on in the porous body is vibrating, would there be any movement relative -- would there be any movement between the cavity wall and the porous body? Would those vibrations just transmit into the cavity wall?

A. It's not talked about in the patent. But there will be an attenuation of vibration the further you go away. Does it move relative to the other? I don't think that it would cause them to be moving. There's no place for item number 25 to move to. It's held in place top, bottom and all the way around on the outside.

Q. And what kind of fit is there between the cavity wall and the porous body?

A. Well, here it's shown as line to line. But these are schematic representations. So we don't know what kind of fit is actually there.

Q. Now, does the porous -- I'm sorry. Does the cavity wall play any role in preventing the porous body from moving relative to the cavity wall?

A. Say that again. Does the --

Q. Does the cavity wall play any role in preventing movement of the porous body relative to the cavity wall?

A. Well, you have a net fit so the two parts work together in cooperation to form one part which doesn't allow for any movement.

Q. Now, looking at Figure 6, there's a -- a heating wire illustrated in row 26, right?

A. Yes.

Q. Okay. And Hon '043 discloses that you could also place a heating sheet there in place of a wire?

A. I think that's the word they use, but I'm not sure.

Q. I think it's on page 9. It's kind of in the middle where it starts talking about the heating element 26.

A. Into a sheet form.

Q. Yeah. Is that what you're looking for?

A. Yes.

Q. Okay. How does Hon '043 disclose anything about the dimensions of this sheet form?

A. It has some dimensions about holes but I don't believe there's any dimensions on the heater.

Q. One skilled in the art in 2006 reading Hon '043 would they have any understanding of what Hon means by a sheet form?

A. I believe so, yes.

Q. What would their understanding be?

A. Well, a device that is more two-dimensional, has a width and a length and some thickness but it's not I aware.

Q. Would it have a greater surface area than the wire?

A. Well, you can't compare them directly because you could make them equal. Depending on how the -- how the dimensions are.

Q. Let me ask a better question.
What understanding if any would one of ordinary skill in the art have as to why Hon '043 is proposing a sheet form as an alternative to a wire?
A. Well, contact area could be one thing. But also it would impact the air flow and the -- the method of mixing air with the aerosol inside the chamber. With the droplets that are ejected in from the high speed holes.
Q. You said contact area could be one be thing. What do you mean by that?
A. The surface area available for transmitting heat so the liquid.
Q. So increasing the surface area?
A. I think that's what -- the objective would be to optimize the surface area to generate the -- an adequate amount of aerosol.
Q. And why -- why optimize the surface area? How does that help you provide an adequate amount of aerosol?
A. When you optimize anything you get all of the various parameters adjusted in proportion such that the thing is efficient and functions the way you want it to and that could include energy consumption it could include the draw, how much

resistance there is to air flow. There could be a whole bunch of things that together need to be optimized to end up with a result that you're looking for.

Q. For example, one of ordinary skill in the art reading Hon would have understood in the 2006 time frame that you could optimize the sheet to have a greater surface area so more droplets hit it, but there are other factors you need to consider, for example, it's impact on air flow through the device is that what you're talking about?

A. They all fit together.

Q. And this would have been within the capabilities of one of ordinary skill in the art in 2006?

A. I believe so, yes.

Q. So it's fair to say one of ordinary skill in the art looking at Hon Figure 6 when it comes to the heating element they would have understood that I could increase the surface area to increase the number of droplets that strike it put there are potentially trade offices such as it's affect on air flow and disrupting air flow through the atomization capital?

A. It's all interconnected.

Q. And one of ordinary skill in the art would have understood that, right?
A. I believe so, yes.
Q. Now, going to paragraph 78, 79 of your declaration, Exhibit 2030?
A. I'm sorry, paragraph 79?
Q. 78 and 79. Just let me know when you're done reviewing it?
A. I'm fine.
Q. Okay. I really want to focus on the -- the -- the curves or graphs that you have at the top of page 37.
Could you explain to me what this graph is showing us?
A. It's a stress strain graph for a number of different materials.
Q. Okay.
A. These are all look to be nickel chromium I think but I'm not sure.
Q. Its like -- it looks like there's one for pure nickel and then there's materials with nickel and chromium, various amounts of chromium, I believe. Is that look right to you?
A. I believe that's the case, yes.
Q. All right. And what is a stress strain

curve?

A. It's a plot of the applied force or stress against the strain which is deformation or movement.

Q. All right. Let's -- let's focus on stress for a minute.

What kind of stress was applied to these materials in making these curves?

A. What kind of stress?

Q. Yeah. What physically what -- what was done to these materials? Where was the stress allied in generating these curves?

A. I'd have to go back and look at the -- the source. But the whole point of these was to show that at very -- at the stress levels that are equated with it, that there's very little strain.

Q. I'm going to show you Exhibit 2035. Which I believe is were the stress strain curves are from and review what you need. I'd like to get an understanding of exactly what kind of stress was applied to these materials, was compressive stress, was it some other type of stress? That's what I'm trying to get at.

A. Okay. Okay. In paragraph two point 3 on page 2 of 10 of this item, it describes the testing.

Q. Okay.

A. That was used to -- that was a compression test. Mechanical testing, I'll read it, was conducted under compression loading on a servohydraulic testing machine with a cross head speed of 0.1 millimeter per second using paraliptide (^CKSP) specimens cut to 10 millimeters by 10 millimeters by 28 millimeters with electro discharge machining.

Q. And where on the material was the compressive stress applied?

A. Well, I don't know if it tells in here. Where on the material? The specimen was put in -- if you're doing a compression test, you put it between a platen and a load cell and compress it at the cross head speed that was mentioned there and you record the stress and strain.

Q. So do you know if the compressive stress was applied in a -- a vertical versus axial direction on the piece, vertical versus horizontal? Do you see what I'm saying? It could be applied this way or this way?

A. You can apply stress in different directions and you may end up getting different results?

Q. Does this paper tell you in which direction the compressive stress was applied?

A. It's been a while since I've read this reference. I have to read the difference.

I think it references the equipment in here but I don't know if it -- it has the details about the -- the way the experiment was run.

Q. So we don't know if a longitudinal stress verse compressive stress was applied to this material?

A. I don't know if it states it in here or not. It may state it but I don't recall off --

Q. All right?

A. At the moment.

Q. All right. If its in there it's in there and if it's not it's not I suppose.

A. Yes.

Q. All right. So let's focus on the X axis on top of your -- on top of 37 of your declaration.

A. Okay.

Q. Okay? Now, there's numbers there on the X axis. What are those numbers representing?

A. Are you talking about the 0 point 511.5.

Q. The 0.1, 0.2, 0.3, 0.4, 0.5, 0.6 on the horizontal axis.

Do you see that?

A. Okay, yeah.
Q. What are those numbers depicting?
A. That's strain.
Q. Right. But what are those numbers referring to?
A. Millimeters per millimeters. So you look at your original sample and you look at the deformation and how much deformation over the original size.
Q. Well, are you using -- well, let me back up.
So these numbers on the X axis are measuring the percentage at which the material is actually compressed and responds to these -- the applied stress?
A. It's millimeters per millimeter. I don't recall if you call that a percentage. But it's -- it's a ratio of the starting dimension to the resulting dimension.
Q. Okay.
A. As it is compressed. It's going to get smaller as you push on it.
Q. It's going to occupy more -- less space as you push on it, the material?
A. It depends. It may occupy less space.

Q. Your compressing it, it gets smaller it's occupying less space?
A. It may bulge out at the side I don't know.
Q. I see okay. Now, is this graph measuring how much of material -- these materials will bend in response to an applied stress?
A. This is a straight compressive load.
Q. That's a different measurement than measuring how much a material may bend in response to a stress?
A. Yes, it is.
Q. And -- strike that.
And on top of page 38 you have another stress strain curve.
Do you see that?
A. Yes.
Q. And what is this curve measuring?
A. Well, paragraph 80 says it's a similar graph for stainless steel fibers taken from this particular report. Exhibit 2036. So this is also a stress strain curve.
Q. And is this curve showing how much these materials would bend in response to stress?
A. Why don't you pull out 2036 so we'll look at it.

Q. Sure.

A. Okay are you ready? What's your question?

Q. Yeah. I was referring to the curve on top -- the curves on top of page 38 and I asked you whether these curves show how much these materials would bend in response to stress.

A. This is according to the document here on page 6 of Exhibit 2036. Figure 5, are compression properties and energy absorption efficiency so this is a compression again and they did it in two directions, longitudinally and transverse to show that the properties are somewhat different depending on which direction the load is applied so to answer your question this isn't bending.

Q. Now, in paragraph 82, you used the term "tensile strength."

Do you see that?

A. Yes.

Q. What's tensile strength?

A. It's a measure of mechanical strength.

Q. How is this strength measured?

A. Typically it's in tension. It's opposed to compression.

Q. And tension is what pulling on the material rather than compressing it?

A. Yes, yes.
Q. Okay. And is it the amount of pulling stress that applies to a material before the material breaks? Is that the measure of tensile strength?
A. That would be yield strength. A tensile curve is just the relationship between applied load and deformation.
Q. Okay. So you apply more load, the material stretches and stretches and stretches, I take it, and at some point it breaks?
A. It could or it could become elastic and just stretch out.
Q. Okay. So let's see if I understand. So you -- you apply a pulling stress, and at some point a material I suppose if you take away that stress will revert back to its original form? Is that a possibility?
A. Not necessarily. It depended on what it's an elastic material or not.
Q. Okay. So an elastic material may go back to its original shape once that stress is removed?
A. It may go, it may go part way. Depends on if you go past the yield point where you get into plastic deformation and get out of elastic

deformation, no, it won't recover.

Q. Okay. So elastic deformation is the situation where the material will recover?

A. To some degree.

Q. Okay. And plastic deformation basically refers to irreversible deformation?

A. Yeah, I'd have to go back up the definitions but in general that's correct.

Q. Now, you refer to paragraph 82 a materials data book, Exhibit 1026. I'll give that to you. I'm going to show you what's been marked as Exhibit 1026 in this IPR. And it's the supplemental declaration of Dr. Sturges. And in paragraph 82, you cite to it looks like page 59 of this supplemental declaration. And you -- I'll give you a chance to locate it. But you cite to page 59 in support of your statement that foam ceramics have a tensile strength similar to that of metal."

Do you see that?

A. Yes.

Q. All right. What are you -- where on page 59 are the tensile strength of foam ceramics disclosed?

A. Well, it's on the left-hand side, hidden under the heading "Ceramics."

Q. Okay which one of these are foam ceramics?

A. Alumina could be. There's a variety of these but the whole point of this was to say that these materials have structural integrity and that they can be arranged similar to having properties similar to metals.

Q. Right. But how do we know these materials listed in the ceramic 6 are foamed?

THE WITNESS: I guess to answer your question I would have to do some research to look at the ceramics that are listed and see which of those could be made in a foam version.

BY MR. GABRIC:

Q. So we don't know if the ceramics listed on page 59 here that you point to for the -- the tensile strength of foam Sturges are, in fact, foam ceramics; isn't that true?

A. First of all, let me correct something. It says 2 here. Under the ceramics it does say for ceramics yield stress is replaced by compression I have is strength which is more relevant in ceramic design so those are all compressive tests similar to the one that was performed for the other graph that I cited for the -- that was a compressive load.

I'll notice that each of the items under ceramics has an asterisk.

Q. Right.

A. And that asterisk refers to the note at the bottom on the page 59 on the right side.

Q. Right. There's two columns here right? There's a column that's labeled yield stress which is footnote tells you --

A. Labeled what?

Q. It's -- the first column is a yield stress but the footnote tells you that the ceramics it's not yield stress it's compressive strength, right?

A. Yes.

Q. But the second column is tensile strength, right?

A. I believe it is. Yes.

Q. Right. And in your declaration, you are pointing to the numbers for tensile strength in support of this statement that foam ceramics have a tensile strength similar to that of metals, correct?

A. That's what it says in my declaration. Yes.

Q. But we don't know if these are foam ceramics, correct?

A. They may be. And I'm not sure. I'd have to go back and look. But I brought no notes with me

and none of my research.
     Q.   There's nothing on this page indicating that these are foam ceramics?
     A.   I don't know that to be the case. Because I'm not an expert in ceramics but if we were to look at these we may find that some of them are made by the foaming process. So I don't think you could say at that none of them are foam.
     Q.   Now, Figure 6 of -- I'm done with that.
          Go back to Hon '043, Figure 6.
     A.   Okay.
     Q.   Now, the porus body 27, the -- one of the functions of this porous body is to re absorb liquid, right?
          MR. HAMILTON: Hold on a second let me get to where we are.
          MR. GABRIC: We'll wait.
          MR. HAMILTON: Objection. Form.
          THE WITNESS: And your question was one of the functions of porous body is to re absorb.
BY MR. GABRIC:
     Q.   I understand the objection now. Let me ask a better question.
          One of the functions of this porous body 27 is to reabsorb large liquid droplets that flow

through the overflow holes, right?
A. That's what the patent says, yes.
Q. Right. Now, there's no heating wire on this portion of the porous body is there?
A. What -- I don't think it's permissible to use your finger to circle something.
Q. I'm sorry, there's --
A. I have no idea what you're talking about.
Q. I'm looking at the porous body 27 and there's no heating wire on the porous body 27, right?
A. There's no heating wire on it? No, there's a heating wire inside the cavity.
Q. Is that -- is this porous body heated by that heating wire?
A. It might warm up a little bit if you leave -- but the way this is used it's on for such a short period of time, I don't think you would notice any temperature increase.
Q. Are you familiar with the terms through hole and blind hole.
A. Yes.
Q. Would one of ordinary skill in the art be familiar with those terms in March of 2006?
A. I believe so. Yes.

Q. What would one of ordinary skill in the art have understood the term "through hole to mean?
MR. HAMILTON: Objection. Form.
THE WITNESS: That's a hole through a material which goes through both sides.
BY MR. GABRIC:
Q. And --
A. So --
Q. I'm sorry?
A. It -- you can look through the hole and see daylight.
Q. And what would one of ordinary skill in the art in March of 2006 have understood the term "cavity -- I'm sorry. Blind hole to mean?
MR. HAMILTON: Objection. Form.
THE WITNESS: A blind hole does not penetrate the second surface so it is a -- a hole that has a measurable depth. You cannot see light through it. Well, unless it was a clear part.
BY MR. GABRIC:
Q. Understood.
Why don't we take a short break. I'm probably pretty close.
MR. HAMILTON: Okay.
THE VIDEOGRAPHER: Off the record. The

time is 11:15.

(Recess taken.)

THE VIDEOGRAPHER: We're back on the record at 11:32.

MR. GABRIC: Mr. Meyst thank you for your time I have nothing further.

MR. HAMILTON: Okay. Why don't we take another break let's see if I have some more questions. At the break we can talk quickly if we want to do lunch for this break or if I can do it beforehand. Let me talk to Mr. Meyst and see what he wants to do and hopefully we can come back and finish the depo.

MR. GABRIC: Okay.

THE VIDEOGRAPHER: Off the record at 11:33.

(Recess taken.)

THE VIDEOGRAPHER: We're back on the record at 12:28.

BY MR. HAMILTON:

Q. Mr. Meyst do you recall this morning discussing with Mr. Gabric the differences between the stress caused by compressive forces and tensile forces?

A. Yes.

Q. And we'll point you to Exhibit 2030. This is your report in support of patent owners opposition in this matter.

2030. Do you have that?

A. There it is. Yes.

Q. If you could turn to pages 37 and 38 which are the pages of the actual report, not the pages of the exhibit. In the exhibit it's 38 and 39?

A. Right.

Q. And you see the two figures you were discussing this morning?

A. Yes.

Q. Regarding stress?

A. Yes, I do.

Q. And did you testify this morning that the -- the force applied in these two charts related to compressive stress?

MR. GABRIC: Object to form.

THE WITNESS: I believe I did, yes.

BY MR. HAMILTON:

Q. Do you believe that's the correct force to apply in your opinions in this case?

MR. GABRIC: Object to form.

THE WITNESS: Based on what they were discussing, which was putting a high pressure into the device, it would have been a compressive force.

BY MR. HAMILTON:

Q. Why would it have been a compressive force?

A. Because.

MR. GABRIC: Object to form. Apologize.

THE WITNESS: Because you're increasing -- go back to the question, why is it a compressive force?

BY MR. HAMILTON:

Q. Yes.

A. The air pressure inside would be pushing down in a compressive manner. The internal pressure was heightened so there would be forces compressing the elements inside.

MR. HAMILTON: All right I have no further questions. Thank you for your time.

MR. GABRIC: I have just a couple. We don't have to take a break.

FURTHER EXAMINATION

BY MR. GABRIC:

Q. Did you discuss this the substance of your testimony with counsel?

A. Did not.

Q. I have nothing -- did you meet with counsel about the substance of your testimony during the break we took before you just gave your redirect testimony?

A. I haven't had any discussions about my testimony today.

Q. Good enough. Thank you.

A. Okay.

THE VIDEOGRAPHER: This will then mark the end of Volume I in the recorded deposition of Richard Meyst. The original recording is retained at DTI at their offices in Chicago, Illinois. Going off the record. The time is 12:32.

Exhibit E

NINTH EDITION

Marks' STANDARD HANDBOOK for MECHANICAL ENGINEERS

Eugene A. Avallone
Theodore Baumeister III

Marks'

Standard Handbook for Mechanical Engineers

Revised by a staff of specialists

EUGENE A. AVALLONE *Editor*

Consulting Engineer; Professor Emeritus of Mechanical Engineering,
The City College of the City University of New York

THEODORE BAUMEISTER III *Editor*

Retired Consultant, Information Systems Department,
E. I. du Pont de Nemours & Co.

Ninth Edition

McGRAW-HILL BOOK COMPANY

New York St. Louis San Francisco Auckland
Bogotá Hamburg London Madrid Mexico
Milan Montreal New Delhi Panama
Paris São Paulo Singapore
Sydney Tokyo Toronto

Library of Congress Cataloged The First Issue
of this title as follows:

Standard handbook for mechanical engineers. 1st–ed.;
1916–
New York, McGraw-Hill.
v. Illus. 18–24 cm.
Title varies: 1916–58: Mechanical engineers' handbook.
Editors: 1916–51, L. S. Marks.—1958– T. Baumeister.
Includes bibliographies.
1. Mechanical engineering—Handbooks, manuals, etc. I. Marks, Lionel Simeon, 1871– ed. II. Baumeister, Theodore, 1897– ed. III. Title: Mechanical engineers' handbook.
TJ151.S82 502'.4'621 16–12915

ISBN 0-07-004127-X
Library of Congress Catalog Card Number: 87-641192

MARKS' STANDARD HANDBOOK FOR MECHANICAL ENGINEERS

234567890 DOC/DOC 89321098

ISBN 0-07-004127-X

First Edition *Eleven Printings*	Third Edition *Seven Printings*	Fifth Edition *Seven Printings*	Seventh Edition *Fifteen Printings*
Second Edition *Seven Printings*	Fourth Edition *Thirteen Printings*	Sixth Edition *Eight Printings*	Eighth Edition *Eleven Printings*

The editors for this book were Betty Sun and David E. Fogarty
and the production supervisor was Teresa F. Leaden.
It was set in Times Roman by University Graphics, Inc.

Printed and bound by R. R. Donnelley & Sons Company.

The editors and the publishers will be grateful to readers who notify them of any inaccuracy or important omission in this book

Section 15

Electrical and Electronics Engineering

BY

C. JAMES ERICKSON *Principal Consultant, Engineering Department, E. I. du Pont de Nemours & Co.*

BYRON M. JONES *Consulting Engineer, Assistant Professor of Electrical Engineering, University of Wisconsin—Platteville.*

15.1 ELECTRICAL ENGINEERING
by C. James Erickson

15.2 ELECTRONICS
by Byron M. Jones

15.1 ELECTRICAL ENGINEERING

by C. James Erickson

REFERENCES: Knowlton, "Standard Handbook for Electrical Engineers," McGraw-Hill. Pender and Del Mar, "Electrical Engineers' Handbook," Wiley. Dawes, "Course in Electrical Engineering," Vols. I and II, McGraw-Hill. Gray, "Principles and Practice of Electrical Engineering," McGraw-Hill. Laws, "Electrical Measurements," McGraw-Hill. Karapetoff-Dennison, "Experimental Electrical Engineering and Manual for Electrical Testing," Wiley. Langsdorf, "Principles of Direct-current Machines," McGraw-Hill. Hehre and Harness, "Electric Circuits and Machinery," Vols. I and II, Wiley. Timbie-Higbie, "Alternating Current Electricity and Its Application to Industry," Wiley. Lawrence, "Principles of Alternating-current Machinery," McGraw-Hill. Puchstein and Lloyd, "Alternating-current Machinery," Wiley. Lovell, "Generating Stations," McGraw-Hill. Underhill, "Coils and Magnet Wire" and "Magnets," McGraw-Hill. Abbott, "National Electrical Code Handbook," McGraw-Hill. Dyke, "Automobile and Gasoline Engine Encyclopedia," The Goodheart-Wilcox Co., Inc. Fink and Carrol, "Standard Handbook for Electrical Engineers," McGraw-Hill.

ELECTRICAL AND MAGNETIC UNITS

System of Units The International System of Units (**SI**) is being adopted universally. The SI system has its roots in the metre, kilogram, second (mks) system of units. Since a centimeter, gram, second (**cgs**) system has been widely used, and will still be used in some instances, Tables 15.1.1 and 15.1.2 are provided for conversion between the two systems. Basic SI units are metre, kilogram (mass), second, ampere, kelvin, mole (quantity), and candela (luminous intensity). Other SI units are derived from these basic units.

Electrical Units

(See Table 15.1.1.)

Current (I, i) The SI unit of current is the **ampere,** which is equal to one-tenth the absolute unit of current (**abampere**). The abampere of current is defined as follows: if 0.01 metre (1 centimetre) of a circuit is bent into an arc of 0.01 metre (1 centimetre) radius, the current is 1 abampere if the magnetic field intensity at the center is 0.01257 ampere per metre (1 oersted), provided the remainder of the circuit produces no magnetic effect at the center of the arc. One **international ampere** (9.99835 amperes) (dc) will deposit 0.001118 gram per second of silver from a standard silver solution.

Quantity (Q) The **coulomb** is the quantity of electricity transported in one second by a current of one ampere.

Potential Difference or Electromotive Force (V, E, emf) The **volt** is the difference of electric potential between two points of a conductor carrying a constant current of one ampere, when the power dissipated between these points is equal to one watt.

Resistance (R, r) The **ohm** is the electrical resistance between two points of a conductor when a constant difference of potential of one volt, applied between these two points, produces in this conductor a current of one ampere, this conductor not being the source of any electromotive force.

Resistivity (ρ) The resistivity of a material is the dc resistance between the opposite parallel faces of a portion of the material having unit length and unit cross section.

Conductance (G, g) The **siemens** is the electrical conductance of a conductor in which a current of one ampere is produced by an electric potential difference of one volt. One siemens is the reciprocal of one ohm.

Conductivity (γ) The conductivity of a material is the dc conductance between the opposite parallel faces of a portion of the material having unit length and unit cross section.

Capacitance (C) is that property of a system of conductors and dielectrics which permits the storage of electricity when potential difference exists between the conductors. Its value is expressed as a ratio of a quantity of electricity to a potential difference. A capacitance value is always positive. The farad is the capacitance of a capacitor between the plates of which there appears a difference of potential of one volt when it is charged by a quantity of electricity equal to one coulomb.

Permittivity or Dielectric Constant (ϵ_0) is the electrostatic energy stored per unit volume of a vacuum for unit potential gradient. The permittivity of a vacuum or free space is 8.85×10^{-12} farads per metre.

Relative Permittivity or Dielectric Constant (ϵ_r) is the ratio of electrostatic energy stored per unit volume of a dielectric for a unit potential gradient to the permittivity (ϵ_0) of a vacuum. The relative permittivity is a number.

Self-inductance (L) is the property of an electric circuit which determines, for a given rate of change of current in the circuit, the emf induced in the same circuit. Thus $e_1 = -L di_1/dt$, where e_1 and i_1 are in the same circuit and L is the coefficient of self-inductance.

The **henry** is the inductance of a closed circuit in which an electromotive force of one volt is produced when the electric current varies uniformly at a rate of one ampere per second.

Mutual Inductance (M) is the common property of two associated electric circuits which determines, for a given rate of change of current in one of the circuits, the emf induced in the other. Thus $e_1 = -M di_2/dt$ and $e_2 = -M di_1/dt$, where e_1 and i_1 are in circuit 1; e_2 and i_2 are in circuit 2; and M is the mutual inductance.

The **henry** is the mutual inductance of two separate circuits in which an electromotive force of one volt is produced in one circuit when the electric current in the other circuit varies uniformly at a rate of one ampere per second.

If M is the mutual inductance of two circuits and k is the coefficient of coupling, i.e., the proportion of flux produced by one circuit which links the other, then $M = k(L_1L_2)^{1/2}$, where L_1 and L_2 are the respective **self-inductances** of the two circuits.

Energy (J) in a system is measured by the amount of work which a system is capable of doing. The **joule** is the work done when the point of application of a force of one newton is displaced a distance of one metre in the direction of the force.

Power (W) is the time rate of transferring or transforming energy. The **watt** is the power which gives rise to the production of energy at the rate of one joule per second.

Table 15.1.1 Electrical Units

Quantity	Symbol	Equation	SI unit	SI unit symbol	CGS unit	Ratio of magnitude of SI to cgs unit
Current	I, i	$I = E/R; I = E/Z; I = Q/t$	Ampere	A	Abampere	10^{-1}
Quantity	Q, q	$Q = it; Q = CE$	Coulomb	C	Abcoulomb	10^{-1}
Electromotive force	E, e	$E = IR; E = W/Q$	Volt	V	Abvolt	10^{8}
Resistance	R, r	$R = E/I; R = \rho l/A$	Ohm	Ω	Abohm	10^{9}
Resistivity	ρ	$\rho = RA/l$	Ohm-metre	Ω·m	Abohm-cm	10^{11}
Conductance	G, g	$G = \gamma A/l; G = A/\rho l$	Siemens	S	Abmho	10^{-9}
Conductivity	γ	$\gamma = 1/\rho; \gamma = l/RA$	Siemens/meter	S/m	Abmho/cm	10^{-11}
Capacitance	C	$C = Q/E$	Farad	F	Abfarad*	10^{-9}
Permittivity	ϵ		Farads/meter	F/m	Stat farad*/cm	8.85×10^{-12}
Relative permittivity	ϵ_r	$\epsilon_r = \epsilon/\epsilon_0$	Numerical		Numerical	1
Self-inductance	L	$L = -N(d\phi/dt)$	Henry	H	Abhenry	10^{9}
Mutual inductance	M	$M = K(L_1L_2)^{1/2}$	Henry	H	Abhenry	10^{9}
Energy	J	$J = eit$	Joule	J	Erg	10^{7}
	kwh	kwh = kw/3600; 3.6 MJ	Kilowatthour	kWh		36×10^{12}
Active power	W	$W = J/t; W = EI \cos\theta$	Watt	W	Abwatt	10^{7}
Reactive power	jQ	$Q = EI \sin\theta$	Var	var	Abvar	10^{7}
Apparent power	VA	$VA = EI$	Volt-ampere	VA		
Power factor	pf	pf $= W/VA$; pf $= W/(W + jQ)$				1
Reactance, inductive	X_L	$X_L = 2\pi fL$	Ohm	Ω	Abohm	10^{9}
Reactance, capacitive	X_C	$X_C = 1/(2\pi fC)$	Ohm	Ω	Abohm	10^{9}
Impedance	Z	$Z = E/I; Z = R + j(X_L - X_C)$	Ohm	Ω	Abohm	10^{9}
Conductance	G	$G = R/Z^2$	Siemens	S	Abmho	10^{-9}
Susceptance	B	$B = X/Z^2$	Siemens	S	Abmho	10^{-9}
Admittance	Y	$Y = I/E; Y = G + jB$	Siemens	S	Abmho	10^{-9}
Frequency	f	$f = 1/T$	Hertz	Hz	Cps, Hz	1
Period	T	$T = 1/f$	Second	s	Second	1
Time constant	T	L/R; RC	Second	s	Second	1
Angular velocity	ω	$\omega = 2\pi f$	Radians/second	rad/s	Radians/second	1

*1 Abfarad (EMU Units) = 9×10^{-20} stat farads (ESU units).

Active Power (P) at the points of entry of a single-phase, two-wire circuit or of a polyphase circuit is the time average of the values of the instantaneous power at the points of entry, the average being taken over a complete cycle of the alternating current. The value of active power is given in **watts** when the rms currents are in amperes and the rms potential differences are in volts. For sinusoidal emf and current, $P = EI \cos\theta$, where E and I are the rms values of volts and currents, and θ is the phase difference of E and I.

Reactive Power (Q) at the points of entry of a single-phase, two-wire circuit, or for the special case of a sinusoidal current and sinusoidal potential difference of the same frequency, is equal to the product obtained by multiplying the rms value of the current by the rms value of the potential difference and by the sine of the angular phase difference by which the current leads or lags the potential difference. $Q = EI \sin\theta$. The unit of Q is the **var** (volt-ampere-reactive). One kilovar = 10^3 var.

Apparent Power (EI) at the points of entry of a single-phase, two-wire circuit is equal to the product of the rms current in one conductor multiplied by the rms potential difference between the two points of entry. Apparent power = EI.

Power Factor (pf) is the ratio of power to apparent power. pf $= P/EI = \cos\theta$, where θ is the phase difference between E and I, both assumed to be sinusoidal.

The **reactance** (X) of a portion of a circuit for a sinusoidal current and potential difference of the same frequency is the product of the sine of the angular phase difference between the current and potential difference times the ratio of the rms potential difference to the rms current, there being no source of power in the portion of the circuit under consideration. $X = (E/I) \sin\theta = 2\pi fL$ ohms, where f is the frequency, and L the inductance in henries; or $X = 1/2\pi fC$ ohms, where C is the capacitance in farads.

The **impedance** (Z) of a portion of an electric circuit to a completely specified periodic current and potential difference is the ratio of the rms value of the potential difference between the terminals to the rms value of the current, there being no source of power in the portion under consideration. $Z = E/I$ ohms.

Admittance (Y) is the reciprocal of impedance. $Y = I/E$ siemens.

The **susceptance** (B) of a portion of a circuit for a sinusoidal current and potential difference of the same frequency is the product of the sine of the angular phase difference between the current and the potential difference times the ratio of the rms current to the rms potential difference, there being no source of power in the portion of the circuit under consideration. $B = (I/E) \sin\theta$.

Magnetic Units

(See Table 15.1.2.)

Magnetic Flux (Φ, ϕ) is the magnetic flow that exists in any magnetic circuit.

The **weber** is the magnetic flux which, linking a circuit of one turn, produces in it an electromotive force of one volt as it is reduced to zero at a uniform rate in one second.

Table 15.1.2 Magnetic Units

Quantity	Symbol	Equation*	SI unit	SI unit symbol	CGS unit	Ratio of magnitude of SI to cgs unit
Magnetic flux	Φ, ϕ	$\phi = F/R$	Weber	wb	Maxwell	10^8
Magnetic flux density	β	$\beta = \phi/A$	Tesla	T	Gauss	10^4
Pole strength	Q_m	$Q_m = F/\beta$; $Q_m = Fl/NI\mu_0\mu_r$	Ampere-turns-metre	A·m		
			Unit pole		Unit pole	0.7958×10^7
Magnetomotive force	$\mathcal{F}$	$\mathcal{F} = NI$	Ampere-turns	A	Gilbert	1.257
Magnetic field intensity	H	$H = \mathcal{F}/l$	Ampere-turns per metre	A/m	Oersted	0.01257
Permeability air	μ_0	$\mu_0 = \beta/H$	Henry per metre	H/m	Gilbert per oersted	1.257×10^{-6}
Relative permeability	μ_r	$\mu_r = \mu/\mu_0$	Numeric		Numeric	1
Reluctivity	γ	$\gamma = 1/\mu_r$	Numeric		Numeric	1
Permeance	P	$P = \mu_0\mu_r A/l$	Henry	H		7.96×10^7
Reluctance	R	$R = l/\mu_0\mu_r A$	1/henry	1/H		1.257×10^{-8}

*l = length in metres; A = area in square metres; F = force in newtons; N = number of turns.

Magnetic Flux Density (β) is the ratio of the flux in any cross section to the area of that cross section, the cross section being taken normal to the direction of flux.

The **tesla** is the magnetic flux density given by a magnetic flux of one weber per square metre.

Unit Magnetic Pole, when concentrated at a point and placed one metre apart in a vacuum from a second unit magnetic pole, will repel or attract the second unit pole with a force of one **newton.**

The **weber** is the magnetic flux produced by a unit pole.

Magnetomotive Force ($\mathcal{F}$, mmf) produces magnetic flux and corresponds to electromotive force in an electric circuit.

The **ampere (turn)** is the unit of mmf.

Magnetic Field Intensity (H) at a point is the vector quantity which is measured by a mechanical force which is exerted on a unit pole placed at the point in a vacuum.

An **ampere per metre** is the unit of field intensity.

Permeability (μ) is the ratio of unit magnetic flux density to unit magnetic field intensity in air (B/H). The permeability of air is 1.257×10^{-6} **henry per metre.**

Relative permeability (μ_r) is the ratio of the magnetic flux in any element of a medium to the flux that would exist if that element were replaced with air, the magnetomotive force (mmf) acting on the element remaining unchanged ($\mu_r = \mu/\mu_0$).

The relative permeability is a number.

Permeance (P) of a portion of a magnetic circuit bounded by two equipotential surfaces, and by a third surface at every point of which there is a tangent having the direction of the magnetic induction, is the ratio of the flux through any cross section to the magnetic potential difference between the surfaces when taken within the portion under consideration. The equation for the permeance of the medium as defined above is $P = \mu_0\mu_r A/l$. Permeance is the **reciprocal of reluctance.**

Reluctivity (γ) of a medium is the reciprocal of its permeability.

Reluctance (R) is the **reciprocal of permeance.** It is the resistance to magnetic flow. In a homogeneous medium of uniform cross section, reluctance is equal to the length divided by the product of the area and permeability, the length and area being expressed in metre units. $R = l/A\mu_0\mu_r$, where $\mu_0 = 1.257 \times 10^{-6}$.

CONDUCTORS AND RESISTANCE

Resistivity, or **specific resistance,** is the resistance of a sample of the material having both a length and cross section of unity. The two most common resistivity samples are the centimetre cube and the cir mil·ft. If l is the length of a conductor of uniform cross section a, then its resistance is

$$R = \rho l/a \tag{15.1.1}$$

where ρ is the resistivity. With a cir mil·ft ρ is the resistance of a cir mil·ft and a is the cross section, cir mils. Since $v = la$ is the volume of a conductor,

$$R = \rho l^2/v = \rho v/a^2 \tag{15.1.2}$$

A **circular mil** is a unit of area equal to that of a circle whose diameter is 1 mil (0.001 in). It is the unit of area which is used almost entirely in this country for wires and cables. To obtain the cir mils of a solid cylindrical conductor, square its diameter expressed in mils. For example, the diameter of 000 AWG solid copper wire is 410 mils and its cross section is $(410)^2 = 168{,}100$ cir mils. The diameter in mils of a solid cylindrical conductor is the square root of its cross section expressed in cir mils

A **cir mil·ft** is a conductor having a length of 1 ft and a uniform cross section of 1 cir mil. In terms of the copper standard the resistance of a cir mil·ft of copper at 20°C is 10.371 Ω. As a first approximation 10 Ω may frequently be used.

At 60°C a cir mil·in of copper has a resistance of 1.0 Ω. This is a very convenient unit of resistivity for magnet coils since the resistance is merely the length of copper in inches divided by its cross section in cir mils.

Temperature Coefficient of Resistance The resistance of the pure metals increases with temperature. The resistance at any temperature t°C is

$$R = R_0(1 + \alpha t) \tag{15.1.3}$$

where R_0 is the resistance at 20°C and α is the **temperature coefficient of resistance.** For copper, $\alpha = 0.00393$.

Table 15.1.3 Properties of Metals and Alloys
(See Table 15.1.27 for properties of resistor alloys)

Metals	Resistivity, 20°C		Temperature coefficient of resistance at 20°C
	μΩ·cm	Ω·cir mil/ft	
Aluminum	2.828	17.01	0.00403
Antimony	42.1	251.0	0.0036
Bismuth	111.0	668.0	0.004
Brass	6.21	37.0	0.0015
Carbon: amorphous	3,800–4,100	......	(−)
Retort (graphite)	720–812*	....	(−)
Copper (drawn)	1.724	10.37	0.00393
Gold	2.44	14.7	0.0034
Iron: electrolytic	10.1	59.9	0.0064
Cast	75.2–98.8	448–588	
Wire	97.8	588	
Lead	22.0	132	0.00387
Molybdenum	5.78	34.8	
Monel metal	43.5	262	0.0019
Mercury	96.8	576	0.00089
Nickel	8.54	50.8	0.0041
Platinum	10.72	63.8	0.003
Platinum silver, 2Ag + 1Pt	24.6†	148.0	0.00031
Silver	1.628	9.8	0.0038
Steel: soft	15.9	95.8	0.0016
Glass hard	45.7	275	
Silicon (4 percent)	51.18	308	
Transformer	11.09	66.7	
Trolley wire	12.7	76.4	
Tin	11.63	70	0.0042
Tungsten	5.51	33.2	0.005
Zinc	5.97	35.58	0.0037

NOTE: Max working temperature: Cu, 260°C; Ni, 600°C; Pt, 1500°C.
*Furnace electrodes, 3,000°C.
†0°C.

With any initial temperature t_1, the resistance at temperature t°C is

$$R = R_1[1 + \alpha_1(t - t_1)] \qquad (15.1.4)$$

where R_1 is the resistance at temperature t_1°C and α_1 is the temperature coefficient of resistance at temperature t_1 [see Eq. (15.1.5)].

For any initial temperature t_1 the value of α_1 is

$$\alpha_1 = 1/(234.5 + t_1) \qquad (15.1.5)$$

Inferred Absolute Zero Between 100 and 0°C the resistance of copper decreases at a rate which is practically uniform and which if continued would give a resistance of zero at −234.5°C (an easy number to remember). If the resistance at t_1°C is R_1 and the resistance at t_2°C is R_2, then

$$R_2/R_1 = (234.5 + t_2)/(234.5 + t_1) \qquad (15.1.6)$$

EXAMPLE. The resistance of a copper coil at 25°C is 4.26 Ω. Determine its resistance at 45°C. Using Eq. (15.1.4) and $\alpha_1 = 1/(234.5 + 25) = 0.00385$, $R = 4.26[1 + 0.00385(45 - 25)] = 4.59$ Ω. Using Eq. (15.1.6) $R = 4.26(234.5 + 45)/(234.5 + 25) = 4.26 \times 1.077 = 4.59$ Ω.

The inferred absolute zero for aluminum is −228.

Materials The materials generally used for the transmission and distribution of electrical energy are copper, aluminum, and sometimes iron and steel. For resistors and heaters, iron, steel, commercial alloys, and carbon are most used.

Copper is the most widely used electrical conductor. It has high conductivity, relatively low cost, good resistance to oxidation, is readily soldered, and has good mechanical characteristics such as tensile strength, toughness, and ductility. Its tensile strength together with its low linear temperature coefficient of expansion are desirable characteristics in its use for overhead transmission lines. The international copper standard for 100 percent conductivity annealed copper is a density of 8.89 g/cm³ (0.321 lb/in³) and resistivity is given in Table 15.1.3. ASTM specifications for minimum conductivities of copper wire are as follows:

Conductor diam, in	*Soft or annealed*	*Medium hard drawn*	*Hard drawn*
0.040–0.324	98.16%	96.60%	96.16%
0.325–0.460	98.16%	97.66%	97.16%

Aluminum is used to considerable extent for high-voltage transmission lines, because its weight is one-half that of copper for the same conductance. Moreover, the greater diameter reduces corona loss. As it has 1.4 times the linear temperature coefficient of expansion, changes in sag with temperature are greater. Because of its lower melting point, spans may fail more readily with arc-overs. In aluminum cable steel-reinforced (ACSR), the center strand is a steel cable, which gives added tensile strength. Aluminum is used occasionally for bus bars because of its large heat-dissipating surface for a given conductance. The greater cross section for a given conductance requires a greater volume of insulation for a given voltage. When the ratio of the cost of aluminum to the cost of copper becomes economically favorable, aluminum is often used for insulated wires and cables. The international aluminum standard for 62 percent conductivity aluminum is a density of 2.70 g/cm^3 (0.0976 lb/in^3) and resistivity as given in Table 15.1.3.

Steel, either **galvanized** or **copper-covered** ("copperweld"), is used for high-voltage transmission spans where tensile strength is more important than high conductance. Steel is also used for third rails.

Copper alloys and **bronzes** are of increasing importance as electrical conductors. They have lower electrical conductivity but greater tensile strength and are resistant to corrosion. **Hitenso, Calsum bronzes, Signal bronze, Phono-electric,** and **Everdur** are bronzes containing phosphorus, silicon, manganese, or zinc. Their conductivities vary from 20 to 85 percent of 100 percent conductivity copper, and they have tensile strengths up to 130,000 lb/in^2, about twice that of hard-drawn copper. Such alloys were frequently used for trolley wires. Copper alloys having lower conductivity are usually classified as resistor materials.

In Table 15.1.3 are given the electrical properties of some of the pure metals and alloys.

American Wire Gage (AWG) The AWG (formerly Brown & Sharpe gage) is based on a constant ratio between diameters of successive gage numbers. The ratio of any diameter to the next smaller is 1.123, and the corresponding ratio of cross sections is $(1.123)^2 = 1.261$, or 1¼ approximately. $(1.123)^6$ is 2.0050, so that diameters differing by 6 gage numbers have a

Table 15.1.4 Working Table, Standard Annealed Copper Wire, Solid
[American Wire Gage (B & S)]

Gage no.	Diam, mils	Cross section		Ω per 1,000 ft		Ω/mi at 25°C (=77°F)	Weight per 1,000 ft, lb
		cir mils	in²	25°C (=77°F)	65°C (=149°F)		
0000	460.0	212,000	0.166	0.0500	0.0577	0.264	641.0
000	410.0	168,000	0.132	0.0630	0.0727	0.333	508.0
00	365.0	133,000	0.105	0.0795	0.0917	0.420	403.0
0	325.0	106,000	0.0829	0.100	0.116	0.528	319.0
1	289.0	83,700	0.0657	0.126	0.146	0.665	253.0
2	258.0	66,400	0.0521	0.159	0.184	0.839	201.0
3	229.0	52,600	0.0413	0.201	0.232	1.061	159.0
4	204.0	41,700	0.0328	0.253	0.292	1.335	126.0
5	182.0	33,100	0.0260	0.319	0.369	1.685	100.0
6	162.0	26,300	0.0206	0.403	0.465	2.13	79.5
7	144.0	20,800	0.0164	0.508	0.586	2.68	63.0
8	128.0	16,500	0.0130	0.641	0.739	3.38	50.0
9	114.0	13,100	0.0103	0.808	0.932	4.27	39.6
10	102.0	10,400	0.00815	1.02	1.18	5.38	31.4
11	91.0	8,230	0.00647	1.28	1.48	6.75	24.9
12	81.0	6,530	0.00513	1.62	1.87	8.55	19.8
13	72.0	5,180	0.00407	2.04	2.36	10.77	15.7
14	64.0	4,110	0.00323	2.58	2.97	13.62	12.4
15	57.0	3,260	0.00256	3.25	3.75	17.16	9.86
16	51.0	2,580	0.00203	4.09	4.73	21.6	7.82
17	45.0	2,050	0.00161	5.16	5.96	27.2	6.20
18	40.0	1,620	0.00128	6.51	7.51	34.4	4.92
19	36.0	1,290	0.00101	8.21	9.48	43.3	3.90
20	32.0	1,020	0.000802	10.4	11.9	54.9	3.09
21	28.5	810	0.000636	13.1	15.1	69.1	2.45
22	25.3	642	0.000505	16.5	19.0	87.1	1.94
23	22.6	509	0.000400	20.8	24.0	109.8	1.54
24	20.1	404	0.000317	26.2	30.2	138.3	1.22
25	17.9	320	0.000252	33.0	38.1	174.1	0.970
26	15.9	254	0.000200	41.6	48.0	220	0.769
27	14.2	202	0.000158	52.5	60.6	277	0.610
28	12.6	160	0.000126	66.2	76.4	350	0.484
29	11.3	127	0.0000995	83.4	96.3	440	0.384
30	10.0	101	0.0000789	105	121	554	0.304
31	8.9	79.7	0.0000626	133	153	702	0.241
32	8.0	63.2	0.0000496	167	193	882	0.191
33	7.1	50.1	0.0000394	211	243	1,114	0.152
34	6.3	39.8	0.0000312	266	307	1,404	0.120
35	5.6	31.5	0.0000248	335	387	1,769	0.0954
36	5.0	25.0	0.0000196	423	488	2,230	0.0757
37	4.5	19.8	0.0000156	533	616	2,810	0.0600
38	4.0	15.7	0.0000123	673	776	3,550	0.0476
39	3.5	12.5	0.0000098	848	979	4,480	0.0377
40	3.1	9.9	0.0000078	1,070	1,230	5,650	0.0200

ratio of approximately 2; cross sections differing by 3 gage numbers also have a ratio of approximately 2. The ratio of cross sections differing by 2 numbers is $(1.261)^2 = 1.590$, or 1.6 approximately. The ratio of cross sections differing by 10 numbers is approximately 10. The gage ordinarily extends from no. 40 to 0000 (4/0). Wires larger than 0000 must be stranded, and their cross section is given in cir mils.

The diameter of no. 10 wire is 102.0 mils. As an approximation this may be considered as being 100 mils; the cross section is 10,000 cir mils; the resistance is 1 Ω per 1,000 ft; and the weight of 1,000 ft is $31.4(10\pi)$ lb. Also the weight of 1,000 ft of no. 2 is 200 lb. These facts give many short cuts in estimating resistances and weights of various gage numbers.

Lay Cables In order to obtain sufficient flexibility, wires larger than 0000 are stranded, and they are designated by their circular mils. Smaller wires may be stranded also since sizes as small as no. 4 when insulated are usually too stiff for easy handling. Lay cables are made up geometrically as shown in Fig. 15.1.1. Six strands will just fit around the single central conductor; the number of strands in each succeeding layer increases by 6. The number of strands that can thus be layed up are 1, 7, 19, 37, 61, 91, 127, etc. In order to obtain sufficient flexibility with large cables, the strands themselves frequently consist of stranded cable.

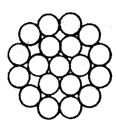


Fig. 15.1.1 Makeup of a 19-strand cable.

The **resistance of cables** is readily computed from Eq. (15.1.1), using the cir mil ft as the unit of resistivity.

EXAMPLE. Determine the resistance of 3,500 ft of 800,000 cir mil cable at 20°C. *Answer:* ρ (of a cir mil·ft) = 10.37. $R = 10.37 \times 3{,}500/800{,}000 = 0.0454\ \Omega$.

$\rho = 10\ \Omega$/cir mil·ft is often sufficiently accurate for practical purposes.

ELECTRICAL CIRCUITS

Ohm's law states that, with a steady current, the current in a circuit is **directly** proportional to the **total** emf acting in the circuit and is **inversely** proportional to the total resistance of the circuit. The law may be expressed by the following three equations:

$$I = E/R \tag{15.1.7}$$
$$E = IR \tag{15.1.8}$$
$$R = E/I \tag{15.1.9}$$

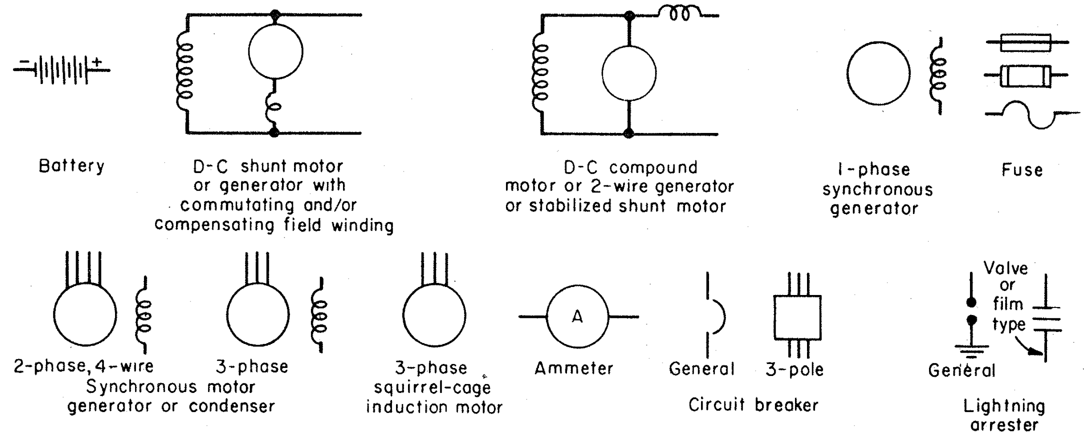
where E is the emf, V; R the resistance, Ω; and I the current, A.

Series Circuits The combined resistance of a number of series-connected resistors is the sum of their separate resistances. When batteries or other sources of emf are connected in series, the total emf of the combination is the sum of the

Table 15.1.5 Bare Concentric Lay Cables of Standard Annealed Copper

AWG no.	cir mils	Ω per 1,000 ft		Weight per 1,000 ft, lb	Standard concentric standing		
		25°C (=77°F)	65°C (=149°F)		No. of wires	Diam of wires, mils	Outside diam, mils
	2,000,000	0.00539	0.00622	6,180	127	125.5	1,631
	1,700,000	0.00634	0.00732	5,250	127	115.7	1,504
	1,500,000	0.00719	0.00830	4,630	91	128.4	1,412
	1,200,000	0.00899	0.0104	3,710	91	114.8	1,263
	1,000,000	0.0108	0.0124	3,090	61	128.0	1,152
	900,000	0.0120	0.0138	2,780	61	121.5	1,093
	850,000	0.0127	0.0146	2,620	61	118.0	1,062
	750,000	0.0144	0.0166	2,320	61	110.9	998
	650,000	0.0166	0.0192	2,010	61	103.2	929
	600,000	0.0180	0.0207	1,850	61	99.2	893
	550,000	0.0196	0.0226	1,700	61	95.0	855
	500,000	0.0216	0.0249	1,540	37	116.2	814
	450,000	0.0240	0.0277	1,390	37	110.3	772
	400,000	0.0270	0.0311	1,240	37	104.0	728
	350,000	0.0308	0.0356	1,080	37	97.3	681
	300,000	0.0360	0.0415	926	37	90.0	630
	250,000	0.0431	0.0498	772	37	82.2	575
0000	212,000	0.0509	0.0587	653	19	105.5	528
000	168,000	0.0642	0.0741	518	19	94.0	470
00	133,000	0.0811	0.0936	411	19	83.7	418
0	106,000	0.102	0.117	326	19	74.5	373
1	83,700	0.129	0.149	258	19	66.4	332
2	66,400	0.162	0.187	205	7	97.4	292
3	52,600	0.205	0.237	163	7	86.7	260
4	41,700	0.259	0.299	129	7	77.2	232

NOTE: See Table 15.1.21 for the carrying capacity of wires.
SOURCE: From *NBS Cir.* 31.



Meters and Instruments

A letter or a letter combination from the following list shall be placed within the circle to indicate the function of the meter or instrument unless some other identification is provided in the circle and explained in the diagram.

A	Ammeter	F	Frequency meter	PF	Power Factor meter	V	Voltmeter
AH	Amp-hr meter	G	Galvanometer	REC	Recording meter	VA	Volt-ammeter
CRO	Cath. Ray Oscill.	μA or UA	Microammeter	SY	Synchroscope	VAR	Varmeter
DM	Demand meter	MA	Milliammeter	+°	Temperature meter	W	Wattmeter
DB	Decibel meter	OHM	Ohmmeter	VARH	Varhour meter	WH	Watthour meter

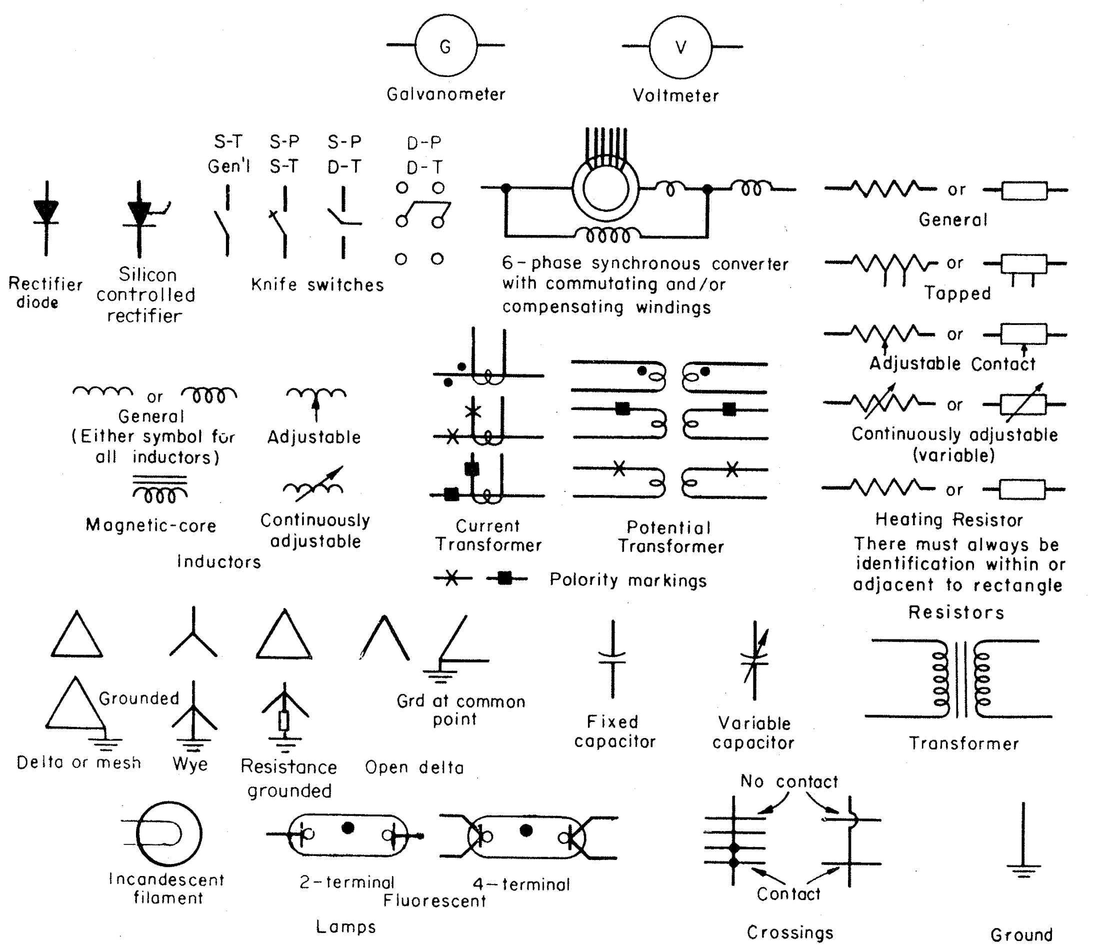


Fig. 15.1.2 Diagrammatic symbols for electrical machinery and apparatus. *(American Standard, "Graphic Symbols for Electrical and Electronic Diagrams," ANSI/IEEE, 315, 1975.)*

separate emfs. The open-circuit emf of a battery is the total generated emf and can be measured at the battery terminals only when no current is being delivered by the battery. The internal resistance is the resistance of the battery alone. The current in a circuit connected in series with a source of emf is $I = E/(R + r)$, where E is the open-circuit emf, R the external resistance, and r the internal resistance of the source of emf.

Parallel Circuits The combined conductance of a number of parallel-connected resistors is the sum of their separate conductances.

$$G = G_1 + G_2 + G_3 + \cdots \quad (15.1.10)$$

$$\frac{1}{R} = \frac{1}{R_1} + \frac{1}{R_2} + \frac{1}{R_3} + \cdots \quad (15.1.11)$$

The equivalent resistance for two parallel resistors having resistances R_1, R_2 is

$$R = R_1R_2/(R_1 + R_2) \quad (15.1.12)$$

The equivalent resistance for three parallel resistors having resistances R_1, R_2, R_3 is

$$R = \frac{R_1R_2R_3}{R_1R_2 + R_2R_3 + R_3R_1} \quad (15.1.13)$$

and for four parallel resistors having resistances R_1, R_2, R_3, R_4

$$R = \frac{R_1R_2R_3R_4}{R_1R_2R_3 + R_2R_3R_4 + R_3R_4R_1 + R_4R_1R_2} \quad (15.1.14)$$

To obtain the resistance of combined series and parallel resistors, the equivalent resistance of each parallel portion is obtained separately and then these equivalent resistances are added to the series resistances according to the principles stated above.

Kirchhoff's laws (derived from Ohm's law) make it possible to solve many circuit networks that would otherwise be difficult to solve. The first law states that: *In any branching network of wires the algebraic sum of the currents in all the wires that meet at a point is zero.* The second law states that: *The sum of all the electromotive forces acting around a complete circuit is equal to the sum of the resistances of its separate parts multiplied each by the strength of the current in it, or the total change of potential around any closed circuit is zero.*

In applying Kirchhoff's laws the following rules should be observed. Currents going toward a junction should be preceded by a plus sign. Currents going away from a junction should be preceded by a minus sign. A rise in potential should be preceded by a plus sign. (This occurs in going through a source of emf from the negative to the positive terminal, and in going through resistance in opposition to the direction of current.) A drop in potential should be preceded by a minus sign. (This occurs in going through a source of emf from the positive to the negative terminal and in going through resistance in conjunction with the current.)

The application of Kirchhoff's laws is illustrated by the following example.

EXAMPLE. Determine the three currents I_1, I_2, and I_3 in the circuit network (Fig. 15.1.3). The arrows show the assumed directions of the three currents.

Applying Kirchhoff's second law to circuit *abcdea*,

$$+4 + 0.2I_1 + 0.5I_1 - 3I_2 + 2 - 0.1I_2 + I_1 = 0$$

or

$$+6 + 1.7I_1 - 3.1I_2 = 0 \quad \text{(I)}$$

and for *edcfge*,

$$-2 + 0.1I_2 + 3I_2 + I_3 + 3 + 0.3I_3 = 0$$

or

$$+1 + 3.1I_2 + 1.3I_3 = 0 \quad \text{(II)}$$

Applying Kirchhoff's first law to junction c,

$$-I_1 - I_2 + I_3 = 0 \quad \text{(III)}$$

Solving (I), (II), and (III) simultaneously gives $I_1 = -2.56$, $I_2 = +0.53$, and $I_3 = -2.03$. The minus signs before I_1 and I_3 show that the actual directions of these two currents are opposite the assumed directions.

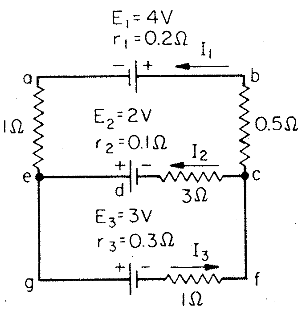


Fig. 15.1.3 Electric network and Kirchhoff's laws.

Electrical Power With direct currents the electrical power is given by the product of the volts and amperes. That is,

$$P = EI \quad \text{W} \quad (15.1.15)$$

Also, by substituting for E and I Eqs. (8) and (7),

$$P = I^2R \quad \text{W} \quad (15.1.16)$$

$$P = E^2/R \quad \text{W} \quad (15.1.17)$$

The watt is too small a unit for many purposes. Hence, the **kilowatt** (kW) is used. 746 watts = 1 hp = 0.746 kW; 1 kW = 1.340 hp. The **kilowatthour (kWh)** is the common engineering unit of electrical energy.

Joule's Law When an electric current flows through resistance, the number of heat units developed is proportional to the square of the current, directly proportional to the resistance, and directly proportional to the time that the current flows. $h = i^2rt$, where h = number of joules; i = current, A; r = resistance, Ω; and t = time, s. h (in Btu) = $0.0009478i^2rt$.

MAGNETISM

Magnetic Circuit

The magnetic circuit is analogous to the electric circuit in that the flux Φ is proportional to the magnetomotive force $\mathcal{F}$ and inversely proportional to the reluctance $\mathcal{R}$ or magnetic resistance. Thus

$$\Phi = \mathcal{F}/\mathcal{R} \quad (15.1.18)$$

Compare with Eq. (15.1.7). Φ is in webers, where the weber is the SI unit of flux, $\mathcal{F}$ in ampere-turns, and R in SI reluctance units. In the cgs system, ϕ is in maxwells, $\mathcal{F}$ is in gilberts, and $\mathcal{R}$ is in cgs reluctance units.

$$\mathcal{R} = l/\mu_r\mu_v A \quad (15.1.19)$$

where μ_r is **relative permeability** (commonly called permeability, μ), a property of the magnetic material, and μ_v is the permeability of evacuated space $= 4\pi \times 10^{-7}$, and A is in square metres. In the cgs system $\mu_v = 1$

$$\mathcal{R} = \frac{l}{\mu_r(4\pi \times 10^{-7})A} = \frac{l}{\mu_r(1.257 \times 10^{-6})A} \qquad \textbf{(15.1.20)}$$

l is in metres and A in square metres.

The unit of **flux density** in the SI system is the tesla, which is equal to the number of webers per square metre taken perpendicular to their direction. One ampere-turn between opposite faces of a metre cube of a magnetic medium produces μ_r tesla. For air, $\mu_r = 4\pi \times 10^{-7}$. In the cgs system the unit of flux density is gauss $= 10^4$T (see Table 15.1.2).

Magnetic-circuit calculations cannot be made with the same degree of accuracy as electric-circuit calculations because of several factors. The cross-sectional dimensions of the magnetic circuit are large relative to its length; magnetic paths are irregular, and their geometry can only be approximated as with the air gap of electric machines, which usually have slots on one or both sides of the gap.

Magnetic flux cannot be confined to definite magnetic paths, but a considerable proportion usually takes paths external to the circuit giving magnetic leakage (see Fig. 15.1.7). The relative permeability of iron varies over wide ranges with the flux density and with the previous magnetic condition (see Fig. 15.1.5). These variations of relative permeability cannot be expressed by any simple equation. Although the foregoing factors prevent the obtaining of extremely high accuracy in magnetic calculations, yet, with experience, it is possible to design magnetic circuits with a precision that is satisfactory for all practical purposes.

The **magnetomotive force** $\mathcal{F}$ in Eq. **(15.1.18)** is expressed in ampere-turns $= NI$, where N is the number of turns linked with the circuit and I is the current, A. The unit of reluctance is the reluctance of a 1-m cube of air. The total reluctance is proportional to the length and inversely proportional to the cross-sectional area of the magnetic circuit, which is analogous to electrical resistance. Hence the reluctance of any given path of uniform cross section A is $l/A\mu$, where l = length of path, cm; A = its cross section, cm^2, and μ = permeability. Reluctances in series are added to obtain their combined reluctance. Ohm's law of the magnetic circuit becomes

$$\Phi = \frac{NI}{l_1/A_1\mu_1 + l_2/A_2\mu_2 + l_3/A_3\mu_3 \cdots} \text{ Mx} \qquad \textbf{(15.1.21)}$$

where l_1, A_1, μ_1, etc., are the lengths, cross sections, and relative permeabilities of each series part of the circuit.

EXAMPLE. In Fig. 15.1.4 is shown a magnetic circuit of cast steel with a 0.4-cm air gap. The cross section of the core is 4 cm square. There are 425 turns wound on the core and the current is 10 A. The relative permeability of the steel at the operating flux density is 1,100.

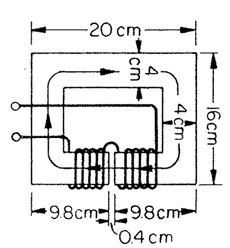


Fig. 15.1.4 Magnetic circuit.

Assume that the path of the flux is as shown, the average path at the corners being quarter circles. Neglect fringing at the air gap and any leakage. Determine the flux and the flux density.

Using the SI system, the length of the iron is 0.522 m, the length of the air gap is 0.004 m, and the cross section of the iron and air gap is 0.0016 m^2.

$$\Phi = \frac{425 \times 10}{\dfrac{0.522}{1{,}100 \times 4\pi \times 10^{-7} \times 0.0016} + \dfrac{0.004}{4\pi \times 10^{-7} \times 0.0016}}$$

$$= 0.00191 \text{ Wb}$$

Using the cgs system, the length of the magnetic path in the iron $= 12 + 8 + 8 + 5.8 + 5.8 + 4\pi = 52.2$ cm. From Eq. **(15.1.21)**,

$$\Phi = \frac{0.4\pi \times 425 \times 10}{[52.2/(16 \times 1{,}100)] + (0.4/16)} = 191{,}000 \text{ Mx}$$

$$B = \frac{191{,}000}{16} = 11{,}940 \text{ G}$$

Magnetization and Permeability Curves The magnetic permeability of air is a constant and is taken as unity. The relative permeability of iron and other magnetic substances varies with the flux density. In Fig. 15.1.5 is shown a magnetization

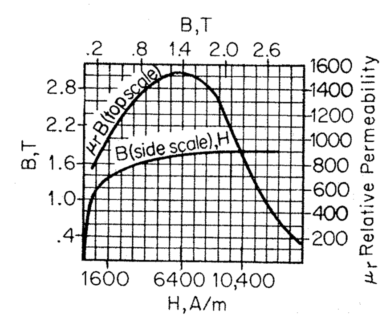


Fig. 15.1.5 Magnetization and relative-permeability curves for cast steel.

curve for cast steel in which the flux density B in tesla is plotted as a function of the field intensity, amperes per metre, H. Also the relative permeability $\mu_r = B/H$ is plotted as a function of the flux density B. Note the wide range over which the relative permeability varies. No satisfactory equation has been found to express the relation between magnetizing force and flux density and between relative permeability and flux density. If an attempt is made to solve Eq. **(15.1.21)** for flux, the factors μ_1, μ_2, etc., are unknown since they are functions of the flux density, which is being determined. The simplest method is one of trial and error, i.e., a value of flux, and the corresponding permeability, is first assumed, the equation solved for the flux, and if the computed flux differs widely from the assumed flux, a second approximation is made, etc. In nearly all magnetic designs either the flux or flux density is the independent variable, and it is required to find the necessary ampere-turns to produce them. Let the flux $\Phi = BA$ where B is the flux density, G. Then

$$\Phi = BA = 0.4\pi NI(l/A\mu_r) \qquad \textbf{(15.1.22)}$$

and

$$NI = Bl/\mu_0\mu_r = 0.796Bl/\mu_r \times 10^6$$

Equation **(15.1.22)** shows that the necessary ampere-turns are proportional to the **flux density** and the length of path and are inversely proportional to the relative permeability.

With air and nonmagnetic substances μ_r [Eq. (15.1.22)] becomes unity, and

$$NI = 0.796Bl \times 10^6 \qquad (15.1.23)$$

in meter units. With inch units

$$NI = 0.313B'l' \qquad (15.1.24)$$

where B' is the flux density, Mx/in^2; and l' the length of the magnetic path, in.

EXAMPLE. The average flux density in the air gap of a generator is 40,000 Mx/in^2, and the effective length of the gap is 0.2 in. How many ampere-turns per pole are necessary for the gap?

$$NI = 0.313 \times 40{,}000 \times 0.2 = 2{,}500$$

Since the relation of μ_r to flux density B in Eq. (15.1.22) is not simple, the relation of ampere-turns per unit length of magnetic circuit to flux density is ordinarily shown graphically.

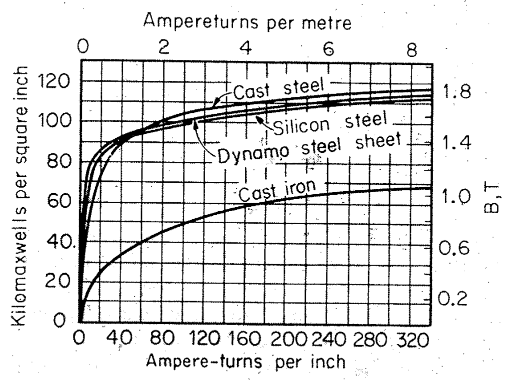


Fig. 15.1.6 Typical magnetization curves.

Typical curves of this character are shown in Fig. 15.1.6, inch units being used although scales of tesla, and ampere turns per metre are also given. To determine the number of ampere-turns necessary to produce a given total flux in a magnetic circuit composed of several parts in series having various lengths, cross sections, and relative permeabilities, determine the flux density if the cross section is fixed, or otherwise choose a cross section to give a suitable flux density. From the magnetization curve obtain the ampere-turns necessary to drive this **flux density** through a unit length of the portion of the circuit considered and multiply by the length. Add together the ampere-turns required for each series part of the magnetic circuit to obtain the total ampere-turns necessary to give the assumed flux.

It is desirable to operate magnetic circuits at as high flux densities as is practicable in order to reduce the amount of iron and copper. The air gaps of dynamos are operated at average densities of 40,000 to 50,000 Mx/in^2. Higher densities increase the exciting ampere-turns and tooth losses. At 45,000 Mx/in^2 the flux density in the teeth may be as high as 120,000 to 130,000 Mx/in^2. The flux densities in transformer cores are limited as a rule by the permissible losses. At 60 Hz and with silicon steel the maximum density is 60,000 to 70,000 Mx/in^2, at 25 Hz the density may run as high as 75,000 to 90,000 Mx/in^2. With laminated cores, the net iron is approximately 0.9 the gross cross section.

Magnetic Leakage It is impossible to confine all magnetic flux to any desired path since there is no known insulator of magnetic flux. Figure 15.1.7 shows the magnetic circuit of a modern four-pole dynamo. A considerable proportion of the useful magnetic flux leaks between the pole shoes and cores, rather than across the air gap. The ratio of the maximum flux, which exists in the field cores, to the useful flux, i.e., the flux

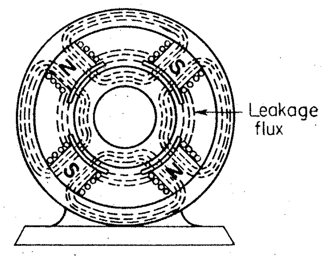


Fig. 15.1.7 Magnetic circuit of four-pole dynamo with leakage flux.

that crosses the air gap, is the **coefficient of leakage.** This coefficient must always be greater than unity and in carefully designed dynamos may be as low as 1.15. It is frequently as high as 1.30. Although the geometry of the leakage-flux paths is not simple, the leakage flux may be determined by approximations with a fair degree of accuracy.

Magnetic Hysteresis The magnetization curves shown in Figs. 15.1.5 and 15.1.6 are called **normal curves.** They are taken with the magnetizing force continuously increased from zero. If at any point the magnetizing force be decreased, a greater value of flux density for any given magnetizing force will result. The effect of carrying iron through a complete cycle of magnetization, both positive and negative, is shown in Fig. 15.1.8.

The curve OKB, taken with increasing values of magnetizing force per centimeter H, is the **normal induction** curve. If after the magnetizing force has reached the value OA, it is

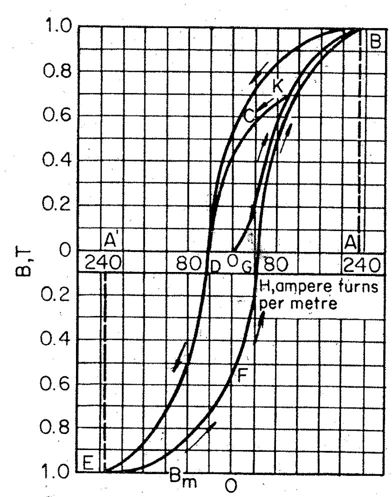


Fig. 15.1.8 Hysteresis loop for dynamo steel.

decreased, the magnetic flux density B will decrease in accordance with curve BCD, between A and O the values being much greater than those given by the normal curve, i.e., the flux density lags the magnetizing force. At zero magnetizing force, the flux density is OC, call the **remanence.** A negative

Table 15.1.6 Steinmetz Coefficients

Hard tungsten steel	0.058	Annealed cast steel	0.008
Hard cast steel	0.025	Ordinary sheet iron	0.004
Forged steel	0.020	Pure iron	0.003
Cast iron	0.013	Annealed iron sheet	0.002
Electrolytic iron	0.009	Best annealed sheet	0.001
Soft machine steel	0.009	Silicon steel sheet	0.00046
		Permalloy	0.0001

magnetizing force OD, called the **coercive force,** is required to bring the flux density to zero. If the magnetizing force is increased negatively to OA', the flux density will be given by the curve DE. If the magnetizing force is increased positively from A' to A, the flux density will be given by the curve $EFGB$, which is similar to the curve $BCDE$. OF is the negative remanence and OG again is the coercive force. The complete curve is called a **hysteresis loop.** When the normal curve reaches the point K, if the magnetizing force is then decreased, another hysteresis loop, a portion of which is shown at KL, will be obtained. It is seen that the flux density lags the magnetizing force throughout.

The energy dissipated per cycle is proportional to the area of the loop and is equal to $(1/4\pi)\int H\, dB$ ergs/(Hz)(cm³). For moderately high densities the energy loss per cycle varies according to the **Steinmetz law**

$$W = 10\eta B_m^{1.6}\ \mathrm{W \cdot s/m^3} \qquad \textbf{(15.1.25)}$$

where B_m is the maximum value of the flux density, T (Fig. 15.1.8). Table 15.1.6 gives values of the Steinmetz coefficient η, for common magnetic steels.

A permanent increase in the hysteresis constant occurs if the temperature of operation remains for some time above 80°C. This phenomenon is known as **aging** and may be much reduced by proper annealing of the iron. Silicon steels containing about 3 percent silicon have a lower hysteresis loss, somewhat larger eddy-current loss, and are practically nonaging.

Eddy-current losses, also known as Foucault-current losses, occur in iron subjected to cyclic magnetization. Eddy-current losses are reduced by laminating the iron, which subdivides the emf and increases greatly the length of path of the parasitic currents. Eddy currents have also a screening effect, which tends to prevent the flux penetrating the iron. Hence laminating also allows the full cross section of the iron to be utilized unless the frequency is too high.

Eddy-current loss in sheets is given by

$$P_e = (\pi t f B_m)^2/6\rho 10^{16} \qquad \mathrm{W/cm^3} \qquad \textbf{(15.1.26)}$$

where t = thickness, cm; f = frequency, Hz; B_m = the maximum flux density, G; ρ = the resistivity, Ω·cm.

Relations of Direction of Magnetic Flux to Current Direction The direction of the magnetizing force of a current is at right angles to its direction of flow. Magnetic lines about a cylindrical conductor carrying current exist in circular planes concentric with and normal to the conductor. This is illustrated in Fig. 15.1.9*a*. The ⊕ sign, corresponding to the feathered end of the arrow, indicates a direction of current away from the observer; a ⊙ sign, corresponding to the tip of an arrow, indicates a direction of current toward the observer.

Corkscrew Rule The direction of the current and that of the resulting magnetic field are related to each other as the forward travel of a corkscrew and the direction in which it is rotated.

Hand Rule Grasp the conductor in the right hand with the thumb pointing in the direction of the current. The fingers will then point in the direction of the lines of flux.

The applications of these rules are illustrated in Fig. 15.1.9. If the currents in parallel conductors are in opposite directions

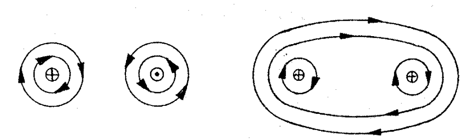


Fig. 15.1.9 Currents in (*a*) opposite directions; (*b*) in the same direction.

(Fig. 15.1.9*a*), the conductors tend to move apart; if the currents in parallel conductors are in the same direction (Fig. 15.1.9*b*), the conductors tend to come together. The magnetic lines act like stretched rubber bands and, in attempting to contract, tend to pull the two conductors together.

The relation of the direction of current in a solenoid helix to the direction of flux is shown in Fig. 15.1.10. Figure 15.1.11 shows the effect on a uniform field of placing a conductor carrying current in that field and normal to it. In (*a*) the direction of the current is toward the observer. By applying the corkscrew rule it is seen that the current weakens the field immediately above it and strengthens the field immediately below it. The reverse is true in (*b*), where the direction of the current is away from the observer.

Figure 15.1.11 is illustrative of the force developed on a conductor carrying current in a magnetic field. In (*a*) the conductor will tend to move upward owing to the stretching of the magnetic lines beneath it. Similarly, the conductor in (*b*) will tend to move downward. This principle is the basis of motor action. (See also Magnets.)

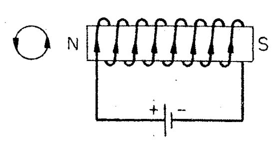


Fig. 15.1.10 Direction of current and poles in solenoid.

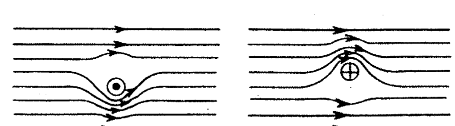


Fig. 15.1.11 Effect of current on uniform magnetic field.

Table 15.1.7 Battery Types and Applications

Battery type	Cell type	Nominal cell voltage	Capacity, wH/kg	Applications
		Primary		
Leclanche (zinc-carbon)	Dry	1.5	22–44	Flashlights, emergency lights, radios
Zinc-mercury (Ruben)	Dry	1.34	90–110	Medical, marine, space, laboratory, and emergency devices
Zinc-alkaline-manganese dioxide	Dry	1.5	66	Models, cameras, shavers, lights
Silver or cuprous chloride-magnesium	Wet		55–120	Disposable devices: torpedoes, rescue beacons, meteorological balloons
		Secondary		
Lead-acid	Wet	2		Automotive, industrial trucks, railway, station service
Lead-calcium	Wet	2		Standby
Edison (nickel-iron)	Wet	1.2		Industrial trucks; boat and train lights
Nickel-cadium	Wet	1.2	28	Engine starting, emergency lighting, station service
Silver oxide-cadium	Wet	1.4	45–65	Space
Silver-zinc	Wet	1.55	90–155	Models, photographic equipment, missiles

BATTERIES

In an **electric cell,** or **battery,** chemical energy is converted into electrical energy. The word *battery* may be used for a single cell or for an assembly of cells connected in series or parallel. A battery utilizes the potential difference which exists between different elements. When two different elements are immersed in electrolyte an emf exists tending to send current within the cell from the negative pole, which is the more highly electropositive, to the positive pole. The **poles,** or **electrodes** of a battery form the junction with the external circuit.

If the external circuit is closed, current flows from the battery at the **positive electrode,** or **anode,** and enters the battery at the **negative electrode,** or **cathode.**

In a **primary battery** the chemically reacting parts **require renewal;** in a **secondary battery,** the electrochemical processes **are reversible** to a high degree and the chemically reacting parts are restored after partial or complete discharge by reversing the direction of current through the battery. See Table 15.1.7 for a summary of battery types and applications.

Electromotive force of a battery is the total potential difference existing between the electrodes on open circuit. When current flows, the potential difference across the terminal drops because of the resistance drop within the cell and because of **polarization.**

Polarization When current flows in a battery, hydrogen is deposited on the cathode. This produces two effects, both of which reduce the terminal voltage of the battery. The hydrogen in contact with the cathode constitutes a hydrogen battery which opposes the emf of the battery; the hydrogen bubbles reduce the contact area of the electrolyte with the cathode, thus increasing the battery resistance. The most satisfactory method of reducing polarization is to have present at the cathode some compound that supplies negative ions to combine with the positive hydrogen ions at the plate. In the Leclanché cell, manganese peroxide in contact with the carbon cathode serves as a depolarizer, its oxygen ion combining with the hydrogen ion to form water.

If E is the emf of the cell, E_p the emf of polarization, r the internal resistance, V the terminal voltage, when current I flows, then

$$V = (E - E_p) - Ir \tag{15.1.27}$$

Primary Batteries

Dry Cells A dry cell is one in which the electrolyte exists in the form of a jelly, is absorbed in a porous medium, or is otherwise restrained from flowing from its intended position, such a cell being completely portable and the electrolyte nonspillable. The Leclanché cell consists of a cylindrical zinc container which serves as the negative electrode and is lined with spe-

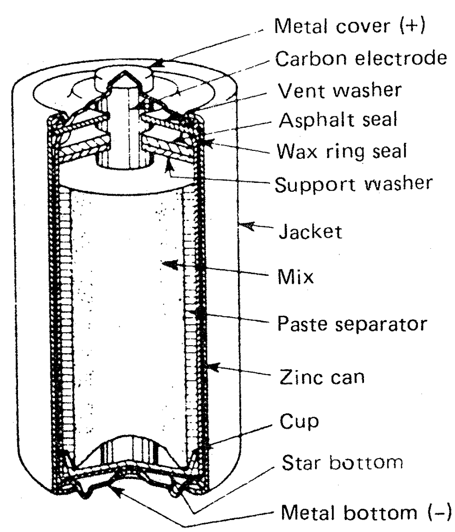


Fig. 15.1.12 Cross section of standard round zinc-carbon cell. *(From "Standard Handbook for Electrical Engineers," Fink and Carrol, McGraw-Hill, NY, copyright 1968.)*

cially prepared paper, or some similar absorbent material, to prevent the mixture of carbon and manganese dioxide, which is tamped tightly around the positive carbon electrode, from coming in contact with the zinc. The absorbent lining and the mixture are moistened with a solution of zinc chloride and sal ammoniac. In smaller cells the manganese-carbon mixture is often molded into a cylinder around the carbon electrode, the whole is then set into the zinc cup, and the space between the molded mixture and the zinc is filled with electrolyte made into a paste in such a manner that it can be solidified by either standing or heating. The top of the cell is closed with a sealing compound, and the cell is placed in a carboard container. The emf of a dry cell when new is 1.4 to 1.6 V.

In **block assembly** the dry cells, especially in the smaller sizes, are assembled in series and sealed in blocks of insulating compound with only two terminals and, sometimes, intermediate taps brought out. This type of battery is used for radio B and C batteries. Another construction is to build the battery up of layers in somewhat the manner of the old voltaic pile. Each cell consists of a layer of zinc, a layer of treated paper, and a flat cake of the manganese-carbon mixture. The cells are separated by layers of a special material which conducts electricity but which is impervious to electrolyte. A sufficient number of such cells are built up to give the required voltage and the whole battery is sealed into the carton.

Leclanché cells are generally available in sizes ranging from small, thin penlight batteries to large assemblies of cells in series or parallel for special high-voltage or high-current applications.

The **efficiency** of a standard-size dry battery depends on the rate at which it is discharged. Up to a certain rate the lower the discharge rate, the greater the efficiency. Above this rate the efficiency decreases (see *Natl. Bur. Stand. Circ.* 79, p. 39).

When used efficiently, a 6-in dry cell will give over 30 A·h of service. As ordinarily used, however, the dry cell gives no more than 8 to 10 A·h of service and at times even less. The 1¼ by 2¼ in flashlight battery is usually employed with a lamp taking 0.25 to 0.35 A. Under these conditions 3 A·h or thereabouts may be expected if the battery is used for not more than an hour or so a day. The so-called "heavy-duty" radio battery will give about 8 to 10 A·h when efficiently used.

For the best results 6-in dry cells should not be used for current drains of over 0.5 A except for very short periods of time. Flashlight batteries should not be used for higher than the preceding current drain, and heavy-duty radio batteries will give best results if the current drain is kept below 25 mA.

Dry cells should be stored in a cool, dry place. Extreme heat during storage will shorten their life. The cell will not be injured by being frozen but will be as good as new after being brought back to normal temperature. In extreme cold weather dry cells may not give more than half of their normal service. At a temperature of about −30°F they freeze solid and give neither voltage nor current.

The amperage of a dry cell by definition is the current that it will give when it is short-circuited (at about 70°F) through an ammeter which with its leads has a resistance of 0.01 Ω.

The **Ruben cell** (Ruben, Balanced Alkaline Dry Cells, *Trans. Electrochem. Soc.*, **92**, 1947) was developed jointly by the Ruben Laboratories and P. R. Mallory & Company during World War II for the operation of radar equipment and other electronic devices which require a high ratio of ampere-hour capacity to the volume of the cell at higher current densities than were considered practicable for the Leclanché type. The anode is of amalgamated zinc, and the cathode is a mercuric oxide depolarizing material intimately mixed with graphite in order to reduce its electrical resistivity. The electrolyte is a solution of potassium hydroxide (KOH) containing potassium zincate. The cell is made in three forms as shown in Fig. 15.1.13, the wound-anode type (*a*), the button type (*b*), and the cylindrical type (*c*).

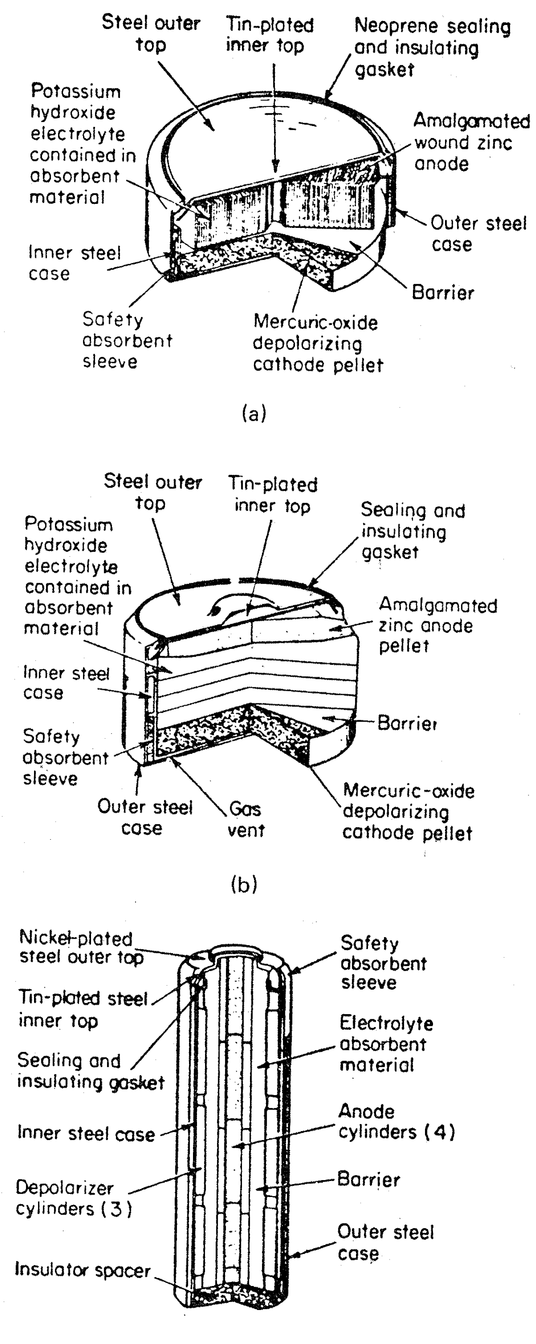


Fig. 15.1.13 Ruben cells. (*a*) Wound-anode; (*b*) button; (*c*) cylindrical. *(From "Standard Handbook for Electrical Engineers," Fink and Carrol, McGraw-Hill, NY, copyright 1968.)*

The no-load emf of the cell is 1.34 V and remains essentially constant irrespective of time and temperature. Advantages of the cell are long shelf life, which enables them to be stored indefinitely; long service life, about four times that of the Leclanché dry cell of equivalent volume; small weight; a flat voltage characteristic which is advantageous for electronic uses in which the characteristics of tubes vary widely with voltage; adaptability to operating at high temperatures without deterioration; high resistance to shock.

The **zinc-alkaline-manganese dioxide cell** is a recently developed cell especially useful in applications that require a dry cell with relatively heavy or continuous drain. The anode is of amalgamated zinc, and the cathode is a manganese dioxide depolarizing material mixed with graphite for conductivity. The electrolyte is a solution of highly alkaline potassium hydroxide immobilized in cellulosic-type separators. These cells are available in standard-size cylindrical construction and wafer (flat) construction for specialized applications.

Wet Cells The **silver** or **cuprous chloride-magnesium cell** is a one-shot battery with a life of days after the electrolyte is added. A wet cell may be stored for years in a dry state. The cathode is either compacted copper chloride and graphite or sheet silver chloride, while the cathode is a thin magnesium sheet. The electrolyte is a solution of sodium chloride. The silver chloride cells are more expensive and are available in more and larger ratings.

The **Weston cell** is a primary cell used as a standard of emf. It consists of a glass H tube in the bottom of one leg of which is mercury which forms the cathode; in the bottom of the other leg is cadmium amalgam forming the anode. The electrolytes consist of mercurous sulfate and cadmium sulfate. There are two forms of the Weston cell: the saturated or normal cell, and the unsaturated cell. In the normal cell the electrolyte is saturated. This is the official standard since it is more permanent than the unsaturated type and can be reproduced with far greater accuracy. When carefully made, the emfs of cells agree within a few parts per million. There is, however, a small temperature coefficient. Although the unsaturated cell is not so reliable as the normal cell and must be standardized, it has a negligible temperature coefficient and is more convenient for general use. The manufacturers recommend that the temperature be not less than 4°C and not more than 40°C and the current should not exceed 0.0001 A. The emf is between 1.0185 and 1.0190 V. Since no appreciable current can be taken from the cell, a null method must be used to utilize its emf.

Storage (Secondary) Batteries

In a **storage battery** the electrolytic action must be **reversible** to a high degree. There are three types of storage batteries; the lead-lead-acid type, the nickel-iron-alkaline type (**Edison** battery), and the nickel-cadmium-alkali type (**Nicad).** In addition, there are various specialized types of cells for scientific and military purposes, and there is continuous development work in the search for higher capacities.

In the manufacture of the **lead-lead-acid cells** there are three general types of plates, or electrodes. In the **Planté type** the active material is electrically formed of pure lead by repeated reversals of the charging current. In the **Faure,** or **pasted-plate,** type, the positive and negative plates are formed by applying a paste, largely of lead oxides (PbO_2, Pb_3O_4), to lead-antimony or lead calcium supporting grids. A current is passed through the plates while they are immersed in weak sulfuric acid, the positive plates being connected as anodes and the negative ones as cathodes. The paste on the positive plates is converted into lead peroxide while that on the negative plate is reduced to spongy lead. The **tubular plate** (iron-clad) type has lead-alloy rods surrounded by perforated dielectric tubes with powdered-lead oxides packed between the rod and tube for the positive plate.

In order to obtain high capacity per unit weight it is necessary to expose a large plate area to the action of the acid. This is done in the Planté plate by "ploughing" with sharp steel disks, and by using corrugated helical inserts as active positive material (Manchester plate). In the pasted plate a large area of the material is necessarily exposed to the action of the acid.

The chemical reactions in a lead cell may be expressed by the following equation, based on the double sulfation theory:

$$\underbrace{PbO_2}_{\text{positive plate}} + \underbrace{Pb}_{\text{negative plate}} + \underbrace{2H_2SO_4}_{\text{sulfuric acid}} \underset{\xrightarrow{\text{discharge}}}{\overset{\xleftarrow{\text{charge}}}{=}} \underbrace{2PbSO_4}_{\text{positive and negative plates}} + \underbrace{2H_2O}_{\text{water}}$$

Between the extremes of complete charge and discharge, complex combinations of lead and sulfate are formed. After complete discharge a hard insoluble sulfate forms slowly on the plates, and this is reducible only by slow charging. This sulfation is objectionable and should be avoided.

Specific Gravity Water is formed with discharge and sulfuric acid is formed on charge, consequently the specific gravity must decrease on discharge and increase on charge. The variation of the specific gravity for a stationary battery is shown in Fig. 15.1.14. With starting and vehicle batteries it is

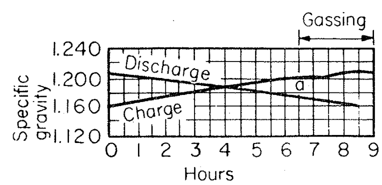


Fig. 15.1.14 Variations of specific gravity in a stationary battery.

necessary to operate the electrolyte from between 1.280 to 1.300 when fully charged to as low as 1.100 when completely discharged. The condition of charge of a battery can be determined by its specific gravity.

Battery electrolyte may be made from concentrated sulfuric acid (oil of vitriol, sp gr 1.84) by **pouring the acid into the water** in the following proportions:

Parts Water to 1 Part Acid

Specific gravity	1.200	1.210	1.240	1.280
Volume	4.3	4.0	3.4	2.75
Weight	2.4	2.2	1.9	1.5

Freezing Temperature of Sulfuric Acid

Specific gravity	1.180	1.200	1.240	1.280
Freezing temp, °F	−6	−16	−51	−90

Voltage The emf of a lead cell when fully charged and idle is 2.05 to 2.10 V. Discharge lowers the voltage in proportion to the current. When charging at constant current and normal rate, the terminal voltage gradually increases from 2.14 to 2.3 V, then increases rapidly to between 2.5 and 2.6 V (Fig. 15.1.15). This latter interval is known as the **gassing period.** When this period is reached, the charging rate should be reduced in order to avoid waste of power and unnecessary erosion of the plates.

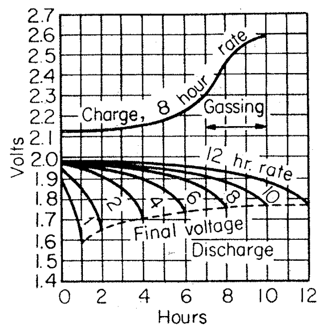


Fig. 15.1.15 Voltage curves on charge and discharge for lead cell.

Practically all batteries have a **normal rating** based on the 8-h rate of discharge. Thus a 320 A·h battery would have a normal rate of 40 A. The ampere-hour capacity of batteries falls off rapidly with increase in discharge rate.

Effect of Discharge Rate on Battery Capacity

Discharge rate, h	8	5	3	1	1/3	1/10
Percentage of rated capacity,						
Planté type	100	88	75	55.8	37	19.5
Pasted type	100	93	83	63	41	25.5

The following rule may be observed in **charging a lead battery.** The charging rate in amperes should be less than the number of ampere-hours out of the battery. For example, if 200 A·h are out of a battery, a charging rate of 200 A may be used until the ampere-hours out of the battery are reduced appreciably.

There are two common methods of charging: the **constant-current method** and the **constant-potential method.** Figure 15.1.16*a* shows a common method of charging with constant current, provided a low-voltage dc power supply is available. The resistor connected in series may be adjusted to give the required current. Several batteries may be connected in series. Figure 15.1.16*b* shows a more common method, using a copper oxide or silicon rectifier, since ac power supply is more common than dc. The rectifier disks, mounted in a stack, are bridge-connected, the directions of rectification being indicated. The polarity of the two wires can readily be determined by means of a dc voltmeter.

The constant-potential method is to be preferred since the rate automatically tapers off as the cell approaches the charged condition. Without resistance the terminal voltage should be 2.3 V per cell, but it is preferable to use 2.4 to 2.5 V per cell with low resistance in series.

When a battery is being charged, its terminal voltage

$$V = E + Ir \qquad (15.1.28)$$

Compare with Eq. **(15.1.27).**

When a battery is fully charged, any rate will produce gassing, but the rate may be reduced to such a low value that gassing is practically harmless. This is called the **finishing rate.**

Portable batteries for automobile starting and lighting, airplanes, industrial trucks, electric locomotives, train lighting, and power boats employ the pasted-type plates because of their high discharge rates for a given weight and size. The separators are either of treated grooved wood; perforated hard rubber; glass-wool mats; perforated rubber, and grooved wood; ribbed microporous rubber. In low-priced short-lived batteries for automobiles, grooved wood alone is used; in the better types, the wood is reinforced with perforated hard rubber. Containers for the low-priced short-lived automobile-type starting batteries are of asphaltic compound; for other portable types they are usually of hard rubber.

The **Exide iron-clad** battery is a portable type designed for propelling electric vehicles. The positive plate consists of a lead-antimony frame supporting perforated hard-rubber tubes. An irregular lead-antimony core runs down the center of each tube, and the lead peroxide paste is packed into these tubes so that shedding of active material from the positive plate cannot occur. Pasted negative plates are used. The separators are flat microporous rubber.

Stationary Batteries The tanks of stationary batteries are made of hard rubber or plastics. When the battery is used for regulating or cycling duty, the positive plates may be of the Planté type because of their long life. However, in most modern installations thick pasted plates are used. Because of the tight fit of the plate assembly within the container and the resulting pressure of the separator against the plate surfaces,

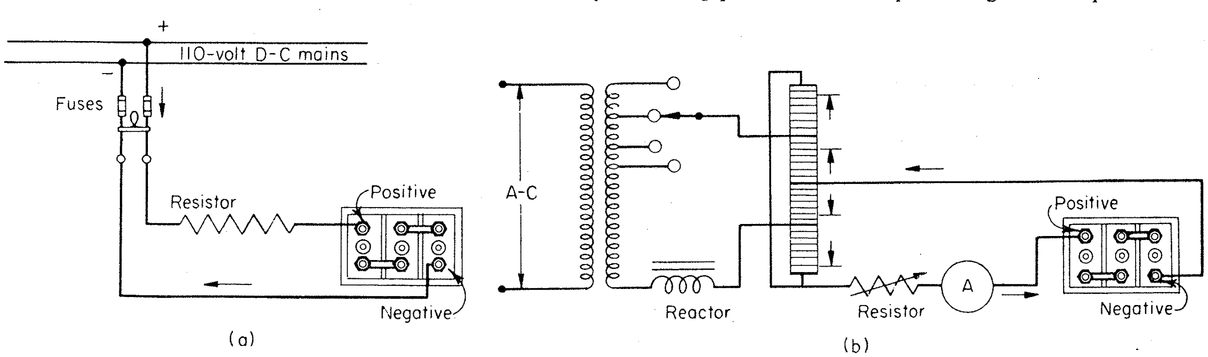


Fig. 15.1.16 Connections for charging storage battery from (*a*) 110-V dc mains, (*b*) copper-oxide rectifier.

shedding of active material is reduced to a minimum and long life is obtained. Pasted negative plates are used in almost all batteries.

A lead storage battery **removed from service** for less than 9 months should be charged once a month if possible; if not, it should be given a heavy overcharge before discontinuing service. If removed for a longer period, siphon off acid (which may be used again) and fill with fresh water. Allow to stand 15 h and siphon off water. Remove and throw away the wood separators. The battery will now stand indefinitely. To put in service again, install new separators, fill with acid (sp gr 1.210) and charge at normal rate 35 h or until gravity has ceased to rise over a period of 5 h. Charge at a low rate a few hours longer.

The **ampere-hour efficiency** of lead batteries is 85 to 90 percent. The **watthour efficiency** obtained from full charge to discharge at the normal rate and at rated amp-hour is 75 to 80 percent. Batteries which do regulating duty only may have a much higher watthour efficiency.

The **Edison storage cell** when fully charged has a positive plate of nickel pencils filled with a higher nickel oxide and a negative plate of flat nickel-plated-steel stampings containing metallic iron in finely divided form. The active material for the positive plate is nickel hydrate and for the negative plate, iron oxide. The electrolyte is a 21 percent solution of potassium hydrate with lithium hydroxides. The initial emf is about 1.4 V and the average emf about 1.1 V throughout discharge. In Fig. 15.1.17

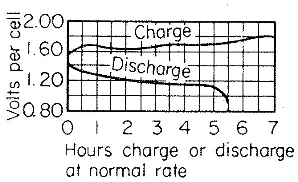


Fig. 15.1.17 Voltage during charge and discharge of Edison cell.

are shown typical voltage characteristics on charge and discharge for an Edison cell. On account of the higher internal resistance of the cell the battery is not so efficient from the energy standpoint as the lead cell. The jar is welded nickel-plated steel. The battery is compact and extremely light and strong and for these reasons is particularly adapted for propelling electric vehicles and for boat- and train-lighting systems. The battery is rugged, and since there is no opportunity for the growth of active material on the plates or flaking of active material, the battery has long life.

Nickel-Cadmium-Alkali (Nicad) Battery The positive active material is nickelic (black) hydroxide mixed with graphite to give it high conductivity. The negative active material is cadmium oxide. Both materials are used in powdered form and are contained within flat perforated steel pockets. These pockets are locked into steel plates, the positive and negative being alike in construction. All steel parts are nickel-plated. A complete plate group consists of a number of positive and negative plates assembled on bolts and terminal posts common to plates of the same polarity. The separators are thin strips of polystyrene, and all other battery insulation is also polystyrene. The entire plate assembly is contained within a welded-steel tank. The electrolyte is potassium hydroxide (KOH), specific gravity 1.210 at 72°F (22°C); it does not enter into any chemical reactions with the electrode materials, and its specific gravity remains constant during charge and discharge, neglecting any slight change due to the small amount of gassing. On charge, the voltage is 1.4 to 1.5 V until near the end when it rises to 1.8 V. On discharge, the voltage is nearly constant at 1.2 V.

Nicad batteries are strong mechanically and are not damaged by overcharge; they hold their charge over long periods of idleness, the active material cannot flake off, the internal resistance is low, there is no corrosion, and the battery has an indefinitely long life. It is a general-purpose battery.

In the **Sonotone** nickel-cadmium battery the positive plates are nickel oxide when the battery is charged, and the negative plates are metallic cadmium. On discharge the positive plates are reduced to a state of lower oxidation, and the negative plates regain oxygen. The electrolyte is a 30 percent solution of potassium hydroxide, the specific gravity of which is 1.29 at room temperature. The case is a transparent plastic. The terminal voltage at the normal discharge rate is 1.2 V per cell.

Rechargeable batteries, exemplified by Gould Nicad cells (Alkaline Battery Division, Gould National Batteries, Inc.), are hermetically sealed nickel-cadmium cells that contain no free alkaline electrolyte. Since there is no spillage or leakage, they can operate in any position, have long life, and require no maintenance or servicing, and their weight is small for their output. They are thus well adapted to power many types of cordless appliances such as tools, hedge shears, cameras, dictating equipment, electric razors, radios, and television sets. The electrodes consist of a plaque of microporous sintered nickel having an extremely high surface area. The electrochemical reactions differ from those of the conventional vented-type alkaline battery, a type which at the end of a charge liberates both oxygen and hydrogen gases as well as electrolytic fumes that must be vented through a valve in the top of the cell. In the sealed nickel-cadmium cell, the negative electrode (at the time that the cell is sealed) never becomes fully charged, and the evolution of hydrogen is completely suppressed. On charging, when the positive electrode has reached its full capacity, the oxygen which has evolved is channeled through the porous separator to the negative electrode and oxidizes the finely divided cadmium of the microporous plate to cadmium hydroxide, which at the same time is reduced to metallic cadmium. The cells are constructed in three different forms: the button type, the cylindrical type, and the prismatic type. Their ratings range from 20 mA·h to 23 A·h. Their average discharge voltage is 1.22 V, and they require 14 h of charge at the normal rate (one-tenth A·h rating), which for a 3.5 A·h cell is 0.35 A.

Precautions in the **care of storage batteries:** An ammeter should not be connected directly across the terminals to test the condition of a cell; a battery should not be left to stand in a discharged condition; a flame should not be brought in the vicinity of a battery that is being charged; the battery should not be allowed to become heated when charging; water should never be added to the concentrated acid—always acid to the water; acid should never be equalized except when the battery is in a charged condition; a battery should never be exposed to the influence of external heat; voltmeter tests should be made when the current is flowing; batteries should always be kept clean. To replace acid lost through slopping, use a solution of 2 parts concentrated sulfuric acid in 5 parts water by weight, unless a hydrometer is at hand to enable the solution to be made up according to the specifications of the makers of the cell.

DIELECTRIC CIRCUIT

Dynamic and Static Electricity Electricity in motion such as an electric current is dynamic electricity; electricity at rest is static electricity. The two are identical physically. Since static electricity is frequently produced at high voltage and small quantity, the two are frequently considered as being two different types of electricity.

Capacitors

Capacitors (formerly condensers) Two conducting bodies, or electrodes, separated by a dielectric constitute a capacitor. If a positive charge is placed on one electrode of a capacitor, an equal negative charge is induced on the other. The medium between the capacitor plates is called a **dielectric.** The dielectric properties of a medium relate to its ability to conduct **dielectric lines.** This is in distinction to its **insulating** properties which relate to its property to conduct **electric current.** For example, air is an excellent insulator but ruptures dielectrically at low voltage. It is not a good dielectric so far as breakdown strength is concerned.

With capacitors

$$Q = CE \qquad \textbf{(15.1.29)}$$
$$C = Q/E \qquad \textbf{(15.1.30)}$$
$$E = Q/C \qquad \textbf{(15.1.31)}$$

where Q = quantity, C; C = capacitance, F; and E = voltage. The unit of capacitance in the practical system is the **farad.** The farad is too large a unit for practical purposes, so that either the **microfarad** (μF) or the **picofarad** (pF) are used. However, in voltage, current, and energy relations the capacitance must be expressed in farads.

The energy stored in a capacitor

$$W = \tfrac{1}{2}QE = \tfrac{1}{2}CE^2 = \tfrac{1}{2}Q^2/C \quad \text{J} \qquad \textbf{(15.1.32)}$$

Capacitance of Capacitors The capacitance of a **parallel-electrode** capacitor (Fig. 15.1.18) is

$$C = \epsilon_r A/(4\pi d \times 9 \times 10^3) \quad \mu\text{F} \qquad \textbf{(15.1.33)}$$

where ϵ_r = relative capacitivity; A = area of one electrode, m^2; and d = distance between electrodes, m.

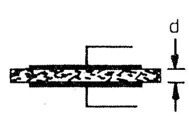


Fig. 15.1.18 Parallel-electrode capacitor.

The capacitance of **coaxial cylindrical** capacitors (Fig. 15.1.19) is

$$C = 0.2171\epsilon_r l \,/\, [9 \times 10^5 \log (R_2/R_1)] \quad \mu\text{F} \qquad \textbf{(15.1.34)}$$

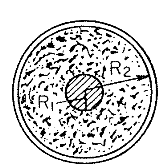


Fig. 15.1.19 Coaxial-cylinder capacitor.

where ϵ_r is the relative capacitivity and l the length, m. Also

$$C = 0.03882\epsilon_r/\log (R_2/R_1) \quad \mu\text{F/mi} \qquad \textbf{(15.1.35)}$$

Equation **(15.1.35)** is useful in that it is applicable to cables.

The capacitance of two **parallel cylindrical conductors** D m between centers and having radii of r m is

$$C = 0.01941/\log (D/r) \quad \mu\text{F/mi} \qquad \textbf{(15.1.36)}$$

In practice, the capacitance to neutral or to an infinite conducting plane midway between the conductors and perpendicular to their plane is usually used. The capacitance to neutral is

$$C = 0.03882/\log (D/r) \quad \mu\text{F/mi} \qquad \textbf{(15.1.37)}$$

Equations **(15.1.36)** and **(15.1.37)** are used for calculating the capacitance of overhead transmission lines. When computing charging current, use voltage between lines in **(15.1.36)** and to neutral in **(15.1.37).**

Capacitances in Parallel The equivalent capacitance of capacitances in parallel (Fig. 15.1.20)

$$C = C_1 + C_2 + C_3 \qquad \textbf{(15.1.38)}$$

Capacitances in parallel are all across the same voltage. If the voltage is E, then the total quantity $Q = CE$ and $Q_1 = C_1E$, etc.

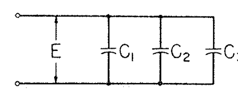


Fig. 15.1.20 Capacitances in parallel.

Capacitances in Series The equivalent capacitance C of capacitances in series (Fig. 15.1.21) is found as follows:

$$1/C = 1/C_1 + 1/C_2 + 1/C_3 \qquad \textbf{(15.1.39)}$$

If the capacitances are not leaky, the charge Q is the same on each. $Q = CE$, $E_1 = Q/C_1$, $E_2 = Q/C_2$, etc.

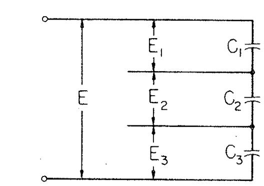


Fig. 15.1.21 Capacitances in series.

Insulators and Dielectrics Insulating materials are applied to electric circuits to prevent the leakage of current. Insulating materials used with high voltage must not only have a high resistance to leakage current, but must also be able to resist dielectric puncture; i.e., in addition to being a good insulator, the material must be a good dielectric. Insulation resistance is usually expressed in MΩ and the resistivity given in MΩ·cm. The dielectric strength is usually given in terms of voltage gradient, common units being V/mil, V/mm, and kV/cm. Insulation resistance decreases very rapidly with increase in temperature. Absorbed moisture reduces the insulation resistance, and moisture and humidity have a large effect on surface leak-

Table 15.1.8 Electrical Properties of Insulating Materials

Material	Volume resistivity, M Ω · cm	Dielectric constant, 60Hz	Dielectric strength	
			V/mil	V/mm
Asbestos board (ebonized)	10^7		55	2×10^3
Bakelite	$5\text{–}30 \times 10^{11}$	4.5–5.5	450–1400	$(17\text{–}55) \times 10^3$
Epoxy	10^{14}	3.5–5	300–400	$(12\text{–}16) \times 10^3$
Fluorocarbons:				
Fluorinated ethylene propylene	10^{18}	2.1	500	20×10^3
Polytetrafluoroethylene	10^{18}	2.1	400	16×10^3
Glass	17×10^9	5.4–9.9	760–3,800	$(3\text{–}15) \times 10^4$
Magnesium oxide		2.2	300–700	$(12\text{–}27) \times 10^3$
Mica	$10^{14} - 10^{17}$	4.5–7.5	1,000–4,000	$(4\text{–}16) \times 10^4$
Nylon	$10^{14} - 10^{17}$	4–7.6	300–400	$(12\text{–}16) \times 10^3$
Neoprene		7.5	600	23.5×10^3
Oils:				
Mineral	21×10^6	2–4.7	300–400	$(12\text{–}16) \times 10^3$
Paraffin	10^{15}	2.41	410–550	$(16\text{–}22) \times 10^3$
Paper		1.7–2.6	110–230	$(4\text{–}9) \times 10^3$
Paper, treated		2.5–4	500–750	$(20\text{–}30) \times 10^3$
Phenolic (glass filled)	$10^{12} - 10^{13}$	5–9	140–400	$(5.5\text{–}16) \times 10^3$
Polyethylene	$10^{15} - 10^{18}$	2.3	450–1,000	$(17\text{–}40) \times 10^3$
Polyimide	$10^{16} - 10^{17}$	3.5	400	16×10^3
Polyvinyl chloride (flexible)	$10^{11} - 10^{15}$	5–9	300–1,000	$(12\text{–}40) \times 10^3$
Porcelain	3×10^8	5.7–6.8	240–300	$(9.5\text{–}12) \times 10^3$
Rubber	$10^{14} - 10^{16}$	2–3.5	500–700	$(20\text{–}27) \times 10^3$
Rubber (butyl)	10^{18}	2.1		

age. In Table 15.1.8 are given the insulating and dielectric properties of several common insulating materials (see also Sec. 6). Dielectric heating of materials is described in Sec. 7.

TRANSIENTS

Induced EMF If a flux ϕ webers linking N turns of conductor changes, an emf

$$e = -N(d\phi/dt) \qquad \text{V} \qquad \textbf{(15.1.40)}$$

is induced.

Self-inductance Let a flux ϕ link N turns. The linkages of the circuit are $N\phi$ weber-turns. If the permeability of the circuit is assumed constant, the number of these linkages per ampere is the **self-inductance** or **inductance** of the circuit. The unit of inductance is the **henry.** The inductance is

$$L = N\phi/(i) \qquad \text{H} \qquad \textbf{(15.1.41)}$$

If the permeability changes with the current

$$L = N(d\phi/di) \qquad \text{H} \qquad \textbf{(15.1.42)}$$

The energy stored in the magnetic field

$$W = \tfrac{1}{2}Li^2 \qquad \text{J} \qquad \textbf{(15.1.43)}$$

EMF of Self-induction If Eq. **(15.1.41)** is written $Li = N\phi$ and differentiated with respect to the time t, $L(di/dt) = N(d\phi/dt)$ and from Eq. **(15.1.40)**

$$e = -L(di/dt) \qquad \text{V} \qquad \textbf{(15.1.44)}$$

e is the emf of self-induction. If a rate of change of current of 1 A/s induces an emf of 1 V, the inductance is then 1 H.

Current in Inductive Circuit If a circuit containing resistance R and inductance L in series is connected across a steady voltage E, the voltage E must supply the iR drop in the circuit and at the same time overcome the emf of self-induction. That is $E = Ri + L\,di/dt$. A solution of this differential equation gives

$$i = (E/R)\,(1 - \epsilon^{-Rt/L}) \qquad \text{A} \qquad \textbf{(15.1.45)}$$

where ϵ is the base of the natural system of logarithms.

Figure 15.1.22 shows this equation plotted when $E = 10$ V, $R = 20\ \Omega$, $L = 0.6$ H. It is to be noted that inductance causes

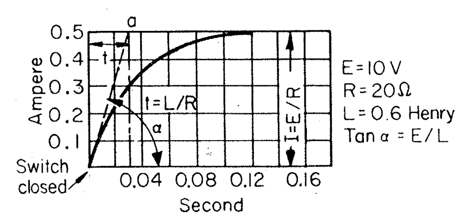


Fig. 15.1.22 Rise of current in inductive circuit.

the current to rise slowly to its Ohm's law value, $I_0 = E/R = {}^{10}\!/_{20} = 0.5$ A. When $t = L/R$, the current has reached 63.2 percent of its Ohm's law value. L/R is the **time constant** of the circuit. In the foregoing circuit, the time constant $L/R = 0.6/20 = 0.03$ s. The initial rate of rise of current is $\tan \alpha = E/L$. If current continued at this rate, it would reach $a = E/R$ in L/R s $[(E/L) \times (L/R) = E/R]$.

If a circuit containing inductance and resistance in series is short-circuited when the current is I_0, the equation of current becomes

$$i = I_0\epsilon^{-Rt/L} \qquad \text{A} \qquad \textbf{(15.1.46)}$$

Figure 15.1.23 shows this equation plotted when $I_0 = 0.5$ A, $R = 20\Omega$, $L = 0.6$ H. It is seen that inductance opposes the decay of current. Inductance always opposes change of current.

Mutual Inductance If two circuits having inductances L_1 and L_2 henrys are so related to each other geometrically that

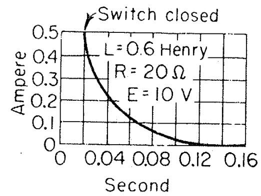


Fig. 15.1.23 Decay of current in inductive circuit.

any portion of the flux produced by the current in one circuit links the other circuit, the two circuits possess **mutual inductance.** It follows that a change of current in one circuit causes an emf to be induced in the other. Let ϵ_2 be induced in circuit 2 by a change di_1/dt in circuit 1. Then

$$e_2 = -M\, di_1/dt \qquad \text{V} \tag{15.1.47}$$

M is the mutual inductance of the two circuits.

$$M = k\sqrt{L_1 L_2} \tag{15.1.48}$$

where k is the **coefficient of coupling** of the two circuits, or the proportion of the flux in one circuit which links the other. Also a change of current di_2/dt in circuit 2 induces an emf e_1 in circuit 1, $e_1 = -M\, di_2/dt$.

The stored energy is

$$W = \tfrac{1}{2}L_1 I_1^2 + \tfrac{1}{2}L_2 I_2^2 + M I_1 I_2 \qquad \text{J} \tag{15.1.49}$$

where I_1 and I_2 are the currents in circuits 1 and 2.

Current in Capacitive Circuit If capacitance C farads and resistance R ohms are connected in series across the steady voltage E, the current is

$$i = (E/R)\epsilon^{-t/CR} \qquad \text{A} \tag{15.1.50}$$

If a capacitor charged to voltage E is discharged through resistance R, the current is

$$i = -(E/R)\epsilon^{-t/CR} \qquad \text{A} \tag{15.1.51}$$

Except for sign, these two equations are identical and are of the same form as Eq. (15.1.46).

In Fig. 15.1.24 is shown the transient current to a capacitor in series with a resistor when $E = 200$ V, $C = 4.0\ \mu$F, $R = 2$ kΩ. When $t = CR$, the current has reached $1/\epsilon = 0.368$ its initial value. CR is the **time constant** of the circuit. The initial rate of decrease of current is $\tan\alpha = -E/CR^2$. If the current continued at this rate it would reach zero when the time is CR s. If, in its fully charged condition, the capacitor of Fig. 15.1.24 is discharged through the resistor R, the curve will be the negative of that shown in Fig. 15.1.24.

Resistance, Inductance, and Capacitance in Series If a circuit having resistance, inductance, and capacitance in series is connected across a source of steady voltage, a transient condition results. If $R > \sqrt{4L/C}$, the circuit is nonoscillatory or overdamped.

The current is

$$i = \frac{EC}{\sqrt{R^2C^2 - 4LC}}\left(\epsilon^{(-\alpha+\beta)t} - \epsilon^{(-\alpha-\beta)t}\right) \qquad \text{A} \tag{15.1.52}$$

where $\alpha = R/2L$ and $\beta = (\sqrt{R^2C^2 - 4LC})/2LC$.

In Fig. 15.1.25 is shown the curve corresponding to Eq. (15.1.52). When $R = \sqrt{4L/C}$, the system is **critically damped** and the transient dies out rapidly without oscillation. The current is

$$i = (E/L)t\epsilon^{-Rt/2L} \qquad \text{A} \tag{15.1.53}$$

Figure 15.1.25 shows also the curve corresponding to Eq. (15.1.53).

If $R < \sqrt{4L/C}$, the transient is oscillatory, being a logarithmically damped sine wave. The current is

$$i = \frac{2EC}{\sqrt{4LC - R^2C^2}}\epsilon^{-Rt/2L}\sin\frac{\sqrt{4LC - R^2C^2}}{2LC}t \qquad \text{A} \tag{15.1.54}$$

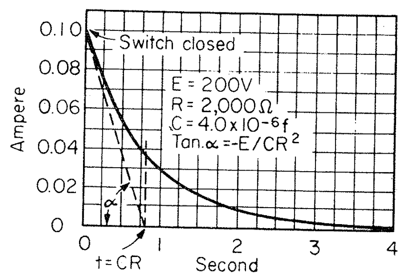


Fig. 15.1.24 Transient current to capacitor.

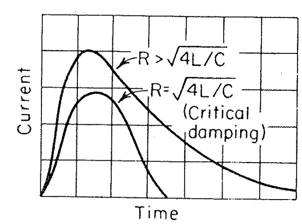


Fig. 15.1.25 Transient current in nonoscillatory circuits.

The transient oscillates at a frequency very nearly equal to $1/(2\pi\sqrt{LC})$ Hz. This is the **natural frequency** of the circuit.

In Fig. 15.1.26 is shown the curve corresponding to Eq. (15.1.54). If the capacitor, after being charged to E V, is discharged into the foregoing series circuits, the currents are

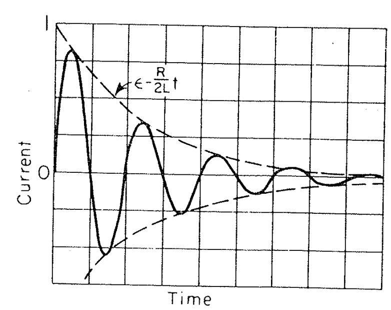


Fig. 15.1.26 Transient current in oscillatory circuit.

given by Eqs. **(15.1.52)** to **(15.1.54)** multiplied by -1. Equations **(15.1.52)** to **(15.1.54)** are the same types obtained with dynamic mechanical systems with friction, mass, and elasticity.

ALTERNATING CURRENTS

Sine Waves In the following discussion of alternating currents, sine waves of voltage and current will be assumed. That is, $e = E_m \sin \omega t$ and $i = I_m \sin (\omega t - \theta)$, where E_m and I_m are maximum values of voltage and current; ω, the angular velocity, in rad/s, is equal to $2\pi f$, where f is the frequency; θ is the angle of phase difference.

Cycle; Frequency When any given armature coil has passed a pair of poles, the emf or current has gone through 360 electrical degrees, or 1 **cycle.** An **alternation** is one-half a cycle. The **frequency** of a synchronous machine in cycles per second (hertz) is

$$f = NP/120 \qquad \text{Hz} \tag{15.1.55}$$

where N is the speed in r/min and P the number of poles. In the United States and Canada the frequency of 60 Hz is almost universal for general lighting and power. For the ac power supply to dc transit systems, and for railroad electrification, a frequency of 25 Hz is used in many installations. In most of Europe and Latin America the frequency of 50 Hz is in general use. In aircraft the frequency of 400 Hz has become standard.

Static inverters make it possible to obtain high and variable frequencies to drive motors at greater than the 3,600 r/min limitation on 60-Hz circuits, and to vary speeds. The textile industry has small motors operating at 12,000 r/min (200 Hz), and larger motors have been run at 6,000 r/min (100 Hz).

The **root-mean-square (rms),** or **effective, value of a current wave** produces the same heating in a given resistance as a direct current of the same ampere value. Since the heating effect of a current is proportional to i^2r, the rms value is obtained by squaring the ordinates, finding their average value, and extracting the square root, i.e., the rms value is

$$I = \sqrt{1/T \int_0^T i^2\, dt} \qquad \text{A} \tag{15.1.56}$$

where T is the time of a cycle. The rms value I of a sine wave equals $(1/\sqrt{2})\, I_m = 0.707 I_m$.

Average Value of a Wave The average value of a sine wave over a complete cycle is zero. For a half cycle the average is $(2/\pi)I_m$, or $0.637\, I_m$, where I_m is the maximum value of the sine wave. The average value is of importance only occasionally. A dc measuring instrument gives the average value of a pulsating wave. The average value is of use (1) when the effects of the current are proportional to the number of coulombs, as in electrolytic work and (2) when converting alternating to direct current.

Form Factor The form factor of a wave is the ratio of rms value to average value. For a sine wave this is $\pi/(2\sqrt{2}) = 1.11$. This factor is important in that it enters equations for induced emf.

Inductive reactance, $2\pi fL$ or ωL, opposes an alternating current in inductance L. It is expressed in Ω. Reactance is usually denoted by the symbol X. Inductive reactance is denoted by X_L.

The **current** in an inductive reactance X_L when connected across the voltage E is

$$I = E/X_L = E/(2\pi fL) \qquad \text{A} \tag{15.1.57}$$

This current lags the voltage by 90 electrical degrees. Inductance absorbs no energy. The energy stored in the magnetic field during each half cycle is returned to the source during the same half cycle.

Capacitive reactance is $1/(2\pi fC) = 1/\omega C$ and is denoted by X_C, where C is in F. If C is given in μF, $X_C = 10^6/2\pi fC$. The current in a capacitive reactance X_C when connected across voltage E is

$$I = E/X_C = 2\pi fCE \qquad \text{A} \tag{15.1.58}$$

This current leads the voltage by 90 electrical degrees. Pure capacitance absorbs no energy. The energy stored in the dielectric field during each half cycle is returned to the source during the same half cycle.

Impedance opposes the flow of alternating current and is expressed in Ω. It is denoted by Z. With resistance and inductance in series

$$Z = \sqrt{R^2 + X_L^2} = \sqrt{R^2 + (2\pi fL)^2} \qquad \Omega \tag{15.1.59}$$

With resistance and capacitance in series

$$Z = \sqrt{R^2 + X_C^2} = \sqrt{R^2 + [1/(2\pi fC)]^2} \qquad \Omega \tag{15.1.60}$$

With resistance, inductance, and capacitance in series

$$Z = \sqrt{R^2 + (X_L - X_C)^2} = \sqrt{R^2 + [2\pi fL - 1/(2\pi fC)]^2} \qquad \Omega \tag{15.1.61}$$

The current is

$$I = E/\sqrt{R^2 + [2\pi fL - 1/(2\pi fC)]^2} \qquad \text{A} \tag{15.1.62}$$

Phasor or Vector Representation Sine waves of voltage and current can be represented by phasors, these phasors being proportional in magnitude to the waves that they represent. The angle between two phasors is also equal to the time angle existing between the two waves that they represent.

Phasors may be combined as forces are combined in mechanics. Both graphical methods and the methods of complex algebra are used. Impedances and also admittances may be similarly combined, either graphically or symbolically. The usual method is to resolve series impedances into their component resistances and reactances, then combine all resistances and all reactances, from which the resultant impedance is obtained. Thus $Z_1 + Z_2 = \sqrt{(r_1 + r_2)^2 + (x_1 + x_2)^2}$, where r_1 and x_1 are the components of Z_1, etc.

Phase Difference With resistance only in the circuit, the current and the voltage are in phase with each other; with inductance only in the circuit, the current lags the voltage by 90 electrical degrees; with capacitance only in the circuit, the current leads the voltage by 90 electrical degrees.

With resistance and inductance in series, the voltage leads the current by angle θ where $\tan \theta = X_L/R$. With resistance and capacitance in series, the voltage lags the current by angle θ where $\tan \theta = -X_C/R$.

With resistance, inductance, and capacitance in series, the voltage may lag, lead, or be in phase with the current.

$$\tan \theta = (X_L - X_C)/R = (2\pi fL - 1/2\pi fC)/R \tag{15.1.63}$$

If $X_L > X_C$ the voltage leads; if $X_L < X_C$ the voltage lags; if $X_L = X_C$ the current and voltage are in phase and the circuit is in resonance.

Power Factor In ac circuits the power $P = I^2R$ where I is the current and R the effective resistance (see below). Also the power

$$P = EI \cos \theta \qquad \text{W} \tag{15.1.64}$$

where θ is the phase angle between E and I. Cos θ is the **power factor** (pf) of the circuit. It can never exceed unity and is usually less than unity.

$$\cos \theta = P/EI \tag{15.1.65}$$

P is often called the true power. The product EI is the volt-amp (V·A) and is often called the apparent power.

Active or **energy** current is the projection of the total current on the voltage phasor. $I_e = I \cos \theta$. Power $= EI_e$.

Reactive, quadrature, or **wattless current** $I_q = I \sin \theta$ and is the component of the current that contributes no power but increases the I^2R losses of the system. In power systems it should ordinarily be made low.

The **vars** (volt-amp-reactive) are equal to the product of the voltage and reactive current. **Vars** $= EI_q$. **Kilovars** $= EI_q/1{,}000$.

Effective Resistance When alternating current flows in a circuit, the losses are ordinarily greater than are given by the losses in the ohmic resistance alone. For example, alternating current tends to flow near the surface of conductors (skin effect). If iron is associated with the circuit, eddy-current and hysteresis losses result. These power losses may be accounted for by increasing the ohmic resistance to a value R, where R is the **effective resistance,** $R = P/I^2$. Since the iron losses vary as $I^{1.8}$ to I^2, little error results from this assumption.

SOLUTION OF SERIES-CIRCUIT PROBLEM. Let a resistor R of 10 Ω, an inductor L of 0.06 H, and a capacitor C of 60 μF be connected in series across 120-V 60-Hz mains (Fig. 15.1.27). Determine (1) the impedance, (2) the current, (3) the voltage across the resistance, the inductance, the capacitance, (4) the power factor, (5) the power, (6) the angle of phase difference.

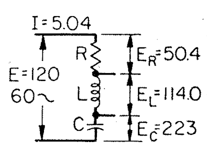


Fig. 15.1.27 Resistor, inductor, and capacitor in series.

(1) $\omega = 2\pi 60 = 377$. $X_L = 0.06 \times 377 = 22.6\Omega$; $X_C = 1/(377 \times 0.000060) = 44.2\ \Omega$; $Z = \sqrt{(10)^2 + (22.6 - 44.2)^2} = 23.8\ \Omega$; (2) $I = 120/23.8 = 5.04$ A; (3) $E_R = IR = 5.04 \times 10 = 50.4$ V; $E_L = IX_L = 5.04 \times 22.6 = 114.0$ V; $E_C = IX_C = 5.04 \times 44.2 = 223$ V; (4) $\tan \theta = (X_L - X_C)/R = -21.6/10 = -2.16$, $\theta = -65.2°$, $\cos \theta = \text{pf} = 0.420$; (5) $P = 120 \times 5.04 \times 0.420 = 254$ W; $P = I^2R = (5.04)^2 \times 10 = 254$ W (check); (6) From (4) $\theta = -65.2°$. Voltage lags. The phasor diagram to scale of this circuit is shown in Fig. 15.1.28.

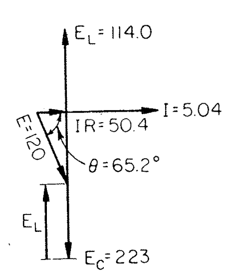


Fig. 15.1.28 Phasor diagram for series circuit.

Since the current is common for all elements of the circuit, its phasor is laid horizontally along the axis of reference.

Resonance If the voltage E and the resistance R [Eq. (15.1.62)] are fixed, the maximum value of current occurs when $2\pi fL - 1/2\pi fC = 0$. The circuit so far as its terminals are concerned behaves like a noninductive resistor. The current $I = E/R$, the power $P = EI$, and the power factor is unity.

The voltage across the inductor and the voltage across the capacitor are opposite and equal and may be many times greater than the circuit voltage. The frequency

$$f = 1/(2\pi \sqrt{LC}) \qquad \text{Hz} \tag{15.1.66}$$

is the **natural frequency** of the circuit and is the frequency at which it will oscillate if the circuit is not acted upon by some external frequency. This is the principle of radio sending and receiving circuits. Resonant conditions of this type should be avoided in power circuits, as the piling up of voltage may endanger apparatus and insulation.

EXAMPLE. For what value of the inductance in the circuit (Fig. 15.1.27) will the circuit be in resonance, and what is the voltage across the inductor and capacitor under these conditions?

From Eq. (15.1.66) $L = 1/(2\pi f)^2C = 0.1173$ H. $I = E/R = 120/10 = 12$ A. $L\omega I = I/C\omega = 0.1173 \times 377 \times 12 = 530$ V. This voltage is over four times the line voltage.

Parallel Circuits

Parallel circuits are used for nearly all power distribution. With several series circuits in parallel it is merely necessary to find the current in each and add all the current phasors vectorially to find the total current. Parallel circuits may be solved analytically.

A series circuit has resistance r_1 and **inductive reactance** x_1. The **conductance** is

$$g_1 = r_1/(r_1^2 + x_1^2) = r_1/Z_1^2 \qquad \text{S} \tag{15.1.67}$$

and the **susceptance** is

$$b_1 = x_1/(r_1^2 + x_1^2) = x_1/Z_1^2 \qquad \text{S} \tag{15.1.68}$$

Conductance is not the reciprocal of resistance unless the reactance is zero; susceptance is not the reciprocal of reactance unless the resistance is zero. With inductive reactance the susceptance is **negative;** with capacitive reactance the susceptance is **positive.**

If a second circuit has resistance r_2 and **capacitive reactance** x_2 in series, $g_2 = r_2/(r_2^2 + x_2^2) = r_2/Z_2^2$; $b_2 = x_2/(r_2^2 + x_2^2) = x_2/Z_2^2$. The total conductance $G = g_1 + g_2$; the total susceptance $B = -b_1 + b_2$. The **admittance** is

$$Y = \sqrt{G^2 + B^2} = 1/Z \qquad \text{S} \tag{15.1.69}$$

The energy current is EG; the reactive current is EB; the power is

$$P = E^2G \qquad \text{W} \tag{15.1.70}$$

$$\text{vars} = E^2B \qquad \text{W} \tag{15.1.71}$$

The power factor is

$$pf = G/Y \tag{15.1.72}$$

Also the following relations hold:

$$r = g/(g^2 + b^2) = g/Y^2 \qquad \Omega \tag{15.1.73}$$

$$x = b/(g^2 + b^2) = b/Y^2 \qquad \Omega \tag{15.1.74}$$

SOLUTION OF A PARALLEL-CIRCUIT PROBLEM. In the parallel circuit of Fig. 15.1.29 it is desired to find the joint impedance, the total current,

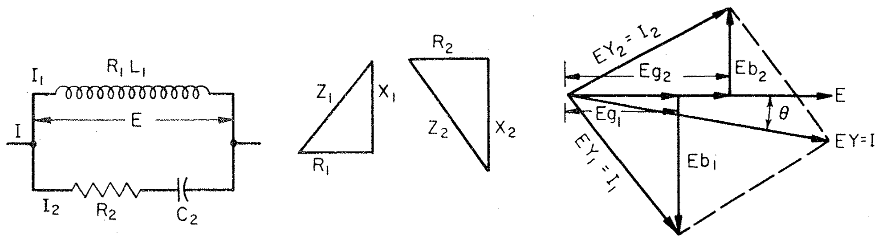


Fig. 15.1.29 Parallel circuit and phasor diagram.

the power in each branch, the total power, and the power factor, when $E = 100, f = 60, R_1 = 2\ \Omega, R_2 = 4\ \Omega, L_1 = 0.00795$ H, $X_1 = 2\pi f L_1 = 3\ \Omega$, $C_2 = 1{,}326\ \mu$F, $X_2 = 1/2\pi f C_2 = 2\ \Omega$, $Z_1 = \sqrt{2^2 + 3^2} = 3.6\ \Omega$, and $Y_1 = 1/3.6 = 0.278$ S.

Solution: $g_1 = R_1/(R_1^2 + X_1^2) = 2/13 = 0.154$; $b_1 = -3/13 = -0.231$; $Z_2 = \sqrt{16 + 4} = 4.47$; $Y_2 = 1/4.47 = 0.224$; $g_2 = R_2/(R_2^2 + X_2^2) = 4/(16 + 4) = 0.2$ S; $b_2 = 2/20 = 0.1$ S; $G = g_1 + g_2 = 0.154 + 0.2 = 0.354$ S; $B = b_1 + b_2 = -0.231 + 0.1 = -0.131$ S; $Y = \sqrt{G^2 + B^2} = \sqrt{0.354^2 + (-0.131)^2} = 0.377$ S, and joint impedance $Z = 1/0.377 = 2.65\ \Omega$. Phase angle $\theta = \tan^{-1}(-0.131/0.354) = -20.3°$. $I = EY = 100 \times 0.377 = 37.7$ A; $P_1 = E^2 g_1 = 100^2 \times 0.154 = 1{,}540$ W; $P_2 = E^2 g_2 = 100^2 \times 0.2 = 2{,}000$ W; total power $= E^2 G = 100^2 \times 0.354 = 3{,}540$ W. Power factor $= \cos\theta = 3{,}540/(100 \times 37.7) = 93.8$ percent.

With parallel circuits, unity power factor is obtained when the algebraic sum of the quadrature currents is zero. That is, $b_1 + b_2 + b_3 \cdots = 0$.

Three-Phase Circuits Ac generators are usually wound with three armature circuits which are spaced 120 electrical degrees apart on the armature. Hence these coils generate emfs 120 electrical degrees apart. The coils are connected either in *Y* (star) or in Δ (mesh) as shown in Fig. 15.1.30. Whether Y- or

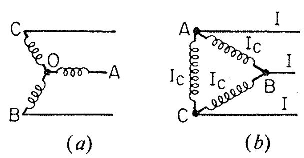


Fig. 15.1.30 Three-phase connections. (*a*) Y connection; (*b*) Δ connection.

Δ-connected, with a balanced load, the three coil emfs E_c and the three coil currents I_c are equal. In the Y connection the line and coil currents are equal, but the line emfs E_{AB}, E_{BC}, E_{CA} are $\sqrt{3}$ times in magnitude the coil emfs E_{OA}, E_{OB}, E_{OC}, since each is the phasor difference of two coil emfs. In the delta connection the line and coil emfs are equal, but *I*, the line current, is $\sqrt{3}\ I_c$, the coil current, i.e., it is the phasor difference of the currents in the two coils connected to the line. The power of a coil is $E_c I_c \cos\theta$, so that the total power is $3E_c I_c \cos\theta$. If θ is the angle between coil current and coil voltage, the angle between line current and line voltage will be $30° \pm \theta$. In terms of line current and emf, the power is $\sqrt{3}\ EI \cos\theta$. A fourth or netural conductor connected to *O* is frequently used with the Y connection. The neutral point *O* is frequently grounded in transmission and distribution circuits. The coil emfs are assumed to be sine waves. Under these conditions they balance, so that in the delta connection the sum of the two coil emfs at each instant is balanced by the third coil emf. Even though the third, ninth, fifteenth . . . harmonics, $3(2n + 1)f$, where $n = 0$ or an integer, exist in the coil emfs, they cannot appear between the three external line conductors of the three-phase Y-connected circuit. In the delta circuit, the same harmonics $3(2n + 1)f$ cause local currents to circulate around the mesh. This may cause a very appreciable heating. In a three-phase system the power

$$P = \sqrt{3}\ EI \cos\theta \qquad \text{W} \tag{15.1.75}$$

the power factor is

$$P/\sqrt{3}\ EI \tag{15.1.76}$$

and the kV·A

$$\sqrt{3}\ EI/1{,}000 \tag{15.1.77}$$

where *E* and *I* are line voltages and currents.

Two-Phase Circuits Two-phase generators have two windings spaced 90 electrical degrees apart on the armature. These windings generate emfs differing in time phase by 90°. The two windings may be independent and power transmitted to the receiver though the two single-phase circuits are entirely insulated from each other. The two circuits may be combined into a two-phase three-wire circuit such as is shown in Fig. 15.1.31, where *OA* and *OB* are the generator circuits (or transformer secondaries) and *A′O′* and *B′O′* are the load circuits. The wire *OO′* is the common wire and under balanced conditions carries a current $\sqrt{2}$ times the current wires *AA′* and *BB′*. For example, if I_c is the coil current, $\sqrt{2}\ I_c$ will be the value of the current in the common conductor *OO′*. If E_c is the voltage across *OA* or *OB*, $\sqrt{2}\ E_c$ will be the voltage across *AB*. The power of a two-phase circuit is twice the power in either coil if the load is balanced. Normally, the voltages *OA* and *OB* are equal, and the current is the same in both coils. Owing to nonsymmetry and the high degree of unbalancing of this system even under balanced loads, it is not used at the present time for transmission and is little used for distribution.

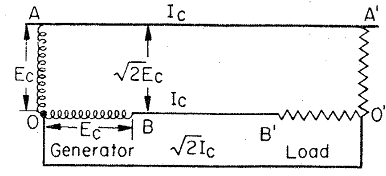


Fig. 15.1.31 Two-phase, three-wire circuit.

Four-Phase Circuit A four-phase or quarter-phase circuit is shown in Fig. 15.1.32. The windings *AC* and *BD* may be independent or connected at *O*. The voltages *AC* and *BD* are 90 electrical degrees apart as in two-phase circuits. If a neutral wire *O-O′* is added, three different voltages can be obtained. Let E_1 = voltage between *O-A, O-B, O-C, O-D*. Voltages between A-B, B-C, C-D, D-A = $\sqrt{2}\,E_1$. Voltages between A-C, B-D = $2E_1$. Because of this multiplicity of voltages and the fact that polyphase power apparatus and lamps may be connected at the same time, this system is still used to some extent in distribution.

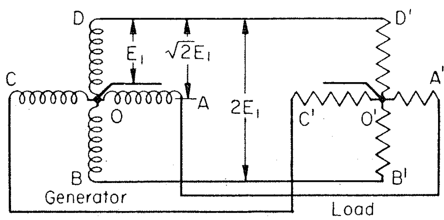


Fig. 15.1.32 Four-phase or quarter-phase circuit.

Advantages of Polyphase Power The advantages of polyphase power over single-phase power are as follows. The output of synchronous generators and most other rotating machinery is from 60 to 90 percent greater when operated polyphase than when operated single phase; pulsating fluxes and corresponding iron losses which occur in many common types of machinery when operated single phase are negligible when operated polyphase; with balanced polyphase loads polyphase power is constant whereas with single phase the power fluctuates over wide limits during the cycle. Because of its minimum number of wires and the fact that it is not easily unbalanced, the three-phase system has for the most part superseded other polyphase systems.

ELECTRICAL INSTRUMENTS AND MEASUREMENTS

Electrical measuring devices that merely indicate, such as ammeters and voltmeters, are called **instruments;** devices that totalize with time such as watthour meters and ampere-hour meters are called **meters.** (See also Sec. 16.) Most types of electrical instruments are available with digital read out.

DC Instruments Direct current and voltage are both measured with an indicating instrument based on the principle of the D'Arsonval galvanometer. A coil with steel pivots and turning in jewel bearings is mounted in a magnetic field produced by permanent magnets. The motion is restrained by two small flat coiled springs, which also serve to conduct the current to the coil. The deflections of the coil are read with a light aluminum pointer attached to the coil and moving over a graduated scale. The same instrument may be used for either current or voltage, but the method of connecting in circuit is different in the two cases. Usually, however, the coil of an instrument to be used as an ammeter is wound with fewer turns of coarser wire than an instrument to be used as a voltmeter and so has lower resistance. The instrument itself is frequently called a **millivoltmeter.** It cannot be used alone to measure voltage of any magnitude since its resistance is so low that it would be burned out if connected across the line. Hence a resistance r' in series with the coil is necessary as indicated in Fig. 15.1.33*a* in which r_c is the resistance of the coil. From 0.2 to 750 V this resistance is usually within the instrument. For higher voltages an external resistance R, called an extension coil or multiplier (Fig. 15.1.33*b*), is necessary. Let e be the reading of the instrument, in volts (Fig. 15.1.33*b*), r the internal resistance of the instrument, including r' and r_c in (15.1.33*a*), R the resistance of the multiplier. Then the total voltage is

$$E = e(R + r)/r \tag{15.1.78}$$

It is clear that by using suitable values of R a voltmeter can be made to have several scales.

Instruments themselves can only carry currents of the magnitudes of 0.01 to 0.06 A. To measure larger values of current the instrument is provided with a shunt R (Fig. 15.1.34). The current divides inversely as the resistances r and R of the instrument and the shunt. A low resistance r' within the instrument is connected in series with the coil. This permits some adjustment to the deflection so that the instrument can be adapted to its shunt. Usually most of the current flows through the shunt, and the current in the instrument is negligible in comparison. Up to 50 and 75 A the shunt can be incorporated within the instrument. For larger currents it is usually necessary to have the shunt external to the instrument and connect the instrument to the potential terminals of the shunt by means of leads. Any given instrument may have any number of ranges

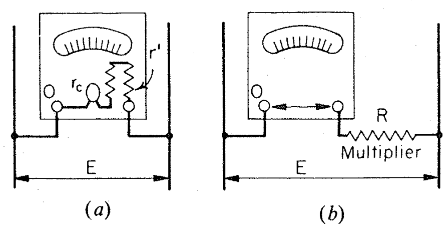


Fig. 15.1.33 Voltmeters. (*a*) Internal resistance; (*b*) with multiplier.

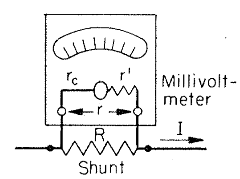


Fig. 15.1.34 Millivoltmeter with shunt.

by providing it with a sufficient number of shunts. The range of the usual instrument of this type is approximately 50 mV. Although the same instrument may be used for voltmeters or ammeters, the moving coils of voltmeters are usually wound with more turns of finer wire. They take approximately 0.01 A so that their resistance is approximately 100 Ω/V. Instruments used as ammeters alone operate with 0.01 to 0.06 A.

Permanent-magnet moving-coil instruments may be used to measure unidirectional pulsating currents or voltages and in such cases will indicate the average value of the periodically varying current or voltage.

AC Instruments Instruments generally used for alternating currents may be divided into five types: **electrodynamometer, iron-vane, thermocouple, rectifier,** and **electronic.** Instruments of the **electrodynamometer** type, the most precise, operate on the principle of one coil carrying current, turning in the magnetic field produced by a second coil carrying current taken from the same circuit. If these circuits or coils are connected in series, the torque exerted on the moving system for a given relative position of the coil system is proportional to the square of the current and is not dependent on the direction of the current. Consequently, the instrument will have a compressed scale at the lower end and will usually have only the upper two-thirds of the scale range useful for accurate measurement. Instruments of this type ordinarily require 0.04 to 0.08 A or more in the moving-coil circuit for full-scale deflection. They read the rms value of the alternating or pulsating current. The wattmeter operates on the electrodynamometer principle. The fixed coil, however, is energized by the current of the circuit, and the moving coil is connected across the potential in series with high resistance. Unless shielded magnetically the foregoing instruments will not, in general, indicate so accurately on direct as on alternating current because of the effects of external stray magnetic fields. Also reversed readings should be taken. **Iron-vane** instruments consist of a fixed coil which actuates magnetically a light movable iron vane mounted on a spindle; they are rugged, inexpensive, and may be had in ranges of 30 to 750 V and 0.05 to 100 A. They measure rms values and tend to have compressed scales as in the case of electrodynamometer instruments.

The compressed part of the scale may, however, be extended by changing the shape of the vanes. Such instruments operate with direct current and are accurate to within 1 percent or so. AC instruments of the **induction type** (Westinghouse Electric Corp.) must be used on ac circuits of the frequency for which they have been designed. They are rugged and relatively inexpensive and are used principally for switchboards where a long-scale range and a strong deflecting torque are of particular advantage. **Thermocouple instruments** operate on the **Seebeck effect.** The current to be measured is conducted through a heater wire, and a thermojunction is either in thermal contact with the heater or is very close to it. The emf developed in the thermojunction is measured by a permanent-magnet dc type of instrument. By controlling the shape of the air gap, a nearly uniform scale is obtained. This type of instrument is well adapted to the measurement of high-frequency currents or voltages, and since it operates on the heating effect of current, it is convenient as a transfer instrument between direct current and alternating current.

In the **rectifier-type instrument** the ac voltage or current is rectified, usually by means of a small copper oxide or a selenium-type rectifier, connected in a bridge circuit to give full-wave rectification (Fig. 15.1.35). The rectified current is measured with a dc permanent-magnet-type instrument M. The instrument measures the **average** value of the half waves that have been rectified, and with the sine waves, the average value is 0.9 the rms value. The scale is calibrated to indicate rms values.

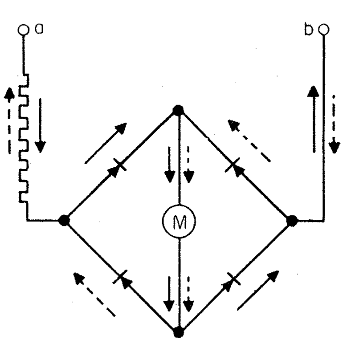


Fig. 15.1.35 Rectifier-type instrument.

With nonsinusoidal waves the ratio of average to rms may vary considerably from 0.9 so that the instrument may be in error up to ± 5 percent from this cause. This type of instrument is widely used in the measurement of high-frequency voltages and currents. **Electronic (vacuum-tube) voltmeters** operate on the principle of the amplification which can be obtained with a triode, or three-element vacuum tube. Since the emf to be measured is applied to the grid, the instruments take practically no current and hence are adapted to measure potential differences which would change radically were any appreciable current taken by the measuring device. This type of instrument can measure voltages from a few tenths of a volt to several hundred volts, and with a potential divider, up to thousands of volts. They are also adapted to frequencies up to 100 MHz.

Particular care must be used in selecting instruments for measuring the nonsinosoidal waves of rectifier and controlled rectifier circuits. The **electrodynamic, iron vane,** and **thermocouple** instruments will read rms values. The **rectifier** instrument will read average values, while the **electronic** instrument may read either rms or average value, depending on the type.

Power Measurement in Single-Phase Circuits Wattmeters are not rated primarily in W, but in A and V. For example, with a low power factor the current and voltage coils may be overloaded and yet the needle be well on the scale. The current coil may be carrying several times its rated current, and yet the instrument reads zero because the potential circuit is not closed, etc. Hence it is desirable to use both an ammeter and a voltmeter in conjunction with a wattmeter when measuring power (Fig. 15.1.36*a*). The instruments themselves consume

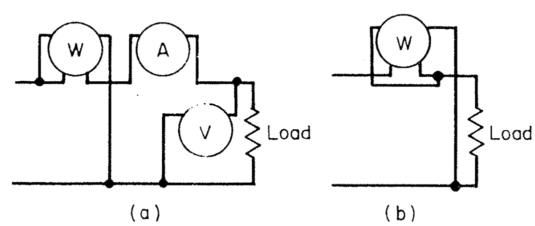


Fig. 15.1.36 Connections of instruments to single-phase load.

appreciable power, and correction is often necessary unless these losses are negligible compared with the power being measured. For example, in Fig. 15.1.36*a*, the wattmeter measures the I^2R loss in its own current coil and in the ammeter (1 to 2 W each), as well as the loss in the voltmeter ($= E^2/R$ where R is the resistance of the voltmeter). The losses in the ammeter and voltmeter may be eliminated by short-circuiting the ammeter and disconnecting the voltmeter when reading the wattmeter. If the wattmeter is connected as shown in Fig. 15.1.36*b*, it measures the power taken by its own potential coil (E^2/R_p) which at 110 V is 5 to 7 W. (R_p is the resistance of the potential circuit.) Frequently correction must be made for this power.

Power Measurement in Polyphase Circuits; Three-Wattmeter Method Let *ao, bo,* and *co* be any Y-connected three-phase load (Fig. 15.1.37). Three wattmeters with their current coils in each line and their potential circuits connected to neutral measure the total power, since the power in each load is measured by one of the wattmeters. The connection *oo′* may, however, be broken, and the total power is still the sum of the three readings; i.e., the power $P = P_1 + P_2 + P_3$. This method is applicable to any system of n wires. The current coil of one wattmeter is connected in each of the n wires. The potential circuit of each wattmeter is connected between its own phase wire and a junction in common with all the other potential circuits. The wattmeters must be connected symmetrically, and the readings of any that read negative must be given the negative sign.

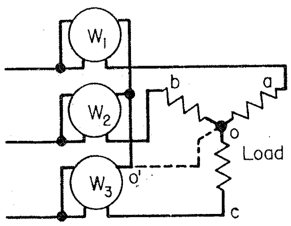


Fig. 15.1.37 Three-wattmeter method.

In the general case any system of n wires requires at least $n - 1$ wattmeters to measure the power correctly. The $n - 1$ wattmeters are connected in series with $n - 1$ wires. The potential circuit of each is connected between its own phase wire and the wire in which no wattmeter is connected (Fig. 15.1.38).

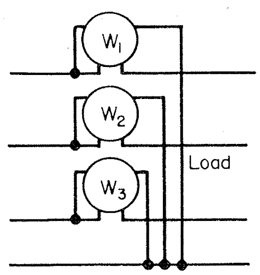


Fig. 15.1.38 Power measurement in n-wire system.

The thermal watt converter is also used to measure power. This instrument produces a dc voltage proportional to three-phase ac power.

Three-Phase Systems The three-wattmeter method (Fig. 15.1.37) is applicable to any three-phase system. It is commonly used with the three-phase four-wire system. If the loads are balanced, $P_1 = P_2 = P_3$ and the power $P = 3P_1$.

The **two-wattmeter** method is most commonly used with three-phase three-wire systems (Fig. 15.1.39). The current coils may be connected in *any* two wires, the potential circuits being connected to the third. It will be recognized that this is adapting the method of Fig. 15.1.38 to three wires. With balanced loads the readings of the wattmeters are $P_1 = Ei\cos(30° + \theta)$, $P_2 = Ei\cos(30° - \theta)$, and $P = P_2 \pm P_1$. θ is the angle of phase difference between coil voltage and current. Since

$$P_1/P_2 = \cos(30° + \theta)/\cos(30° - \theta) \qquad \textbf{(15.1.79)}$$

the power factor is a function of P_1/P_2. Table 15.1.9 gives values of power factor for different ratios of P_1/P_2.

$P = P_2 + P_1$ when $\theta < 60°$.

When $\theta = 60°$, pf $= \cos 60° = 0.5$, $P_1 = \cos(30° + 60°) = 0$, $P = P_2$. When $\theta > 60°$, pf < 0.5, $P = P_2 - P_1$. Also,

$$\tan\theta = \sqrt{3}\,(P_2 - P_1)/(P_2 + P_1) \qquad \textbf{(15.1.80)}$$

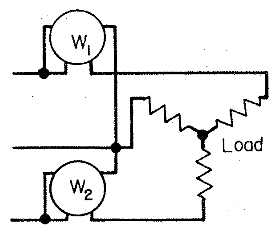


Fig. 15.1.39 Two-wattmeter method.

In a **polyphase wattmeter** the two single-phase wattmeter elements are combined to act on a single spindle. Hence the adding and subtracting of the individual readings are done automatically. The total power is indicated on one scale. This type of instrument is almost always used on switchboards. The connections of a portable type are shown in Fig. 15.1.40.

In the foregoing instrument connections, Y-connected loads are shown. These methods are equally applicable to delta-connected loads. The two-wattmeter method (Fig. 15.1.39) is obviously adapted to the two-phase three-wire system (Fig. 15.1.31).

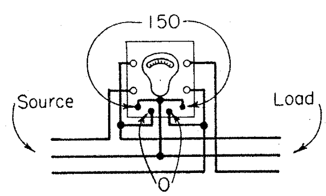


Fig. 15.1.40 Connections for polyphase wattmeter in three-phase circuit.

Measurement of Energy

Watthour meters record the energy taken by a circuit over some interval of time. Correct registration occurs if the angular velocity of the rotating element at every instant is proportional

Table 15.1.9 Ratio P_1/P_2 and Power Factor

P_1/P_2	Power factor	P_1/P_2	Power factor	P_1/P_2	Power factor	P_1/P_2	Power factor
+1.0	1.000	+0.4	0.804	−0.1	0.427	−0.6	0.142
+0.9	0.996	+0.3	0.732	−0.2	0.360	−0.7	0.102
+0.8	0.982	+0.2	0.656	−0.3	0.296	−0.8	0.064
+0.7	0.956	+0.1	0.576	−0.4	0.240	−0.9	0.030
+0.6	0.918	0.0	0.50	−0.5	0.188	−1.0	0.000
+0.5	0.866						

to the power. The method of accomplishing this with dc meters is illustrated in Fig. 15.1.41. The meter is in reality a small motor. The field coils *FF* are in series with the line. The armature *A* is connected across the line, usually in series with a resistor *R*. The movable field coil *F'* is in series with the armature *A* and serves to compensate for friction. *C* is a small commutator, either of copper or of silver, and the two small brushes are usually of silver. An aluminum disk, rotating between the

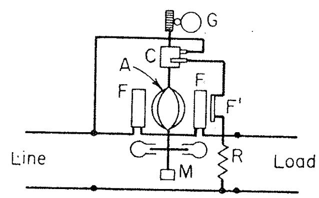


Fig. 15.1.41 DC watthour meter.

poles of permanent magnets *M*, acts as a magnetic brake the retarding torque of which is proportional to the angular velocity of the disk. A small worm and the gears *G* actuate the recording dials.

The following relation, or an equivalent, holds with most types of meter. With each revolution of the disk, *K* Wh are recorded, where *K* is the **meter constant** found usually on the disk. It follows that the average watts *P* over any period of time *t* sec is

$$P = 3{,}600KN/t \tag{15.1.81}$$

where *N* is the revolutions of the disk during that period. Hence, the meter may be calibrated by connecting standardized instruments to measure the average power taken by the load and by counting the revolutions *N* for *t* s. Near full load, if the meter registers fast, the magnets *M* should be moved outward radially; if it registers slow, the magnets should be moved inward. If the meter registers fast at light (5 to 10 percent) load, the starting coil *F'* should be moved further away from the armature; if it registers slow, *F'* should be moved nearer the armature. A meter should not register more than 1.5 percent fast or slow, and with calibrated standards it can be made to register to within 1 percent of correct.

The **induction watthour meter** is used with alternating current. Although the dc meter registers correctly with alternating current, it is more expensive than the induction type, the commutator and brushes may cause trouble, and at low power factors compensation is necessary. In the induction watthour meter the driving torque is developed in the aluminum disk by the joint action of the alternating magnetic flux produced by the potential circuit and by the load current. The driving torque and the retarding torque are both developed in the same aluminum disk, hence no commutator and brushes are necessary. The rotating element is very light, and hence the friction torque is small. Equation (15.1.81) applies to this type of meter. When calibrating, the average power *W* for *t* s is determined with a calibrated wattmeter. The friction compensation is made at light loads by changing the position of a small hollow stamping with respect to the potential lug. The meter should also be adjusted at low power factor (0.5 is customary). If the meter is slow with lagging current, resistance should be cut out of the compensating circuit; if slow with leading current, resistance should be inserted.

Power-Factor Measurement The usual method of determining power factor is by the use of voltmeter, ammeter, and wattmeter. The wattmeter gives the watts of the circuit, and the product of the voltmeter reading and the ammeter reading gives the volt-amperes. The power factor is the ratio of the two [see Eqs. (15.1.65) and (15.1.76)]. Also single-phase and three-phase power-factor indicators, which can be connected directly in circuit, are on the market.

Instrument Transformers

With voltages higher than 600 V, and even at 600 V, it becomes dangerous and inaccurate to connect instruments and meters directly into power lines. It is also difficult to make potential instruments for voltages in excess of 600 V and ammeters in excess of 60-A ratings. To insulate such instruments from high voltage and at the same time to permit the use of low-range instruments, instrument transformers are used. **Potential transformers** are identical with power transformers except that their volt-ampere rating is low, being 40 to 500 W. Their primaries are wound for line voltage and their secondaries for 110 V. **Current transformers** are designed to go in series with the line, and the rated secondary current is 5 A. The secondary of a current transformer **should always be closed** when current is flowing; it should never be allowed to become open circuited under these conditions. When open-circuited the voltage across the secondary becomes so high as to be dangerous and the flux becomes so large in magnitude that the transformer overheats. Semiconductors that break down at safe voltages and short current-transformer secondaries are available to ensure that the secondary is closed. The secondaries of both potential and current transformers should be well grounded at one point (Figs. 15.1.42 and 15.1.43). Instrument transformers introduce slight errors because of small variations

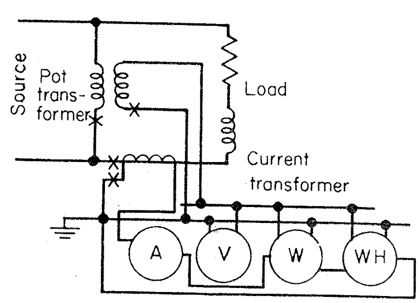


Fig. 15.1.42 Single-phase connections of instruments with transformers.

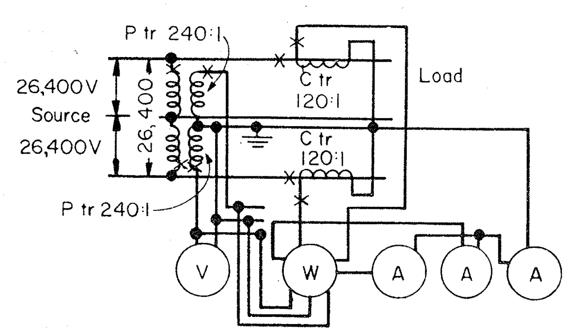


Fig. 15.1.43 Three-phase connections of instruments and instrument transformers.

in their ratio with load. Also there is slight phase displacement in both current and potential transformers. The readings of the instruments must be multiplied by the instrument transformer ratios. The scales of switchboard instruments are usually calibrated to take these ratios into account.

Figure 15.1.42 shows the use of instrument transformers to measure the voltage, current, power, and kilowatthours of a single-phase load. Figure 15.1.43 shows the connections that would be used to measure the voltage, current, and power of a 26,400-V 600-A three-phase load.

Measurement of High Voltages Potential transformers such as those shown in Figs. 15.1.42 and 15.1.43 may be used even for very high voltages, but for voltages above 132 kV they become so large and expensive that they are used only sparingly. A convenient method used with testing transformers is the employment of a voltmeter coil, which consists of a coil of a few turns interwoven in the high-voltage winding and insulated from it. The voltage ratio is the ratio of the turns in the high-voltage winding to those in the voltmeter coil. A capacitance voltage divider consists of two or more capacitors connected in series across the high voltage to be measured. A high-impedance voltmeter, such as an electronic one, is connected across the capacitor at the grounded end. The high voltage $V = V_m C_m / C$ V, where V_m is the voltmeter reading, C the capacitance (in μF) of the entire divider, and C_m the equivalent capacitance (in μF) of the capacitor at the grounded end. A bushing potential device consists of a high-voltage-transformer bushing having a capacitance tap brought out from one of the metallic electrodes within the bushing which is near ground potential. This device is obviously a capacitance voltage divider. For testing, **sphere gaps** are used for the very high voltages. Calibration data for sphere gaps are given in the ANSI/IEEE Std. 4-1978 Standard Techniques for High Voltage Testing. Even when it is not being used for the measurement of voltage, it is frequently advisable to connect a sphere gap in parallel with the specimen being tested so as to prevent overvoltages. The gap is set to a slightly higher voltage than that which is desired.

Measurement of Resistance

Voltmeter-Ammeter Method A common method of measuring resistance, known as the voltmeter-ammeter or fall-in-potential method, makes use of an ammeter and a voltmeter. In Fig. 15.1.44, the resistance to be measured is R. The current in the resistor R is I A, which is measured by the ammeter A in series. The drop in potential across the resistor R is measured by the voltmeter V. The current shunted by the voltmeter is so small that it may generally be neglected. A correction may be applied if necessary, for the resistance of the voltmeter is generally given with the instrument. The potential difference divided by the current gives the resistance included between

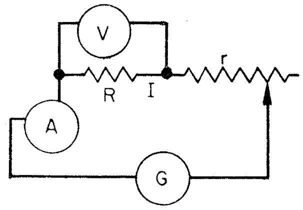


Fig. 15.1.44 Voltmeter-ammeter method for resistance measurement.

the voltmeter leads. As a check, determinations are generally made with several values of current, which may be varied by means of the controlling resistor r. If the resistance to be measured is that of the armature of a dc machine and the voltmeter leads are placed on the brush holders, the resistance determined will include that of the brush contacts. To measure the resistance of the armature alone, the voltmeter leads should be placed directly on the commutator segments on which the brushes rest but not under the brushes.

Insulation Resistance Insulation resistance is so high that it is usually given in megohms (10^6 ohms, MΩ) rather than in ohms. Insulation resistance tests are important, for although they may not be conclusive they frequently reveal flaws in insulation, poor insulating material, presence of moisture, etc. Such tests are applied to the insulation of electrical machinery from the windings to the frame, to underground cables, to insulators, capacitors, etc.

For moderately low resistances, 1 to 10 MΩ, the voltmeter method given in Fig. 15.1.45, which shows insulation measure-

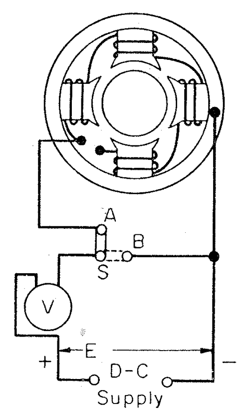


Fig. 15.1.45 Voltmeter method for insulation resistance measurement.

ment to the frame of the field winding of a generator, may be used. To measure the current when a voltage E is impressed across the resistor R, a high-reading voltmeter V is connected in series with R. The current under this condition with the switch connecting S and A is $E/(R + r)$, where r is the resistance of the voltmeter. A high-resistance voltmeter is necessary, since the method is in reality a comparison of the unknown insulation resistance R with the known resistance r of the voltmeter. Hence, the resistance of the voltmeter must be comparable with the unknown resistance, or the deflection of the instrument will be so small that the results will be inaccurate. To determine the impressed voltage E, the same voltmeter is used. The switch S connects S and B for this purpose. With these two readings, the unknown resistance is

$$R = r(E - e)/e \qquad (15.1.82)$$

where e is the deflection of the voltmeter when in series with the resistance to be measured as when S is at A. If a special voltmeter, having a resistance of 100 kΩ per 150 V, is available, a resistance of the order of 2 to 3 MΩ may be measured very accurately.

When the insulation resistance is too high to be measured with a voltmeter, a sensitive galvanometer may be used. The connections for measuring the insulation resistance of a cable are shown in Fig. 15.1.46. The battery should have an emf of at least 100 V. Radio B batteries are convenient for this purpose. The method involves comparing the unknown resistance with a standard 0.1 MΩ. To calibrate the galvanometer the

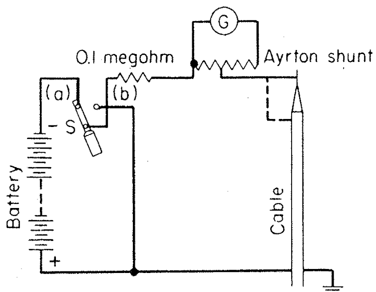


Fig. 15.1.46 Measurement of insulation resistance with galvanometer.

cable is short-circuited (dotted line) and the switch S is thrown to position *(a)*. Let the galvanometer deflection be D_1 and the reading of the Ayrton shunt S_1. The short circuit is then removed. The 0.1 MΩ is left in circuit since it is usually negligible in comparison with the unknown resistance X. Let the reading of the galvanometer now be D_2 and the reading of the shunt S_2. Then

$$X = 0.1\ S_2D_1/S_1D_2 \qquad \text{M}\Omega \tag{15.1.83}$$

When the switch S is thrown to position *(b)*, the cable is short-circuited through the 0.1 MΩ and becomes discharged.

The **Megger** insulation tester is an instrument that indicates insulation resistance directly on a scale. It consists of a small hand or motor-driven generator which generates 500 V, 1,000 V, 2,500 V, or 5,000 V. A clutch slips when the voltage exceeds the rated value. The current through the unknown resistance flows through a moving element consisting of two coils fastened rigidly together, but which move in different portions of the magnetic field. A pointer attached to the spindle of the moving element indicates the insulation resistance directly. These instruments have a range up to 10,000 MΩ and are very convenient where portability and convenience are desirable.

The insulation resistance of electrical machinery may be of doubtful significance as far as dielectric strength is concerned. It varies widely with temperature, humidity, and cleanliness of the parts. When the insulation resistance falls below the prescribed value, it can (in most cases of good design) be brought to the required standard by cleaning and drying the machine. Hence it may be useful in determining whether or not the insulation is in proper condition for a dielectric test. IEEE Std. 62-1978 specifies minimum values of insulation resistance in MΩ = (rated voltage)/(rating in kW + 1,000). If the operating voltage is higher than the rated voltage, the operating voltage should be used. The rule specifies that a dc voltage of 500 be used in testing. If not, the voltage should be specified.

Wheatstone Bridge Resistors from a fraction of an Ω to 100 kΩ and more may be measured with a high degree of precision with the Wheatstone bridge (Fig. 15.1.47). The bridge consists of four resistors $ABCX$ connected as shown. X is the unknown resistance; A and B are ratio arms, the resistance units of which are in even decimal Ω as 1, 10, 100, etc. C is the rheostat arm. A battery or low-voltage source of direct current is connected across ab. A galvanometer G of moderate sensitivity is connected across cd. The values of A and B are so chosen that three or four significant figures in the value of C are obtained. As a first approximation it is well to make A and B equal. When the bridge is in balance,

$$X/C = A/B \tag{15.1.84}$$

The positions of the battery and galvanometer are interchangeable. There are many modifications of the bridge which adapt it to measurements of very low resistances and also to ac measurements.

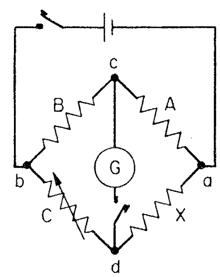


Fig. 15.1.47 Wheatstone bridge.

Kelvin Double Bridge The simple Wheatstone bridge is not adapted to measuring very low resistances since the contact resistances of the test specimen become comparable with the specimen resistance. This error is avoided in the Kelvin double bridge, the diagram of which is shown in Fig. 15.1.48. The specimen X, which may be a short length of copper wire or bus bar, is connected in series with an adjustable calibrated resistor R whose resistance is comparable with that of the specimen.

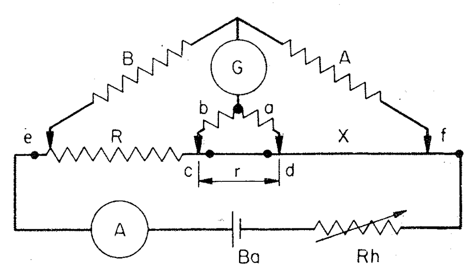


Fig. 15.1.48 Kelvin double bridge.

The arms A and B of the bridge are ratio arms usually with decimal values of 1, 10, 100 Ω. One terminal of the galvanometer is connected to X and R by means of two resistors a and b. If these resistors are set so that $a/b = A/B$, the contact resistance r between X and R is eliminated in the measurement. The contact resistances at c and d have no effect since at balance the galvanometer current is zero. The contact resistances at f and e need only be negligible compared with the resistances of arms A and B both of which are reasonably high. By means of the variable resistor Rh the value of current, as indicated by ammeter A, may be adjusted to give the necessary sensitivity. When the bridge is in balance,

$$X/R = A/B \tag{15.1.85}$$

Potentiometer The principle of the potentiometer is shown in Fig. 15.1.49. *ab* is a slide wire, and *bc* consists of a number of equal individual resistors between contacts. A battery *Ba* the emf of which is approximately 2 V supplies current to this wire through the adjustable rheostat *R*. A slider *m* makes contact with *ab*, and a contactor *m′* connects with the contacts in

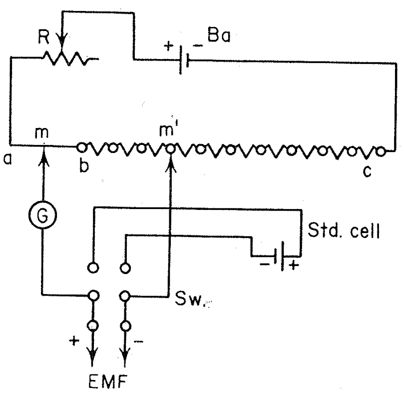


Fig. 15.1.49 Potentiometer principle.

bc. A galvanometer *G* is in series with the wire connecting to *m*. By means of the double-throw double-pole switch *Sw*, either the standard cell or the unknown emf (*EMF*) may be connected to *mm′* through the galvanometer *G*. The potentiometer is standardized by throwing *Sw* to the standard-cell side, setting *mm′* so that their positions on *ab* and *bc* correspond to the emf of the standard cell. The rheostat *R* is then adjusted until *G* reads zero. (In commercial potentiometers a dial which may be set directly to the emf of the standard cell is usually provided.) The unknown emf is measured by throwing *Sw* to *EMF* and adjusting *m* and *m′* until *G* reads zero. The advantage of this method of measuring emf is that when the potentiometer is in balance no current is taken from either the standard cell or the source of emf. Potentiometers seldom exceed 1.6 V in range. To measure voltage in excess of this, a **volt box** which acts as a multiplier is used. To measure current, the voltage drop across a standard resistor of suitable value is measured with the potentiometer. For example, with 50 A a 0.01-Ω standard resistance gives a voltage drop of 0.5 V which is well within the range of the potentiometer.

Potentiometers of low range are used extensively with thermocouple pyrometers. Figure 15.1.49 merely illustrates the principle of the potentiometer. There are many modifications, conveniences, etc., not shown in Fig. 15.1.49.

DC GENERATORS

All electrical machines are comprised of a magnetic circuit of iron (or steel) and an electric circuit of copper. In a generator the armature conductors are rotated so that they cut the magnetic flux coming from and entering the field poles. In the dc generator (except the unipolar type) the emf induced in the individual conductors is alternating, but this is rectified by the commutator and brushes, so that the current to the external circuit is unidirectional.

The induced emf in a generator (or motor)

$$E = \phi ZNP/60P'10^8 \quad \text{V} \tag{15.1.86}$$

where ϕ = flux in webers entering the armature from one north pole; Z = total number of conductors on the armature; N = speed, r/min; P = number of poles; and P' = number of parallel paths through the armature.

Since with a given generator, Z, P, P' are fixed, the induced emf

$$E = K\phi N \quad \text{V} \tag{15.1.87}$$

where K is a constant. When the armature delivers current, the terminal volts are

$$V = E - I_a R_a \tag{15.1.88}$$

where I_a is the armature current and R_a the armature resistance including the brush and contact resistance, which vary somewhat.

There are three standard types of dc generators: the **shunt generator,** the **series generator,** and the **compound generator.**

Shunt Generator The field of the shunt generator in series with its rheostat is connected directly across the armature as shown in Fig. 15.1.50. This machine maintains approximately constant terminal voltage over its working range of load. An external characteristic of the generator is shown in Fig. 15.1.51. As load is applied the terminal voltage drops owing to the armature-resistance drop [Eq. (15.1.88)] and armature reac-

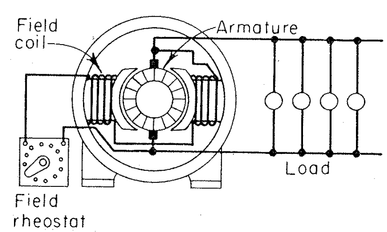


Fig. 15.1.50 Shunt generator.

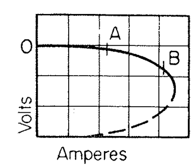


Fig. 15.1.51 Shunt generator characteristic.

tion which decreases the flux. The drop in terminal voltage reduces the field current which in turn reduces the flux, hence the induced emf, etc. At some point *B*, usually well above rated current, the foregoing reactions become cumulative and the generator starts to break down. The current reaches a maximum value and then decreases to nearly zero at short circuit. With large machines, point *B* is well above rated current, the operating range being between *O* and *A*. The voltage may be maintained constant by means of the field rheostat. Automatic regulators which operate through field resistance are frequently used to maintain constant voltage.

Shunt generators are used in systems which are all tied together where their stability when in parallel is an advantage. If a generator fails to build up, (1) the load may be connected; (2) the field resistance may be too high; (3) the field circuit may be open; (4) the residual magnetism may be insufficient; (5) the field connection may be reversed.

Series Generator In the series generator (Fig. 15.1.52) the entire load current flows through the field winding, which consists of relatively few turns of wire of sufficient size to carry the entire load current without undue heating. The field excitation, and hence the terminal voltage, depends on the magnitude of the load current. The generator supplies an essentially constant current and for years was used to supply series arc lamps for street lighting requiring direct current. Except for some special applications, the series generator is now obsolete.

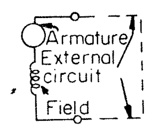


Fig. 15.1.52 Series generator.

Compound-wound Generators By the addition of a series winding to a shunt generator the terminal voltage may be automatically maintained very nearly constant, or, by properly proportioning the series turns, the terminal voltage may be made to increase with load to compensate for loss of voltage in the line, so that approximately constant voltage is maintained at the load. If the shunt field is connected outside the series field (Fig. 15.1.53), the machine is **long shunt;** if the shunt field is connected inside the series field, i.e., directly to the armature terminals, it is **short shunt.** So far as the operating characteristic is concerned, it makes little difference which way a machine is connected.

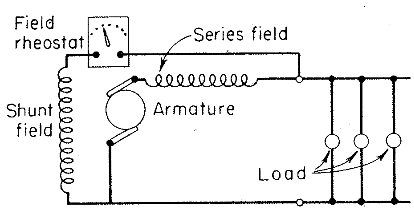


Fig. 15.1.53 Compound-wound dc generator.

Compound-wound generators are chiefly used for small isolated plants and for generators supplying a purely motor load subject to rapid fluctuations such as in railway work. When first putting a compound generator in service, the shunt field must be so connected that the machine builds up. The series field is then connected so that it aids the shunt field. Figure 15.1.54 gives the characteristics of an overcompounded 200-kW 600-V compound-wound generator.

Amplidynes The amplidyne is a dc generator in which a small amount of power supplied to a control field controls the generator output, the response being nearly proportional to the

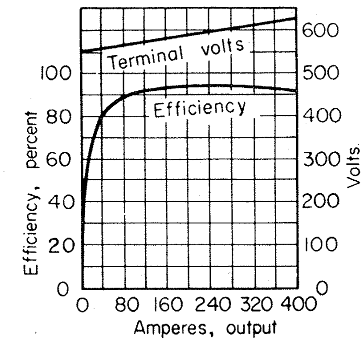


Fig. 15.1.54 Characteristics of a 200-kW, compound-wound, dc generator.

control field input. The amplidyne is a dc amplifier which can supply large amounts of power. The amplifier operates on the principle of armature reaction. In Fig. 15.1.55, NN and SS are the conventional north and south poles of a dc generator with

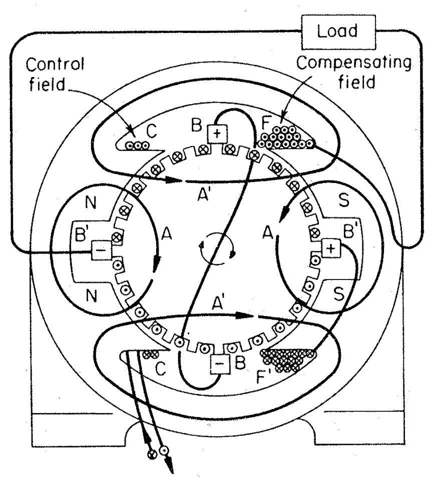


Fig. 15.1.55 Amplidyne.

Table 15.1.10 Approximate Test Performance of Compound-wound DC Generators with Commutating Poles

kW	Rpm	Volts	Amperes	Efficiencies, percent		
				1/4 load	1/2 load	Full load
5	1,750	125	40	77.0	80.5	82.0
10	1,750	125	80	80.0	83.0	85.0
25	1,750	125	200	84.0	86.5	88.0
50	1,750	125	400	83.0	86.0	88.0
100	1,750	125	800	87.0	88.5	90.0
200	1,750	125	1,600	88.0	90.5	91.0
400	1,750	250	1,600	91.7	91.9	91.7
1,000	1,750	250	4,000	92.1	92.6	92.1

SOURCE: Westinghouse Electric Corp.

central cavities. BB are the usual brushes placed at right angles to the pole axes of NN and SS. A control winding CC of small rating, as low as 100 W, is wound on the field poles. In Fig. 15.1.55, for simplicity, the control winding is shown as being wound on one pole only. The brushes BB are short-circuited, so that a small excitation mmf in the control field produces a large short-circuit current along the brush axis BB. This large short-circuit current produces a large armature-reaction flux AA along brush axis BB. The armature rotating in this field produces a large voltage along the brush axis B′B′. The load or working current is taken from brushes B′B′ as shown. In Fig. 15.1.55 the working current only is shown by the crosses and dots in the circles. The short-circuit current would be shown by crosses in the conductors to the left of brushes BB and by dots in the conductors to the right of brushes BB.

A small current in the control winding produces a high output voltage and current as a result of the large short-circuit current in brushes BB.

In order that the brushes B′B′ shall not be short-circuiting conductors which are cutting the flux of poles NN and SS, cavities are cut in these poles. Also the load current from brushes B′B′ produces an armature reaction mmf in opposition to flux A′A′ produced by the control field CC. Were this mmf not compensated, the flux A′A′ and the output of the machine would no longer be determined entirely by the control field. Hence there is a compensating field FF′ in series with the armature, which neutralizes the armature-reaction mmf which the load current produces. For simplicity the compensating field is shown on one field pole only.

The amplidyne is capable of controlling and regulating speed, voltage, current, and power with accurate and rapid response. The amplification is from 10,000 to 250,000 times in machines rated from 1 to 50 kW. Amplidynes are frequently used in connection with **selsyns** and are employed for gun and turret control and for accurate controls in many industrial power applications.

Parallel Operation of Shunt Generators It is desirable to operate generators in parallel so that the station capacity can be adpated to the load. Shunt generators, because of their drooping characteristics (Fig. 15.1.51), are inherently stable when in parallel. To connect shunt generators in parallel it is necessary that the switches be so connected that like poles are connected to the same bus bars when the switches are closed. Assume one generator to be in operation; to connect another generator in parallel with it, the incoming generator is first brought up to speed and its terminal voltage adjusted to a value slightly greater than the bus-bar voltage. This generator may then be connected in parallel with the other without difficulty. The proper division of load between them is adjusted by means of the field rheostats and is maintained automatically if the machines have similar voltage-regulation characteristics.

Parallel Operation of Compound Generators As a rule, compound generators have either flat or rising voltage characteristics. Therefore, when connected in parallel, they are inherently unstable. Stability may, however, be obtained by using an equalizer connection, Fig. 15.1.56, which connects the terminals of the generator at the junctions of the series fields. This connection is of low resistance so that any increase of current divides proportionately between the series fields of the two machines. The equalizer switch (E.S.) should be closed first and opened last, if possible. In practice, the equalizer switch is often one blade of a three-pole switch, the other two being the bus switch S, as in Fig. 15.1.56. When compound generators are used on a three-wire system, two series fields—one at each armature terminal—and two equalizers are necessary. It is possible to operate any number of compound generators in parallel provided their characteristics are not too different and the equalizer connection is used.

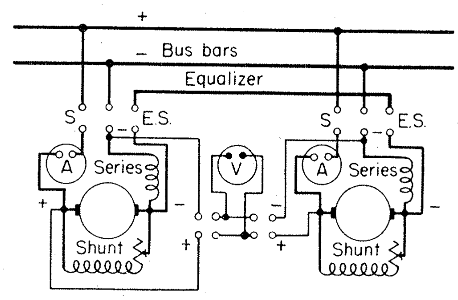


Fig. 15.1.56 Connections for compound-wound generators operating in parallel.

DC MOTORS

Motors operate on the principle that a conductor carrying current in a magnetic field tends to move at right angles to that field (see Fig. 15.1.11). The ordinary dc generator will operate entirely satisfactorily as a motor and will have the same rating. The conductors of the motor rotate in a magnetic field and therefore must generate an emf just as does the generator. The induced emf

$$E = K\phi N \qquad \textbf{(15.1.89)}$$

where K = constant, ϕ flux entering the armature from one north pole, and N = r/min [see Eq. **(15.1.87)**]. This emf is in opposition to the terminal voltage and tends to oppose current entering the armature. Its value is

$$E = V - I_a R_a \qquad \textbf{(15.1.90)}$$

where V = terminal voltage, I_a = armature current, and R_a = armature resistance [compare with Eq. **(15.1.88)**]. From Eq. **(15.1.89)** it is seen that the speed

$$N = K_s E/\phi \qquad \textbf{(15.1.91)}$$

when $K_s = 1/K$. This is the fundamental speed equation for a motor. By substituting in Eq. **(15.1.90)**

$$N = K_s(V - I_a R_a)/\phi \qquad \textbf{(15.1.92)}$$

which is the general equation for the speed of a motor.

The internal or electromagnetic torque developed by an armature is proportional to the flux and to the armature current; i.e.,

$$T_t = K_t \phi I_a \qquad \textbf{(15.1.93)}$$

when K_t is a constant. The torque at the pulley is slightly less than the internal torque by the torque necessary to overcome the rotational losses, such as friction, windage, eddy-current and hysteresis losses in the armature iron and in the pole faces.

The total mechanical power developed internally

$$P_m = EI_a \quad \text{W} \quad = \quad EI_a/746 \quad \text{hp} \qquad \textbf{(15.1.94)}$$

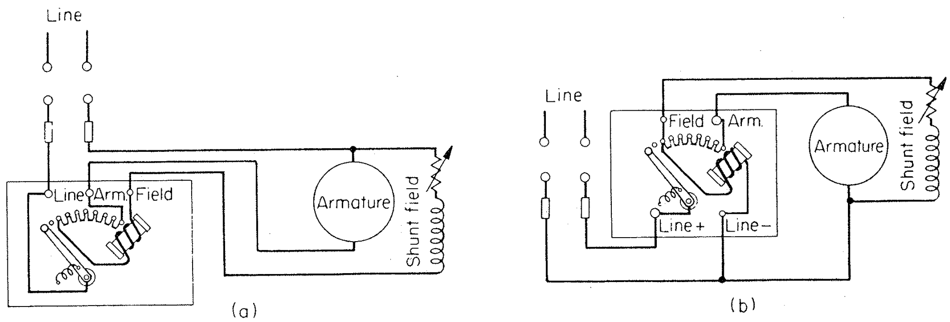


Fig. 15.1.57 Connections for shunt dc motors and starters. (*a*) Three-point box; (*b*) four-point box.

The internal torque thus becomes

$$T = EI_a 33{,}000/(2\pi \times 746N) = 7.04EI_a/N \qquad (15.1.95)$$

Let VI be the motor input. The output is $VI\eta$ where η is the efficiency. The horsepower is

$$P_H = VI\eta/746 \qquad (15.1.96)$$

and the torque is

$$T = 33{,}000P_H/2\pi N = 5{,}260P_H/N \quad \text{lb}\cdot\text{ft} \qquad (15.1.97)$$

where N is r/min.

Shunt Motor In the shunt motor (Fig. 15.1.57) the flux is substantially constant and I_aR_a is 2 to 6 percent of V. Hence from Eq. (15.1.92), the speed varies only slightly with load (Fig. 15.1.58), so that the motor is adapted to work requiring constant speed. The speed regulation of constant-speed motors is defined by ANSI/IEEE Std. 100-1977 Standard Dictionary of Electrical and Electronic terms as follows:

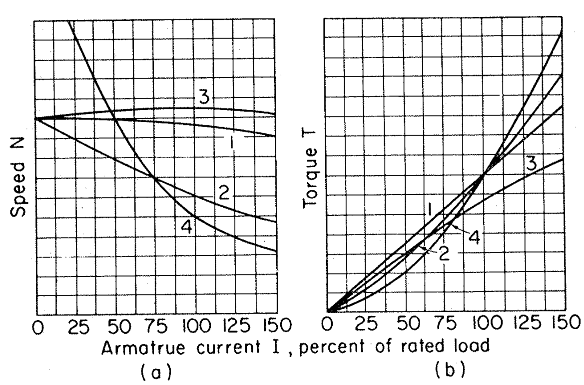


Fig. 15.1.58 Speed and torque characteristics of dc motors. (1) Shunt motor; (2) cumulative compound motor; (3) differential compound motor; (4) series motor.

> The speed regulation of a constant-speed direct-current motor is the change in speed when the load is reduced gradually from the rated value to zero with constant applied voltage and field rheostat setting expressed as a percent of speed at rated load.

In Fig. 15.1.61 the speed regulation under each condition is $100(ac - bc)/bc$ (Fig. 15.1.58*a*). Also from Eq. (15.1.93) it is seen that the torque is practically proportional to the armature current (see Fig. 15.1.58*b*). The motor is able to develop full-load torque and more on starting, but the ordinary starter is not designed to carry the current necessary for starting under load. If a motor is to be started under load, the starter should be provided with resistors adapted to carry the required current without overheating. A controller is also adapted for starting duty under load.

Commutating poles have so improved commutation in dc machines that it is possible to use a much shorter air gap than formerly. Since, with the shorter air gap, fewer field ampere-turns are required, the armature becomes magnetically strong with respect to the field. Hence, a sudden overload might weaken the field through armature reaction, thus causing an increase in speed; the effect may become cumulative and the motor run away. To prevent this, modern shunt motors are usually provided with a **stabilizing winding,** consisting of a few turns of the field in series with the armature and aiding the shunt field. The resulting increase of field ampere-turns with load will more than compensate for any weakening of the field through armature reaction. The series turns are so few that they have no appreciable compounding effect. The shunt motor is used to drive constant-speed line shafting, for machine tools, etc. Since its speed may be efficiently varied, it is very useful when **adjustable speeds** are necessary, such as individual drive for machine tools.

Shunt-Motor Starters At standstill the counter emf of the motor is zero and the armature resistance is very low. Hence, except in motors of very small size, series resistance in the armature circuit is necessary on starting. The field must, however, be connected across the line so that it may obtain full excitation.

Figure 15.1.57 shows the two common types of starting boxes used for starting shunt motors. The armature resistance remains in circuit only during starting. In the **three-point box** (Fig. 15.1.57*a*) the starting lever is held, against the force of a spring, in the running position, by an electromagnet in series with the field circuit, so that, if the field circuit is interrupted or the line voltage becomes too low, the lever is released and the armature circuit is opened automatically. In the **four-point starting box** the electromagnet is connected directly across the line, as shown in Fig. 15.1.57*b*. In this type the arm is released instantly upon failure of the line voltage. In the three-point type some time elapses before the field current drops enough to effect the release. Some starting rheostats are provided with an overload device so that the circuit is automatically interrupted if too large a current is taken by the armature. The four-point box is used where a wide speed range is obtained by means of the field rheostat. The electromagnet is not then affected by changes in field current.

In large motors and in many small motors, automatic starters are widely used. The advantages of the automatic starter are that the current is held between certain maximum and minimum values so that the circuit does not become opened by too rapid starting as may occur with manual operation; the acceleration is smooth and nearly uniform. Since workers can stop and start a motor merely by the pushing of a button, there results considerable saving by the shutting down of the motor when it is not needed. Automatic controls are essential to elevator motors so that smooth rapid acceleration with frequent starting and stopping may be obtained. Also automatic starting is very necessary with multiple-unit operation of electric-railway cars and with rolling-mill motors which are continually subjected to rapid acceleration, stopping, and reversing.

Series Motor In the series motor the armature and field are in series. Hence, if saturation is neglected, the flux is proportional to the current and the torque [Eq. (**15.1.93**)] varies as the current squared. Therefore any increase in current will produce a much greater proportionate increase in torque (see Fig. 15.1.58*b*). This makes the motor particularly well adapted to traction work, cranes, hoists, fork-lift trucks, and other types of work which require large starting torques. A study of Eq. (**15.1.92**) shows that with increase in current the numerator changes only slightly, whereas the change in the denominator is nearly proportional to the change in current. Hence the speed of the series motor is practically inversely proportional to the current. With overloads the speed drops to very low values (see Fig. 15.1.58*a*). With decrease in load the speed approaches infinity, theoretically. Hence the series motor should always be connected to its load by a direct drive, such as gears, so that it cannot reach unsafe speeds (see Speed Control of Motors). A series-motor starting box with no-voltage release is shown in Fig. 15.1.59.

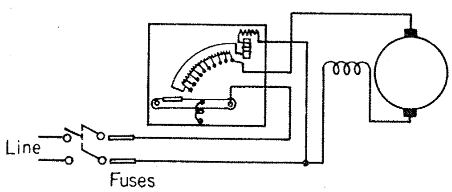


Fig. 15.1.59 Series motor starter, no-voltage release.

Differential Compound Motors The cumulative compound winding of a generator becomes a differential compound winding when the machine is used as a motor. Its speed may be made more nearly constant than that of a shunt motor, or, if desired, it may be adjusted to increase with increasing load.

The speed as a function of armature current is shown in Fig. 15.1.58*a* and the torque as a function of armature current is shown in Fig,. 15.1.58*b*.

Since the speed of the shunt motor is sufficiently constant for most purposes and the differential motor tends toward instability, particularly in starting and on overloads, the differential motor is little used.

Cumulative compound motors develop a more rapid increase in torque with load than shunt motors (Fig. 15.1.58*b*); on the other hand, they have much poorer speed regulation (Fig. 15.1.58*a*). Hence they are used where larger starting torque than that developed by the shunt motor is necessary, as in some industrial drives. They are particularly useful where large and intermittent increases of torque occur as in drives for shears, punches, rolling mills, etc. In addition to the sudden increase in torque which the motor develops with sudden applications of load, the fact that it slows down rapidly and hence causes the rotating parts to give up some of their kinetic energy is another important advantage in that it reduces the peaks on the power plant. Performance data for compound motors are given in Table 15.1.11.

Commutation The brushes on the commutator of either a motor or generator should be set in such a position that the induced emf in the armature coils undergoing commutation and hence short-circuited by the brushes, is zero. In practice, this condition can at best be only approximately realized. Frequently conditions are such that it is far from being realized. At no load, the brushes should be set in a position corresponding to the geometrical neutral of the machine, for under these conditions the induced emf in the coils short-circuited by the brushes is zero. As load is applied, two factors cause sparking under the brushes. The mmf of the armature, or **armature reaction,** distorts the flux; when the current in the coils undergoing commutation reverses, an emf of self-induction $-\ L\ di/dt$ tends to prolong the curent flow which produces sparking. In a generator, armature reaction distorts the flux in the direction of rotation and the brushes should be advanced. In order to neutralize the emf of self-induction the brushes should be set a little ahead of the neutral plane so that the emf induced in the short-circuited coils by the cutting of the flux at the fringe of the next pole is opposite to this emf of self-induction. In a motor the brushes are correspondingly moved backward in the direction opposite rotation.

Theoretically, the brushes should be shifted with every change in load. However, practically all dc generators and

Table 15.1.11 Test Performance of Compound-wound DC Motors

Hp	Rpm	115 volts		230 volts		550 volts	
		Amp	Full-load eff, percent	Amp	Full-load eff, percent	Amp	Full-load eff, percent
1	1,750	8.4	78	4.3	79	1.86	73.0
2	1,750	16.0	80	8.0	81	3.21	82.0
5	1,750	40.0	82	20.0	83	8.40	81.0
10	1,750	75.0	85.6	37.5	85	15.4	86.5
25	1,750	182.0	87.3	91.7	87.5	38.1	88.5
50	850	.....	....	180.0	89	73.1	90.0
100	850	.....	....	350.0	90.5	149.0	91.0
200	1,750	.....	....	700.0	91	295.0	92.0

SOURCE: Westinghouse Electric Corp.

R.J. Reynolds Vapor Exhibit 1034-00309

motors now have **commutating poles** (or **interpoles**) and with these the brushes can remain in the no-load neutral plane, and good commutation can be obtained over the entire range of load. Commutating poles are small poles between the main poles (Fig. 15.1.60) and are excited by a winding in series with the armature. Their function is to neutralize the flux distortion in the **neutral plane** caused by armature reaction and also to supply a flux that will cause an emf to be induced in the conductors undergoing commutation, opposite and equal to the emf of self-induction. Since armature reaction and the emf of self-induction are both proportional to the armature current, saturation being neglected, they are neutralized theoretically at every load. Commutating poles have made possible dc generators and motors of very much higher voltage, greater speeds, and larger kW ratings than would otherwise be possible.

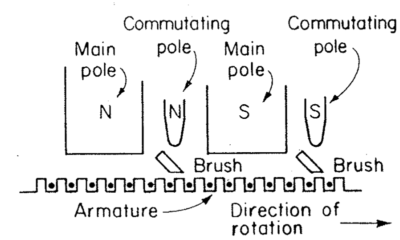


Fig. 15.1.60 Commutating poles in motor.

Occasionally, the commutating poles may be connected incorrectly. In a motor, passing from an N main pole in the direction of rotation of the armature, an N commutating pole should be encountered as shown in Fig. 15.1.60. In a generator under these conditions an S commutating pole should be encountered. The test can easily be made with a compass. If poor commutation is caused by too strong interpoles, the winding may be shunted. If the poles are too weak and the shunting cannot be reduced, they may be strengthened by inserting sheet-iron shims between the pole and the yoke thus reducing the air gap.

Although the emfs induced in the coils undergoing commutation are relatively small, the resistance of the coils themselves is low so that unless further resistance is introduced, the short-circuit currents would be large. Hence, with the exception of certain low-voltage generators, carbon brushes that have relatively large contact resistance are almost always used. Moreover, the graphite in the brushes has a lubricating action, and the usual carbon brush does not score the commutator.

Speed Control of Motors

Shunt Motors In Eq. (15.1.91) the speed of a shunt motor $N = K_s E/\phi$, where K_s is a constant involving the design of the motor such as conductors on armature surface and number of poles. Obviously, in order to change the speed of a motor, without changing its construction, two factors may be varied, the counter emf E and the flux ϕ.

Armature-Resistance Control The counter emf $E = V - I_a R_a$, where V is the terminal voltage, assumed constant. R_a must be small so that the armature heating can be maintained within permissible limits. Under these conditions the speed change with load is small. By inserting an external resistor, however, into the armature circuit the counter emf E may be made to decrease rapidly with increase in load; that is, $E = V - I_a(R_a + R)$ [see Eq. (15.1.90)] where R is the resistance of the external resistor. The resistor R must be inserted in the **armature** circuit only. The advantages of this method are its simplicity, the full torque of the motor is developed at any speed, and the method introduces no commutating difficulties. Its disadvantages are the increased speed regulation with change of load (Fig. 15.1.61), the low efficiency, particularly at the lower speeds, and the fact that provision must be made to dissipate the comparatively large power losses in the series resistor. Figure 15.1.61 shows typical speed-load curves without and with series resistors in the armature circuit. The armature efficiency is nearly equal to the ratio of the operating speed to the no-load speed. Hence at 25 percent speed the armature efficiency is practically 25 percent. Frequently the controlling and starting resistors are one, and the device is called a **controller.** Starting rheostats themselves are not designed to carry the armature current continuously and must not be used as controllers. The armature-resistance method of speed control is frequently used to regulate the speed of ventilating fans where the power demand diminishes rapidly with decrease in speed.

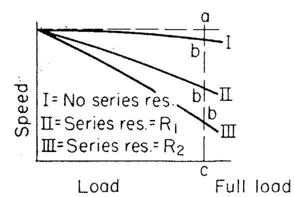


Fig. 15.1.61 Speed-load characteristics with armature resistance control.

Control by Changing Impressed Voltage From Eq. (15.1.92) it is evident that the speed of a motor may be changed if V is changed by connecting the armature across different voltages. Speed control by this method is accomplished by having mains (usually four), which are maintained at different voltages, available at the motor.

The shunt field of the motor is generally permanently connected to one pair of mains, and the armature circuit is provided with a controller by means of which the operator can readily connect the armature to any pair of mains. Such a system gives a series of distinct and widely separated speeds and generally necessitates the use of field-resistance control, in combination, to obtain intermediate speeds. This method, known as the **multivoltage method,** has the disadvantage that the system is expensive, for it requires several generating machines, a somewhat complicated switchboard, and a number of service wires. The system is used somewhat in machine shops and is extensively used for dc elevator starting and speed control.

In the **Ward Leonard system,** the variable voltage is obtained from a separately excited generator whose armature terminals are connected directly to the armature terminals of the working motor. The generator is driven at essentially constant speed by a dc shunt motor if the power supply is direct current, or by an induction motor or a synchronous motor if the power supply is alternating current. The field circuit of the generator and that of the motor are connected across a constant-voltage dc supply. The terminal voltage of the generator, and hence the voltage applied to the armature of the motor, is varied by changing the generator-field current with a field rheostat. The rheostat has a wide range of resistance so that the speed of the motor may be varied smoothly from 0 to 100 percent. Since three machines are involved the system is costly, somewhat complicated, and has low power efficiency. However, because the system is flexible and the speed can be smoothly varied over wide ranges, it has been used in many applications, such as elevators, mine hoists, large printing presses, paper machines, and electric locomotives.

The **Ward-Leonard system** has been largely replaced by a **static converter system** where a silicon-controlled-rectifier bridge is used to convert three-phase alternating current, or single-phase on smaller drives, to dc voltage that can be smoothly varied by **phase-angle firing** of the rectifiers from full voltage to zero. This system is smaller, lighter, and less expensive than the motor-generator system. Care must be taken to assure that the dc motor will accept the **harmonics** present in the dc output without overheating.

Control by Changing Field Flux Equation (**15.1.91**) shows that the speed of a motor is inversely proportional to the flux ϕ. The flux can be changed either by varying the shunt-field current or by varying the reluctance of the magnetic circuit. The variation of the **shunt-field current** is the simplest and most efficient of all the methods of speed control.

With the ordinary motor, speed variation of 1.5 to 1.0 is obtainable with this method. If attempt is made to obtain greater ratios, severe sparking at the brushes results, owing to the field distortion caused by the armature mmf becoming large in comparison with the weakened field of the motor. Speed ratios of 5:1 and higher are, however, obtainable with motors which have commutating poles. Since the field current is a small proportion of the total current (1 to 3 percent), the rheostat losses in the field circuit are always small. This method is efficient. Also for any given speed adjustment the speed regulation is excellent, which is another advantage. Because of its simplicity, efficiency, and excellent speed regulation, the control of speed by means of the field current is by far the most common method. Output power remains constant when the field is weakened, so output torque varies inversely with motor speed.

Speed Control of Series Motors The series motor is fundamentally a variable-speed motor, the speed varying widely from light load to full load and more (see Fig. 15.1.58*a*). From Eq. (**15.1.92**) the speed for any value of ϕ, or current, can be changed by varying the impressed voltage. Hence the speed can be controlled by inserting resistance in series with the motor. This method, which is practically the same as the armature-resistance control method for shunt motors, has the same objections of low efficiency and poor regulation with fluctuating loads. It is extensively used in controlling the speed of hoist and crane motors.

The **series-parallel** system of series-motor speed control is almost universally used in electric traction. At least two motors are necessary. The two motors are first connected in series with each other and with the starting resistor. The starting resistor is gradually cut out and, since each motor then operates at half line voltage, the speed of each is approximately half speed. Both motors take the same current, and each can develop full torque. This condition of operation is efficient since there is no external resistance in circuit. When the controller is moved to the next position, the motors are connected in parallel with each other and each in series with starting resistors. Full speed of the motors is obtained by gradually cutting out these resistors. Connecting the two motors in series on starting reduces the current to one-half the value that would be required for a given torque were both motors connected in parallel on starting. The power taken from the trolley is halved, and an intermediate running speed is efficiently obtained.

In the **multiple-unit** method of speed control which is used for electric railway trains, the starting contactors, reverser, etc., for each car are located under that car. The relays operating these control devices are actuated by energy taken from the train line consisting usually of seven wires. The train line runs the entire length of the train, the connections between the individual cars being made through the couplers. The train line is energized by the action of the motorman operating any one of the small master controllers which are located in each car. Hence corresponding relays, contactors, etc., in every car all operate simultaneously. High accelerations may be reached with this system because of the large tractive effort exerted by the wheels on every car.

SYNCHRONOUS GENERATORS

The synchronous generator is the only type of ac generator now in general use at power stations.

Construction In the usual synchronous generator the armature or stator, is the stationary member. This construction has many advantages. It is possible to make the slots any reasonable depth, since the tooth necks increase in cross section with increase in depth of slot; this is not true of the rotor. The large slot section which is thus obtainable gives ample space for copper and insulation. The conductors from the armature to the bus bars can be insulated throughout their entire lengths, since no rotating or sliding contacts are necessary. The insulation in a stationary member does not deteriorate as rapidly as that on a rotating member, for it is not subejcted to centrifugal force or to any considerable vibration.

The **rotating member** is ordinarily the field. There are two general types of field construction: the **salient-pole type** and the **cylindrical,** or **nonsalient-pole, type.** The salient-pole type is used almost entirely for slow and moderate-speed generators since this construction is the least expensive and permits ample space for the field ampere-turns.

It is not practicable to employ salient poles in high-speed turboalternators because of the excessive windage and the difficulty of obtaining sufficient mechanical strength. The **cylindrical type** consists of a cylindrical steel forging with radial slots in which the field copper, usually in strip form, is placed. The fields are ordinarily excited at low voltage, 125 and 250 V, the current being conducted to the rotating member by means of slip rings and brushes. An ac generator armature to supply **field voltage** can be mounted on the generator shaft and supply dc to the motor field through a static rectifier bridge, also mounted on the shaft, eliminating all slip rings and brushes. The field power is ordinarily only 1.5 percent and less of the rated power of the machine (see Table 15.1.12).

Classes of Synchronous Generators Synchronous generators may be divided into three general classes: (1) the slow-speed engine-driven type; (2) the moderate-speed waterwheel-driven type; and (3) the high-speed turbine-driven type. In (1) a hollow box frame is used as the stator support, and the field consists of a spider to which a larger number of salient poles are attached, usually bolted. The speed seldom exceeds 75 to

Table 15.1.12 Performance Data for Synchronous Generators

80 PERCENT PF, 3 PHASE, 60 CYCLE, 240 TO 2,400 VOLTS, HORIZONTAL-COUPLED OR BELTED-TYPE ENGINE-

kVA	Poles	Rpm	Excitation, kW	Efficiencies, percent			Approx net weight lb
				1/2 load	3/4 load	Full load	
25	4	1,800	0.8	81.5	85.7	87.6	900
93.8	8	900	2	87	89.5	90.9	2,700
250	12	600	5	90	91.3	92.2	6,000
500	18	400	8	91.7	92.6	93.2	10,000
1,000	24	300	14.5	92.6	93.4	93.9	16,100
3,125	48	150	40	93.4	94.2	94.6	52,000

INDUSTRIAL-SIZE TURBINE GENERATORS, DIRECT-CONNECTED TYPE, 80 PERCENT PF, 3 PHASE, 60 CYCLE, AIR-COOLED

kVA	Poles	Rpm	Excitation		Efficiency, percent			Volume of air, cfm	Voltage	Approx wt, including exciter, lb
			kW	Volts	1/2 load	3/4 load	Full load			
1,875	2	3,600	18	125	95.3	96.1	96.3	3,500	480–6,900	21,900
2,500	2	3,600	22	125	95.3	96.1	96.3	5,000	2,400–6,900	22,600
3,125	2	3,600	24	125	95.3	96.3	96.5	5,500	2,400–6,900	25,100
3,750	2	3,600	24	125–250	95.3	96.3	96.6	6,500	2,400–6,900	27,900
5,000	2	3,600	29	125–250	95.3	96.3	96.6	11,000	2,400–6,900	40,100
6,250	2	3,600	38	125–250	95.3	96.3	96.7	12,000	2,400–13,800	43,300
7,500	2	3,600	42	125–250	95.5	96.5	96.9	15,000	2,400–13,800	45,000
9,375	2	3,600	47	125–250	95.5	96.5	96.9	16,500	2,400–13,800	61,200

CENTRAL-STATION-SIZE TURBINE GENERATORS, DIRECT-CONNECTED TYPE, 85 PERCENT PF, 3 PHASE, 60 CYCLE, 11,500 TO 14,400 VOLTS

kVA	Poles	Rpm	Excitation		Efficiency, percent			Volume, of air, cfm	Ventilation	Approx wt, including exciter, lb
			kW	Volts	1/2 load	3/4 load	Full load			
13,529	2	3,600	70	250	96.3	97.1	97.3	22,000	Air-cooled	116,700
17,647	2	3,600	100	250	97.7	97.9	97.9	22,000	H_2-cooled	115,700
23,529	2	3,600	115	250	98.0	98.2	98.2	25,000	H_2-cooled	143,600
35,294	2	3,600	145	250	98.1	98.3	98.3	34,000	H_2-cooled	194,800
47,058	2	3,600	155	250	98.3	98.5	98.5	42,000	H_2-cooled	237,200
70,588	2	3,600	200	250	98.4	98.7	98.7	50,000	H_2-cooled	302,500

SOURCE; Westinghouse Electric Corp.

90 r/min, although it may run as high as 150 r/min. Waterwheel generators also have salient poles which are usually dovetailed to a cylindrical spider consisting of steel plates riveted together. Their speeds range from 80 to 900 r/min and sometimes higher, although the 9,000-kVA Keokuk synchronous generators rotate at only 58 r/min, operating at a very low head. The speed rating of direct-connected waterwheel generators decreases with decrease in head. It is desirable to operate synchronous generators at the highest permissible speed since the weight and costs diminish with increase in speed. Waterwheel-driven generators must be able to run at double speed, as a precaution against accident, should the governor fail to shut the gate sufficiently rapidly in case the circuit breakers open or should the governing mechanism become inoperative.

Turbine-driven generators operate at speeds of 720 to 3,600 r/min. Direct-connected exciters, belt-driven exciters from the generator shaft, and separately driven exciters are used. In large stations separately driven (usually motor) exciters may supply the excitation energy to excitation bus bars. Steam-driven exciters and storage batteries are frequently held in reserve. With slow-speed synchronous generators, the belt-driven exciter is frequently used because it can be driven at higher speed, thus reducing the cost.

Synchronous-Generator Design At the present time single-phase generators are seldom built. For single-phase service two phases of a standard three-phase Y-connected generator are used. A single-phase load or unbalanced three-phase load produces flux pulsations in the magnetic circuits of synchronous generators, which increase the iron losses and introduce harmonics into the emf wave. Two-phase windings consist of two similar single-phase windings displaced 90 electrical space degrees on the armature and ordinarily occupying all the slots on the armature. The most common tye of winding is the three-phase lap-wound two-layer type of winding. In three-phase windings three windings are spaced 120 electrical space degrees apart, the individual phase belts being spaced 60° apart. Usually, all the slots on the armature are occupied. Standard voltages are 550, 1,100, 2,200, 6,600, 13,200, and 20,000 V. It is much more difficult to insulate for 20,000 V than it is for the lower voltages. However, if the power is to be transmitted at this voltage, its use would be justified by the saving of transformers. In machines of moderate and larger ratings it is common to generate at 6,600 and 13,200 V if transformers must be used. The higher voltage is preferable, particularly for the higher ratings, because it reduces the cross section of the connecting leads and bus bars.

The **standard frequency** in the United States for lighting and power systems is 60 Hz; the few former 50-Hz systems have practically all been converted to 60 Hz. The frequency of 25 Hz is commonly used in street-railway and subway systems to supply power to the synchronous converters and other ac-dc conversion apparatus; it is also commonly used in railroad electrification, particularly for single-phase series-motor locomotives (see Sec. 11). At 25 Hz incandescent lamps have noticeable flicker. In European (and most other) countries 50 Hz is standard. The frequency of a synchronous machine

$$f = P \times \text{r/min}/120 \qquad \text{Hz} \tag{15.1.98}$$

where P is the number of poles. Synchronous generators are rated in kVA rather than in kW, since heating, which determines the rating, is dependent only on the current and is independent of power factor. If the kilowatt rating is specified, the power factor should also be specified.

Induced EMF The induced emf per phase in synchronous generator is

$$E = 2.22 k_b k_p \Phi f Z \qquad \text{V/phase} \tag{15.1.99}$$

where k_b = breadth factor or belt factor (usually 0.9 to 1.0), which depends on the number of slots per pole per phase, 0.958 for three-phase, four slots per pole per phase; k_p = pitch factor = 1.0 for full pitch, 0.966 for ⅚ pitch; Φ = total flux Wb, entering armature from one north pole and is assumed to be sinusoidally distributed along the air gap; f = frequency; and Z = number of series conductors per phase.

Synchronous generators usually are Y-connected. The advantages are that for a given line voltage the voltage per phase is $1/\sqrt{3}$ that of the delta-connected winding; third-harmonic currents and their multiples cannot circulate in the winding as with a delta-connected winding; third-harmonic emfs and their multiples cannot exist in the line emfs; a neutral point is available for grounding.

Regulation The terminal voltage of synchronous generator at constant frequency and field excitation depends not only on the current load but on the power factor as well. This is illustrated in Fig. 15.1.62, which shows the voltage-current characteristics of a synchronous generator with lagging current, leading current and in-phase current (pf = 1.00). With leading current the voltage may actually rise with increase in load; the rate of voltage decrease with load becomes greater as the lag of the current increases. The regulation of a synchronous generator is defined by the ANSI/IEEE Std. 100-1977 Standard Dictionary of Electrical and Electronic Terms as follows:

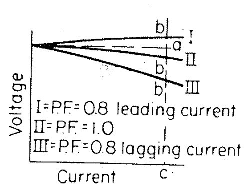


Fig. 15.1.62 Synchronous generator characteristics.

The voltage regulation of a synchronous generator is the rise in voltage with constant field current, when, with the synchronous generator operated at rated voltage and rated speed, the specified load at the specified power factor is reduced to zero, expressed as a percent of rated voltage.

For example, in Fig. 15.1.62 the regulation under each condition is

$$100(ac - bc)/bc \tag{15.1.100}$$

With leading current the regulation may be negative.

Three factors affect the regulation of synchronous generators; the **effective armature resistance,** the **armature leakage reactance,** and the **armature reaction.** With alternating current the armature loss is greater than the value obtained by multiplying the square of the armature current by the ohmic resistance. This is due to hysteresis and eddy-current losses in the iron adjacent to the conductor and to the alternating flux-producing losses in the conductors themselves. Also the current is not distributed uniformly over conductors in the slot, but the current density tends to be greatest in the top of the slot. These factors all have the effect of increasing the resistance. The ratio of effective to ohmic resistance varies from 1.2 to 1.5. The **armature leakage reactance** is due to the flux produced by the armature current linking the conductors in the slots and also the end connections.

The armature mmf reacts on the field to change the value of the flux. With a single-phase generator and with an unbalanced load on a polyphase generator, the armature mmf is pulsating and causes iron losses in the field structure. With polyphase machines under a constant balanced load, the armature mmf is practically constant in magnitude and fixed in its relation to the field poles. Its direction with relation to the field-pole axis is determined by the power factor of the load.

A component of current in phase with the no-load induced emf, or the excitation emf, merely distorts the field by strengthening the trailing pole tip and weakening the leading pole tip. A component of current lagging the excitation emf by 90° weakens the field without distortion. A component of current leading the excitation emf by 90° strengthens the field without distortion. Ordinarily, both cross magnetization and one of the other components are acting simultaneously.

The foregoing effects are called **armature reaction.** Frequently the effects of armature reactance and armature reaction can be combined into a single quantity.

It is difficult to determine the regulation of synchronous generator by actual loading, even when in service, owing to the difficulty of obtaining, controlling, and absorbing the large balanced loads. Hence methods of **predetermining regulation without actually loading** the machine are used.

Synchronous Impedance Method Both armature reactance and armature reaction have the same effect on the terminal voltage. In the synchronous impedance method the generator is considered as having no armature reaction, but the armature reactance is increased a sufficient amount to account for the effect of armature reaction. The phasor diagram for a current I lagging the terminal voltage V by an angle θ is shown in Fig. 15.1.63. In a polyphase generator the phasor diagram is applicable to one phase, a balanced load almost always being assumed.

The power factor of the load is $\cos\theta$; IR is the effective armature resistance drop and is parallel to I; IX_s is the synchronous reactance drop and is at right angles to I and leading it by 90°. IX_s includes both the reactance drop and the drop in voltage due to armature reaction. That part of IX_s which replaces armature reaction is in reality a fictitious quantity. The synchronous impedance drop is given by IZ_s. The no-load or open-circuit (excitation) voltage

$$E = \sqrt{(V\cos\theta + IR)^2 + (V\sin\theta \pm IX_s)^2} \quad \text{V} \tag{15.1.101}$$

All quantities are per phase. The negative sign is used with leading current.

$$\text{Regulation} = 100(E - V)/V \tag{15.1.102}$$

With leading current E may be less than V and a negative regulation results.

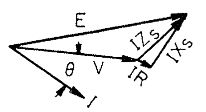


Fig. 15.1.63 Phasor diagram for synchronous impedance method.

The synchronous impedance is determined from an open-circuit and a short-circuit test, made with a weak field. The voltage E' on open circuit is divided by the current I' on short circuit for the same value of field current.

$$Z_s = E'/I' \qquad X_s = \sqrt{Z_s^2 - R^2} \quad \Omega \tag{15.1.103}$$

R is so small compared with X_s that for all practical purposes $X_s = Z_s$. R may be determined by measuring the ohmic resistance per phase and multiplying by 1.4 to 1.5 to obtain the effective resistance. This value of R and the value of X_s obtained from Eq. (15.1.103) may then be substituted in Eq. (15.1.101) to obtain E at the specified load and power factor.

Since the synchronous reactance is determined at low saturation of the iron and used at high saturation, the method gives regulations that are too large; hence it is called the **pessimistic method.**

MMF Method In the mmf method the generator is considered as having no armature reactance but the armature reaction is increased by an amount sufficient to include the effect of reactance. That part of armature reaction which replaces the effect of armature reactance is in reality a fictitious quantity. To obtain the data necessary for computing the regulation, the generator is short-circuited and the field adjusted to give rated current in the armature. The corresponding value of field current I_a is read. The field is then adjusted to give voltage E' equal to rated terminal voltage + IR drop (= $V + IR$, as phasors, Fig. 15.1.64) on open circuit and the field current I' read.

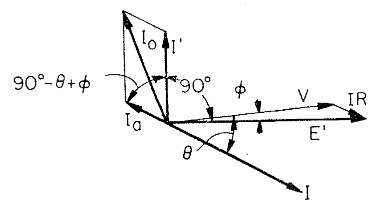


Fig. 15.1.64 Phasor diagram for mmf method.

I_a is 180° from the current phasor I, and I' leads E' by 90° (Fig. 15.1.64). The angle between I' and I_a is $90 - \theta + \phi$, but since ϕ is small, it can usually be neglected. The phasor sum of I_a and I' is I_o. The open-circuit voltage E corresponding to I_o is the no-load voltage and can be found on the saturation curve. The regulation is then found from Eq. (15.1.102). This method gives a value of regulation less than the actual value and hence is called the **optimistic method.** The actual regulation lies somewhere between the values obtained by the two methods but is more nearly equal to the value obtained by the mmf method.

ANSI Method The ANSI method (American Standard 50, Rotating Electrical Machinery) which has become the accepted standard for the predetermination of synchronous generator operation, eliminates in large measure the errors due to saturation which are inherent in the synchronous impedance and mmf methods. In Fig. 15.1.65*a* is shown the saturation curve OAF of the generator. The axis OP is not only the field-current axis but also the axis of the current phasor I as well. V the terminal voltage is drawn θ deg from I or OP, where θ is the power factor angle. The effective resistance drop IR and the **leakage reactance** drop IX are drawn parallel and perpendicular to the current phasor. E_a, the phasor sum of V, IR, and IX, is the internal **induced** emf. Arcs are swung with O as the center and V and E_a as radii to intercept the axis of ordinates at B and C. OK, tangent to the straight portion of the saturation curve, is the **air-gap line.** If there is no saturation, I_v is the

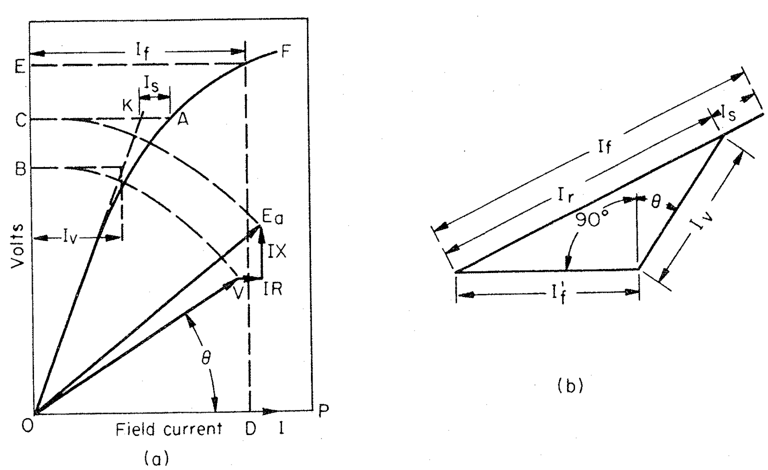


Fig. 15.1.65 ANSI method of synchronous generator regulation.

field current necessary to produce V, and CK is the field current necessary to produce E_a. The field current I_s is the increase in field current necessary to take into account the saturation corresponding to E_a.

The corresponding phasor diagram to a larger scale is shown in Fig. 15.1.65*b*. I'_f, the field current necessary to produce rated current at short circuit, corresponding to I_a (Fig. 15.1.64), is drawn horizontally. The field current I_v is drawn at an angle θ to the right of a perpendicular erected at the right-hand end of I'_f. I_r is the resultant of I'_f, and I_v. I_s is added to I_r giving I_f the resultant field current. The no-load emf E is found on the saturation curve, Fig. 15.5.65*a*, corresponding to $I_f = OD$.

Excitation is commonly supplied by a small dc generator driven from the generator shaft. On account of commutation, except in the smaller sizes, the dc generator cannot be driven at 3,600 r/min, the usual speed for turbine generators, and belt or gear drives are necessary. The use of the silicon rectifier has made possible simpler means of excitation as well as voltage regulation. In one system the exciter consists of a small rotating-armature synchronous generator (which can run at high speed) mounted directly on the main generator shaft. The three-phase armature current is rectified by six silicon rectifiers and is conducted directly to the main generator field without any sliding contacts. The main generator field current is controlled by the current to the stationary field of the exciter generator. In another system there is no rotating exciter, the generator excitation being supplied directly from the generator terminals, the 13,800 V, three-phase, being stepped down to 115 V, three-phase, by small transformers and rectified by silicon rectifiers. Voltage regulation is obtained by saturable reactors actuated by potential transformers connected across the generator terminals.

Most regulators such as the following operate through the field of the exciter. In the **Tirrill regulator** the field resistance of the exciter is short-circuited temporarily by contacts when the bus-bar voltage drops. Actually, the contacts are vibrating continuously, the time that they are closed depending on the value of the bus-bar voltage. The General Electric Co. manufactures a direct-acting regulator in which the regulating rheostat is part of the regulator itself. The rheostat consists of stacks of graphite plates, each plate being pivoted at the center. Tilting the plates changes the path of the current through the rheostat and thus changes the resistance. The plates are tilted by a sensitive torque armature which is actuated by variations of voltage from the normal value (for regulators employing silicon rectifiers).

Parallel Operation of Synchronous Generators The kilowatt division of load between synchronous generators in parallel is determined entirely by the speed-load characteristics of their prime movers and not by the characteristics of the generators themselves. No appreciable adjustment of kilowatt load between synchronous generators in parallel can be made by means of their field rheostats, as with dc generators. Consider Fig. 15.1.66, which gives the speed-load characteristics in terms of frequency, of two synchronous generators, no. 1 and no. 2, these characteristics being the speed-load characteristics of their prime movers. These speed-load characteristics are drooping, which is necessary for stable parallel operation. The total load on the two machines is $P_1 + P_2$ kW. Both machines must be operating at the same frequency F_1. Hence generator 1 must be delivering P_1 kW, and generator 2 must be delivering P_2 kW (the small generator losses being neglected). If, under the foregoing conditions, the field of either machine is strengthened, it cannot deliver a greater kilowatt load, for its prime mover can deliver more power only by dropping its speed. This is impossible, for both generators must operate always at the same frequency f_1. For any fixed total power load, the division of kilowatt load between synchronous generators can be changed only by modifying in some manner the

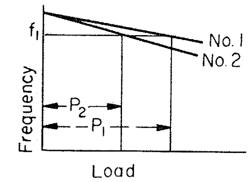


Fig. 15.1.66 Speed-load characteristics of synchronous generators in parallel.

speed-load characteristics of their prime movers, such, for example, as changing the tension in the governor spring. Synchronous generators in parallel are of themselves in stable equilibrium. If the driving torque of one machine is increased, the resulting electrical reactions between the machines cause a circulating current to flow between machines. This current puts more electrical load on the machine whose driving torque is increased and tends to produce motor action in the other machines. In an extreme case, the driving torque of one prime mover may be removed entirely, and its generator will operate as a synchronous motor, driving the prime mover mechanically.

Variations in driving torques cause currents to circulate between synchronous generators, transferring power which tends to keep the generators in synchronism. If the power transfer takes the form of recurring pulsations, it is called **hunting**, which may be reduced by building heavy copper grids called **amortisseurs**, or **damper windings**, into the pole faces. Turbine- and water-wheel-driven synchronous generators are much better adapted to parallel operation than are synchronous generators which are driven by reciprocating engines, because of their uniformity of torque.

Increasing the field current of synchronous generators in parallel with others causes it to deliver a greater lagging component of current. Since the character of the load determines the total current delivered by the system, the lagging components of current delivered by the other generators must decrease and may even become leading components. Likewise if the field of one generator is weakened, it delivers a greater leading component of current and the other machines deliver components of current which are more lagging. These leading and lagging currents do not affect appreciably the division of **kilowatt** load between the synchronous generators. They do, however, cause unnecessary heating in their armatures. The fields of all synchronous generators should be so adjusted that the heating due to the quadrature components of currents is a minimum. With two generators having equal armature resistances, this occurs when both deliver equal quadrature currents.

Armature reactance in the armature of machines in parallel is desirable. If not too great, it stabilizes their operation by producing the synchronizing action. Synchronous generators with too little reactance are sensitive, and if connected in parallel with slight phase displacement or inequality of voltage, considerable disturbance results. Armature reactance also reduces the current on short circuit, particularly during the first few cycles when the short-circuit current is a maximum. Frequently, external power-limiting reactances are connected in series to protect the generators and equipment from injury that would result from the tremendous short-circuit currents. For these reasons, poor regulation in large synchronous generators is frequently considered to be an advantage rather than a disadvantage.

Ground Resistors Most power systems operate with a grounded neutral. When the station generators deliver current directly to the system (without intervening transformers), it is customary to ground the neutral (of the Y-connected windings) of one generator in a station; this is usually done through a grounding resistor of from 2 to 6 Ω. If the neutral of more than one generator is grounded, third-harmonic (and multiples thereof) currents can circulate between the generators. The ground resistor reduces the short-circuit currents when faults to ground occur, and hence reduces the violence of the short circuit as well as the duty of the circuit breakers. Grounding reactors are sometimes used but have limited application owing to the danger of high voltages resulting from resonant conditions.

INDUCTION GENERATORS

The **induction generator** is an induction motor driven above synchronous speed. The rotor conductors cut the rotating field in a direction to convert shaft mechanical power to electrical power. Load increases as speed increases, so the generator is self-regulating and can be used without governor control. On short circuits the induction generator will deliver current to the fault for only a few cycles because, unlike the synchronous generator, it is not self-exciting. Since it is not self-excited, an induction generator must always be used in parallel with an electrical system where there are some synchronous machines, or capacitor banks, to deliver lagging current (VARs) to the induction generator for excitation.

Induction generators have found favor in industrial **cogeneration** applications and as **wind-driven** generators where they provide a small part of the total load.

CELLS

Fuel cells convert chemical energy of fuel and oxygen directly to electrical energy. **Solar cells** convert solar radiation to electrical energy. At present, these conversion methods are not economically competitive with historic generating techniques, but they have found applications in isolated areas, such as microwave relaying stations and satellites, where power requirements are small and costs for transmitting electrical power from more conventional sources is prohibitive.

TRANSFORMERS

Transformer Theory The transformer is a device that transfers energy from one electric circuit to another without change of frequency and usually, but not always, with a change in voltage. The energy is transferred through the medium of a magnetic field: it is supplied to the transformer through a primary winding and is delivered by means of a secondary winding. Both windings link the same magnetic circuit. With no load on the secondary, a small current, called the exciting current, flows in the primary and produces the alternating flux. This flux links both primary and secondary windings and induces the same volts per turn in each. With a sine wave the emf is

$$E = 4.44\Phi_m nf \quad \text{V} \tag{15.1.104}$$

where Φ_m = maximum instantaneous flux in webers, n = turns on either winding, and f = frequency. Equation **(15.1.104)** may also be written

$$E = 4.44B_m Anf \quad \text{V} \tag{15.1.105}$$

B_m = maximum instantaneous flux density in iron and A = net cross section of iron. If B_m is in T, A is in m^2; if B_m is in Mx/in^2, A is in in^2.

In English units, Eq. **(15.1.104)** becomes

$$E = 4.44\Phi_m nf10^{-8} \quad \text{V} \tag{14.1.104a}$$

where Φ_M is in maxwells; Eq. **(15.1.105)** becomes

$$E = 4.44B_m Anf10^{-8} \quad \text{V} \tag{15.1.105a}$$

where B_m is in Mx/in^2 and A is in in^2.

B_m is practically fixed. In large transformers with silicon steel it varies between 60,000 and 75,000 Mx/in^2 at 60 Hz and between 75,000 and 90,000 Mx/in^2 at 25 Hz. It is desirable to operate the iron at as high density as possible in order to minimize the weight of iron and copper. On the other hand, with too high densities the eddy-current and hysteresis losses become too great, and with low frequency the exciting current may become excessive. It follows from Eq. **(15.1.104)** that

$$E_1/E_2 = n_1/n_2 \tag{15.1.106}$$

where E_1 and E_2 are the primary and secondary emfs and n_1 and n_2 are the primary and secondary turns. Since the impedance drops in ordinary transformers are small, the terminal voltages of primary and secondary are also practically proportional to their number of turns. As the change in secondary terminal voltage in the ordinary constant-potential transformer over its range of operation is small (1.5 to 3 percent), the flux must remain substantially constant. Therefore, the added ampere-turns produced by any secondary load must be balanced by opposite and equal primary ampere-turns. Since the exciting current is small compared with the load current (1.5 to 5 percent) and the two are usually out of phase, the exciting current may ordinarily be neglected. Hence,

$$n_1 I_1 = n_2 I_2 \tag{15.1.107}$$

$$I_1/I_2 = n_2/n_1 \tag{15.1.108}$$

where I_1 and I_2 are the primary and secondary currents.

When load is applied to the secondary of a transformer, the secondary ampere-turns reduce the flux slightly. This reduces the counter emf of the primary, permitting more current to enter and thus supply the increased power demanded by the secondary.

Both primary and secondary windings must necessarily have resistance. All the flux produced by the primary does not link the secondary; the counter ampere-turns of the secondary produce some flux which does not link the primary. These **leakage fluxes** produce reactance in each winding. The combined effect of the resistance and reactance produces an impedance drop in each winding when current flows. These impedance drops produce a slight drop in the secondary terminal voltage with load.

Transformer Testing Transformer regulation and losses are so small that it is far more accurate to compute the regulation and efficiency than to determine them by actual measurement. The necessary measurements and computations are comparatively simple, and little power is involved in making the tests. In the **open-circuit** test, the power input to either winding is measured at its rated voltage. Usually it is more convenient to make this test on the low-voltage winding, particularly if it is rated at 110, 220, or 550 V. The open-circuit power practically all goes to supply the core losses, consisting of eddy-current and hysteresis losses. Let this value of power be P_0. The eddy-current loss varies as the square of the voltage and frequency; the hysteresis loss varies as the 1.6 power of the voltage, and directly as the frequency. In the **short-circuit** test one winding is short-circuited, and the current in the other is adjusted to near its rated value. The voltage V_c, the current I_1, and the power input P_c are measured. When one winding of a transformer is short-circuited, the voltage across the other winding is 3 to 4 percent of rated value when rated current flows. Since a voltage range of from 110 to 250 V is best adapted to measuring instruments, that winding whose rated voltage, multiplied by 0.03 or 0.04, is closest to this voltage range should be used for making the short-circuit test, the other winding being short-circuited. Practically all the power on short circuit goes to supply the copper loss of primary and secondary. If the measurements are made on the primary,

$$R_{01} = P_c/I_1^2 \qquad \textbf{(15.1.109)}$$
$$Z_{01} = V_c/I_1 \qquad \textbf{(15.1.110)}$$
$$X_{01} = \sqrt{Z_{01}^2 - R_{01}^2} \qquad \textbf{(15.1.111)}$$

where R_{01}, Z_{01}, and X_{01} are the equivalent resistance, impedance, and reactance referred to the primary. Also $R_{02} = R_{01}(n_2/n_1)^2$; $Z_{02} = Z_{01}(n_2/n_1)^2$; $X_{02} = X_{01}(n_2/n_1)^2$, these quantities being the equivalent resistance, impedance, and reactance referred to the secondary. If the dc resistances R_1 and R_2 of the primary and secondary are measured,

$$R_{01} = R_1 + (n_1/n_2)^2R_2 \qquad \textbf{(15.1.112)}$$
$$R_{02} = R_2 + (n_2/n_1)^2R_1 \qquad \textbf{(15.1.113)}$$

The ac or effective resistances, determined from Eq. **(15.1.109)**, are usually 10 to 15 percent greater than these values.

Regulation The regulation may be computed from the foregoing data as follows:

$$V_1' = \sqrt{(V_1 \cos\theta + I_1R_{01})^2 + (V_1 \sin\theta \pm I_1X_{01})^2} \qquad \textbf{(15.1.114)}$$
$$\text{Regulation} = 100(V_1' - V_1)/V_1 \qquad \textbf{(15.1.115)}$$

V_1 = rated primary terminal voltage; $\cos\theta$ = load power factor; I_1 = rated primary current; R_{01} = equivalent resistance referred to primary [from Eq. **(15.1.109)**]; X_{01} = equivalent reactance referred to primary. The + sign is used with lagging current and the − sign with leading current. Equations **(15.1.114)** and **(15.1.115)** are equally applicable to the secondary if the subscripts are changed.

Efficiency The only two losses in a constant-potential transformer are the core loss in W, P_0, which is practically independent of load, and P_c the copper loss in W, which varies as the load current squared. The efficiency for any current I_1 is

$$\eta = V_1I_1 \cos\theta/(V_1I_1 \cos\theta + P_0 + I_1^2R_{01}) \qquad \textbf{(15.1.116)}$$

Equation **(15.1.116)** applies equally well to the secondary if the subscripts are changed. The maximum efficiency occurs when the core and copper losses are equal.

All-Day Efficiency Since transformers must usually be on the line 24 h/day, part of which time the load may be very light, the all-day efficiency is important. This is equal to the total energy or watthour output divided by the total energy or watthour input for the 24 h. That is,

$$\eta = \frac{(V_1I_1 \cos\theta_1)t_1 + \cdots}{(V_1I_1 \cos\theta_1)t_1 + \cdots + (I_1^2R_{01})t_1 + \cdots + 24\,P_0} \qquad \textbf{(15.1.117)}$$

where t_1 = time in hours that load $V_1I_1 \cos\theta_1$ is being delivered, etc.

Polyphase Transformer Connections Three-phase transformer banks may be connected Δ-Δ, Δ-Y, Y-Y, and Y-Δ. The Δ-Δ connection is very common, particularly at the lower voltages, and has the important advantage that the bank will operate V-connected if one transformer is disabled. The Δ-Y connection is advantageous for stepping up to high voltages since the secondary of the transformers need be wound only for 58 percent ($1/\sqrt{3}$) of the line voltage; it is also necessary when a four-wire three-phase system is obtained from a three-wire three-phase system since "a floating neutral" on the secondary cannot occur. The Y-Y system may be used for stepping up voltage. It should not be used for obtaining a three-phase four-wire system from a three-phase three-wire system, because of the "floating neutral" on the secondary and the resulting high degree of unbalance of the secondary voltages. The Y-Δ system may be used to step down high voltages, the reverse of the Δ-Y connection. Y-connected windings also preclude third harmonic, and their multiples, circulating currents in the transformer windings. In the Δ-Y and Y-Δ systems the ratio of line voltage is obviously not that of the individual transformers. Because of different phase displacement between primaries and secondaries, a Δ-Δ bank cannot be connected in parallel (on both sides) with a Δ-Y bank, etc., even if they both have the correct voltage ratios between lines.

Three-phase transformers combine the magnetic circuits of three single-phase transformers so that they have parts in common. A material saving in cost, in weight, and in space results, the greatest saving occurring in the core and oil. The advantages of three-phase transformers are often outweighed by their lack of flexibility. The failure of a single phase shuts

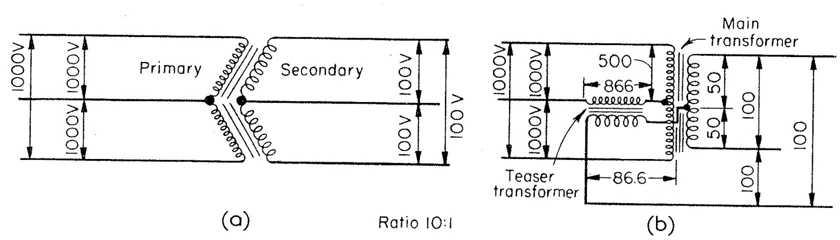


Fig. 15.1.67 Transformer connections for transforming moderate amounts of three-phase power. (*a*) Open delta connection; (*b*) T connection.

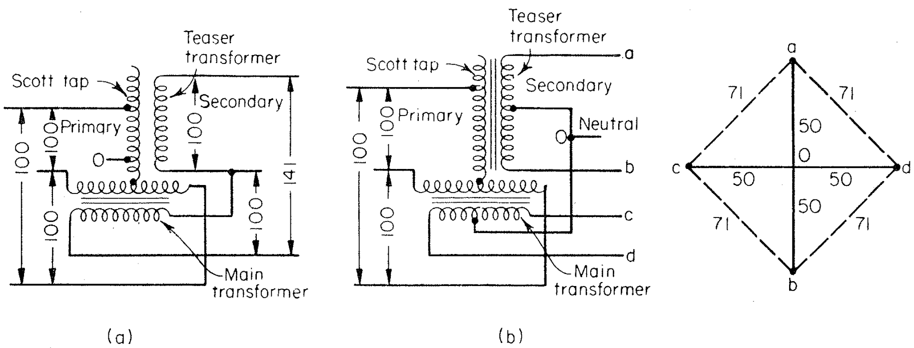


Fig. 15.1.68 Connections for transforming from three-phase to two- and four-phase power.

down the entire transformer. With three single units, one unit may be readily replaced with a single spare. The primaries of single-phase transformers may be connected in Y or Δ at will and the secondaries properly phased. The primaries, as well as the secondaries of three-phase transformers, must be phased.

For the transformation of moderate amounts of power from three-phase to three-phase, two transformers employing either the **open delta** or **T connection** (Fig. 15.1.67) may be used. With each connection the ratio of line voltages is the same as the transformer ratios. In the figure, ratios 10:1 are shown. In the T connection the primary and the secondary of the main transformer must be provided with a center tap to which one end of the teaser transformer is connected. The ratings of these systems are only 58 percent of the rating of the system using three similar transformers, one for each phase. Owing to dissymmetry, the terminal voltages become somewhat unbalanced even with a balanced load.

To transform from two- to three-phase or the reverse, the **T connection** of Fig. 15.1.68 is used. To make the secondary voltages symmetrical a tap (called a **Scott tap**) is brought out at 86.6 percent ($\sqrt{3}/2$) of the primary winding of the teaser transformer as shown in Fig. 15.1.68. With balanced no-load voltages the voltages become slightly unbalanced even under a symmetrical load, owing to unequal phase differences in the individual coils. The three-phase neutral O is one-third the winding of the teaser transformer from the junction. In Fig. 15.1.68*a* the transformation is from three-phase to a two-phase three-wire system. In Fig. 15.1.68*b* the transformation is from three-phase to a four-phase, five-wire system. The voltages are given on the basis of 100-V primaries with 1:1 transformer ratios.

An **autotransformer,** also called **compensator,** consists essentially of a single winding linking a magnetic circuit. Part of the energy is transformed, and the remainder flows through conductively. Suitable taps are provided so that, if the primary voltage is applied to two of the taps, a voltage may be taken from any other two taps. The ratio of voltages is equal practically to the ratio of the turns between their taps. An autotransformer should be installed only when the ratio of transformation is not large. The ratio of power transformed to total power is $1 - n$, where n is the ratio of low-voltage to high-voltage emf. This gives the saving over the ordinary transformer and is greatest when the ratio is not far from unity. Figure 15.1.69*a* shows 100 kW being changed from 3,300 to 2,300 V; 30.3 kW only are being actually transformed, and the remainder of the power flows through conductively. Figure 15.1.69*b* shows how an ordinary 10:1, 10-kW lighting transformer may be connected to boost 110 kW 10 percent in voltage. In Fig. 15.1.69*b*, however, the 230-V secondary must be insulated for 2,300 V to the core and ground. The voltage may likewise be reduced by reversing the 230-V coil. An autotransformer should never be used when it is desired to keep dangerous primary potentials from the secondary. It is used for starting induction motors (Fig. 15.1.71) and for a number of similar purposes.

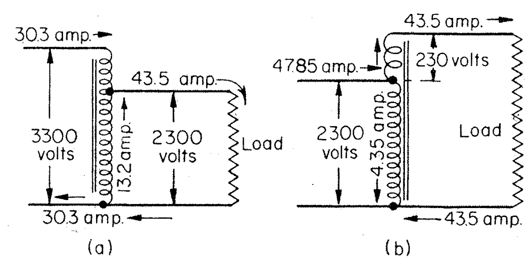


Fig. 15.1.69 Autotransformer.

Data on Transformers Single-phase 55° self-cooled oil-insulated transformers for 2,300-V primaries, 230/115-V secondaries, and in sizes from 5 to 200 kVA for 60(25) Hz have efficiencies from one-half to full load of about 98 (97 to 98.7) percent and regulation of 1.5 (1.1 to 2.1) percent with pf = 1, and 3.5 (2.7 to 4.1) percent with pf = 0.8. Power transformers with 13,200-V primaries and 2,300-V secondaries in sizes from 667 to 5,000 kVA and for both 60 and 25 Hz have efficiencies from one-half to full load of about 99.0 percent and regulation of about 1.0 (4.2) percent with pf = 1 (0.8).

AC MOTORS

Polyphase Induction Motor The polyphase induction motor is the most common type of motor used. It consists ordinarily of a stator which is wound in the same manner as the synchronous-generator stator. If two-phase current is supplied to a two-phase winding or three-phase current to a three-phase winding, a rotating magnetic field is produced in the air gap. The number of poles which this field has is the same as the number of poles that a synchronous generator employing the same stator winding would have. The speed of the rotating field, or the **synchronous speed,**

$$N = 120f/P \quad \text{r/min} \tag{15.1.118}$$

where f = frequency and P = number of poles.

There are two general types of rotors. The **squirrel-cage type** consists of heavy copper or aluminum bars short-circuited by end rings, or the bars and end rings may be an integral aluminum casting. The **wound rotor** has a polyphase winding of the same number of poles as the stator, and the terminals are brought out to slip rings so that external resistance may be introduced. The rotor conductors must be cut by the rotating field, hence the rotor cannot run at synchronous speed but must slip. The per unit **slip** is

$$s = (N - N_2)/N \qquad \textbf{(15.1.119)}$$

where N_2 = the rotor speed, r/min. The rotor frequency

$$f_2 = sf \qquad \textbf{(15.1.120)}$$

The torque is proportional to the air-gap flux and the components of rotor current in space phase with it. The rotor currents tend to lag the emfs producing them, because of the rotor-leakage reactance. From Eq. **(15.1.120)** the rotor frequency and hence the rotor reactance ($X_2 = 2\pi f_2 L_2$) are low when the motor is running near synchronous speed, so that there is a large component of rotor current in space phase with the flux. With large values of slip the increased rotor frequency increases the rotor reactance and hence the lag of the rotor currents behind their emfs, and therefore considerable space-phase difference between these currents and the flux develops. Consequently even with large values of current the torque may be small. The torque of the induction motor increases with slip until it reaches a maximum value called the **break-down torque,** after which the torque decreases (see Fig. 15.1.72). The break-down torque varies as the square of the voltage, inversely as the stator impedance and rotor reactance, and is independent of the rotor resistance.

The squirrel-cage motor develops moderate torque on starting ($s = 1.0$) even through the current may be three to seven times rated current. For any value of slip the torque of the induction motor varies as the square of the voltage. The torque of the squirrel-cage motor which, on starting, is only moderate may be reduced in the larger motors because of starting voltage drop from inrush and possible necessity of applying reduced-voltage starting.

Polyphase squirrel-cage motors are used for constant-speed work. They are used widely on account of their rugged construction and the absence of moving electrical contacts, which makes them suitable for operation when exposed to flammable dust or gas. General-purpose squirrel-cage motors have starting torques of 100 to 250 percent of full load torque at rated voltage. The highest torques occur at the higher rated speeds. The locked rotor currents vary between four and seven times full-load current. In the **double-squirrel-cage** type of motor there is a high-resistance winding in the top of the rotor slots and a low-resistance winding in the bottom of the slots. The low-resistance winding is made to have a high leakage reactance, either by separating the windings with a magnetic bridge, Fig. 15.1.70*a*, or by making the slot very narrow in the area

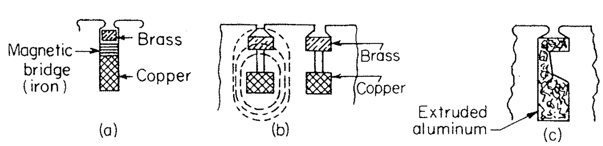


Fig. 15.1.70 Types of slots for squirrel-cage windings.

between the two windings, Fig. 15.1.70*b*. On starting, because of the high reactance of the low-resistance winding, most of the rotor current will flow in the high-resistance winding, giving the motor a large starting torque. As the rotor approaches the low value of slip at which it normally operates, the rotor frequency and hence the rotor reactance become low and most of the rotor current now flows in the low-resistance winding. Cage bars can be shaped so that one winding gives comparable results. See Fig. 15.1.70*c*. The rotor operates with a low value of slip. The high starting torque of the high-resistance motor and the excellent constant-speed operating characteristics of the low-resistance squirrel-cage rotor are combined in one motor.

The **single shaped** cage bar is more economical, and almost any shape for required characteristics can be extruded from aluminum. Single bars also eliminate the problems of differential expansion with double cage bars.

Nameplates of polyphase integral-hp squirrel-cage induction motors carry a *code* letter and a *design* letter. These provide information about motor characteristics, the former on locked rotor or starting inrush current (see Table 15.1.26) and the latter on torque characteristics. National Electrical Manufacturers Association standards publication No. MG1-1978 defines four design letters: A, B, C, and D. In all cases the motors are designed for full voltage starting. Locked rotor current and torque, pull-up torque and breakdown torque are tabulated according to horsepower and speed. Designs A, B, and C have full load slips less than 5 percent and design D more than 5 percent. The nature of the various designs can be understood by reference to the full voltage values for a 100-hp, 1800-r/min motor which follow:

	Design			
	A	B	C	D
Locked rotor torque	125*	125*	200*	275*
Pull-up torque	100†	100†	140†	...
Breakdown torque	200†	200†	190†	...
Locked rotor current	...	600*	600*	600*
Full load slip (%)	5§	5§	5§	5‡

NOTE: All quantities, except slip, are percent of full load value.
*Upper limit.
†Not less than.
‡Greater than.
§Less than.

Starting It is desirable to start induction motors by direct connection across the line, since reduced voltage starters are expensive and almost always reduce the starting torque. The capacity of the distribution system dictates when reduced voltage starting must be used to limit voltage dips on the system. On stiff industrial systems 25,000-hp motors have been successfully started **across-the-line.**

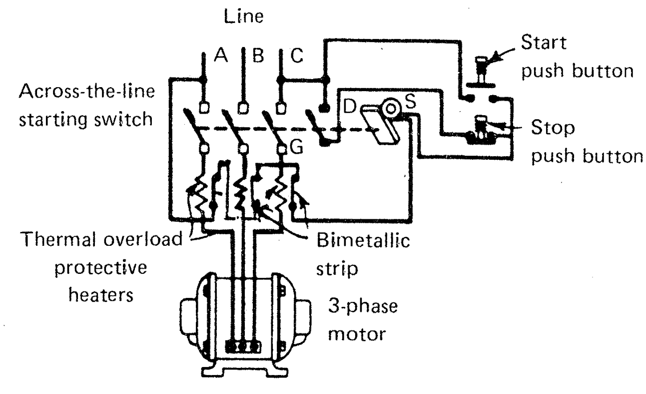
In Fig. 15.1.71*a* is shown an **"across-the-line" starter** which may be operated from different push-button stations. The START push button closes the solenoid circuit between phases *C* and *A* through three bimetallic strips in series. This energizes solenoid *S*, which attracts armature *D*, which in turn closes the starting switch and the auxiliary blade *G*. This blade keeps the solenoid circuit closed when the START push button is released. Pressing the STOP push button opens the solenoid circuit, permitting the starting switch to open. A prolonged heavy overload raises the temperature of the heaters by an amount that will cause at least one of the bimetallic strips to open the solenoid circuit, releasing the starting switch.

A common method of applying reduced-voltage start is to use a **compensator** or **autotransformer** or **autostarter** (Fig. 15.1.71*b*). When the switch is in the starting position, the three windings *AB* of the three-phase autotransformer are connected in Y across the line and the motor terminals are connected to the taps which supply reduced voltage. When the switch is in the running position, the starter is entirely disconnected from the line. In modern practice, motors are protected by thermal or magnetic overload relays (Fig. 15.1.71) which operate to trip the circuit breaker. Since a time element is involved in the operation of such relays, they do not respond to large starting currents, because of their short duration. To limit the current to as low a value as possible, the lowest taps that will give the motor sufficient voltage to supply the required starting torque should be used. As the torque of an induction motor varies as the square of the voltage, the compensator produces a very low starting torque.

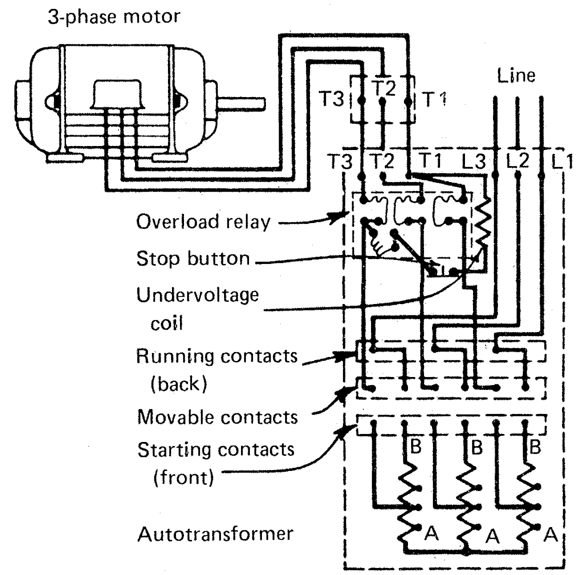
Resistors in series with the stator may also be used to start squirrel-cage motors. They are inserted in each phase and are gradually cut out as the motor comes up to speed. The resistors are generally made of wire-type resistor units or of graphite disks enclosed within heat-resisting porcelain-lined iron tubes. The disadvantage of resistors is that if the motor is started slowly the resistor becomes very hot and may burn out. Resistor starters are less expensive than autotransformers. Their application is to motors that start with light loads at infrequent intervals.

A **phase-controlled,** silicon-controlled rectifier (SCR) may be used to limit the motor-starting current to any value that will provide sufficient starting torque by reducing voltage to the motors. The SCR can also be used to start and stop the motors. A positive opening device such as a contactor or circuit breaker should be used in series with the SCR to stop the motor in case of a failed SCR, which will normally fail shorted and apply full voltage to the motor.

By **introducing resistance into the rotor circuit** through slip rings, the rotor currents may be brought nearly into phase with the air-gap flux and, at the same time, any value of torque up to maximum torque obtained. As the rotor develops speed, resistance may be cut out until there is no external resistance in the rotor circuit. The speed may also be controlled by inserting resistance in the rotor circuit. However, like the armature-resistance method of speed control with shunt motors, this method is also inefficient and gives poor speed regulation. Figure 15.1.72 shows graphically the effect on the torque of applying reduced voltage (curves *b*, *c*) and of inserting resistance in the rotor circuit (curve *d*). As shown by curves *b* and *c*, the torque for any given slip is proportional to the square of the line voltage. The effect of introducing resistance into the rotor circuit is shown by curve *d*. The point of maximum torque is shifted toward higher values of slip. The maximum torque at starting (slip = 1.0) occurs when the rotor resistance is equal to the rotor reactance at standstill. The wound-rotor motor is used where large starting torque is necessary as in railway work, hoists, and cranes. It has better starting characteristics than the squirrel-cage motor, but, because of the necessarily higher resistance of the rotor, it has greater slip even with the rotor resistance all cut out. Obviously, the wound rotor, controller, and external resistance make it more expensive than the squirrel-cage type.



(a) Across-the-line-starter



(b) Autostarter

Fig. 15.1.71 Starters for squirrel-cage induction motor. (*a*) Across-the-line starter; (*b*) autostarter.

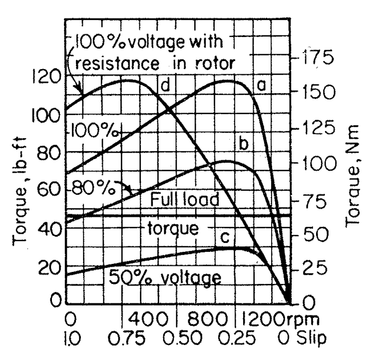


Fig. 15.1.72 Speed-torque curve for 10-hp, 60-Hz, 1,140-r/min induction motor.

One **disadvantage of induction motors** is that they take lagging current, and the power factor at half load and less is low. The speed- and torque-load characteristics of induction motors are almost identical with those of the shunt motor. The speed decreases slightly to full load, the slip being from 10 percent in small motors to 2 percent in very large motors. The torque is almost proportional to the load nearly up to the breakdown torque. The power factor is 0.8 to 0.9 at full load. The direction of rotation of any three-phase motor may be reversed by interchanging any two stator wires.

Speed Control of Induction Motors The induction motor inherently is a constant-speed motor. From Eqs. **(15.1.118)** and **(15.1.119)** the rotor speed is

$$N_2 = 120f(1 - s)/P \qquad (15.1.121)$$

The speed can be changed only by changing the frequency, poles, or slip. In some applications where the motors constitute the only load on the generators, as with electric propulsion of ships, their speed may be changed by changing the frequency. Even then the range is limited, for both turbines and generators must operate near their rated speeds for good efficiency. By employing two distinct windings or by reconnecting a single winding by switching it is possible to change the number of poles. Complications prevent more than two speeds being readily obtained in this manner. Elevator motors frequently have two distinct windings. Another objection to changing the number of poles is the fact that the design is a compromise, and sacrifices of desirable characteristics usually are necessary at both speeds.

The change of slip by introducing resistance into the rotor circuit has been discussed under the wound-rotor motor. It is also possible to introduce an **inverter** into the rotor circuit and convert the slip frequency power to line frequency power and return it to the distribution system.

Inverters are also used to convert line frequencies to variable frequencies to operate squirrel-cage motors at almost any speed up to their mechanical limitations. Inverters also reduce the starting stresses on a motor.

Polyphase voltages should be evenly **balanced** to prevent phase current **unbalance.** If voltages are not balanced, the motor must be derated in accordance with National Electrical Manufacturers Association (NEMA), Publication No. MG-1-14.34.

Approximately 5 percent voltage unbalance would cause about 25 percent increase in temperature rise at full load. The input current unbalance at full load would probably be 6 to 10 times the input voltage unbalance.

Single phasing, one phase open, is the ultimate unbalance and will cause overheating and burnout if the motor is not disconnected from the line.

Single-Phase Induction Motor Single-phase induction motors are usually made in fractional horsepower ratings, but they are listed by NEMA in integral ratings up to 10 hp. They have relatively high rotor resistances and can operate in the

Table 15.1.13 Drip-proof Motors

hp	Weight, lb	Amp	Power factor, percent: ½ load	Power factor, percent: ¾ load	Power factor, percent: Full load	Efficiency, percent*: ½ load	Efficiency, percent*: ¾ load	Efficiency, percent*: Full load
3 phase, 230 V, 60 Hz, 1,750 rpm, squirrel cage								
1	40	3.8	45.3	64.8	66.2	63.8	71.4	75.5
2	45	6.0	54.3	67.6	76.5	75.2	79.9	81.5
5	65	14.2	61.3	73.7	80.7	77.0	80.9	81.5
10	110	26.0	66.3	77.7	83.1	83.5	86.0	86.5
20	190	53.6	69.5	78.4	80.9	85.0	87.0	86.5
40	475	101.2	69.8	79.4	83.6	86.8	88.3	88.5
100	830	230.0	83.7	88.2	89.0	92.2	92.4	91.7
200	1,270	446.0	82.5	87.8	89.2	93.5	94.2	94.1
3 phase, 230 V, 60 Hz, 1,750 rpm, wound rotor								
5	155	15.6	50.7	64.1	74.0	77.6	81.0	81.4
10	220	27.0	66.8	77.5	82.5	84.3	85.4	84.8
25	495	60.0	79.2	86.6	89.4	88.1	88.6	87.8
50	650	122.0	72.4	82.3	86.8	86.7	87.9	87.7
100	945	234.0	78.4	86.2	89.4	90.0	90.4	89.8
200	2,000	446.0	79.8	87.6	90.8	90.5	92.1	92.5
3 phase, 2,300 V, 60 Hz, 1,750 rpm, squirrel cage								
300	2,300	70.4	76.2	83.4	86.0	90.8	92.5	92.8
700	3,380	155.0	85.5	89.4	90.0	91.7	93.4	93.7
1,000	4,345	221.0	86.2	89.5	90.0	92.1	93.8	94.1
3 phase, 2,300 V, 60 Hz, 1,750 rpm, wound rotor								
300	3,900	68	84.4	86.2	89.9	90.8	92.5	92.8
700	5,750	154	82.7	87.7	90.9	91.7	93.4	93.7
1,000	8,450	218	82.9	87.9	91.1	92.1	93.8	94.1

SOURCE: Westinghouse Electric Corp.
*High-efficiency motors are available at premium cost.

single-phase mode without overheating. Single-phase induction motors are not self-starting.

However, the single-phase motor runs in the direction in which it is started. There are several methods of starting single-phase induction motors. Short-circuited turns, or **shading coils,** may be placed around the pole tips which retard the time phase of the flux in the pole tip, and thus a weak torque in the direction of rotation is produced. A high-resistance starting winding, displaced 90 electrical degrees from the main winding, produces poles between the main poles and so provides a rotating field which is weak but is sufficient to start the motor. This is called the **split-phase** method. In order to minimize overheating this winding is ordinarily cut out by a centrifugal device when the armature reaches speed. In the larger motors a repulsion-motor start is used. The rotor is wound like an ordinary dc armature with a commutator, but with short-circuited brushes pressing on it axially rather than radially. The motor starts as a repulsion motor, developing high torque. When it nears its synchronous speed, a centrifugal device pushes the brushes away from the commutator, and at the same time causes the segments to be short-circuited, thus converting the motor into a single-phase induction motor.

Capacitor Motors Instead of splitting the phase by means of a high-resistance winding, it has become almost universal practice to connect a capacitor in series with the auxiliary winding (which is displaced 90 electrical degrees from the main winding). With capacitance, it is possible to make the flux produced by the auxiliary winding lead that produced by the main field winding by 90° so that a true two-phase rotating field results and good starting torque develops. However, the 90° phase relation between the two fields is obtainable at only one value of speed (as at starting), and the phase relation changes as the motor comes up to speed. Frequently the auxiliary winding is disconnected either by a centrifugal switch or a relay as the motor approaches full speed, in which case the motor is called a **capacitor-start** motor. With proper design the auxiliary winding may be left in circuit permanently (frequently with additional capacitance introduced). This improves both the power factor and torque characteristics. Such a motor is called **permanent-split capacitor motor.**

Phase Converter If a polyphase induction motor is operating single-phase, polyphase emfs are generated in its stator by the combination of stator and rotor fluxes. Such a machine can be utilized, therefore, for converting single-phase power into polyphase power and, when so used, is called a **phase converter.** Unless corrective means are utilized, the polyphase emfs at the machine terminals are somewhat unbalanced. The power input, being single-phase and at a power factor less than unity, not only fluctuates but is negative for two periods during each cycle. The power output being polyphase is steady, or nearly so. The cyclic differences between the power output and the power input are accounted for in the kinetic energy stored in the rotating mass of the armature. The armature accelerates and decelerates, but only slightly, in accordance with the difference between output and input. The phase converter is used principally on railway locomotives, since a single trolley wire can be used to deliver single-phase power to the locomotive, and the converter can deliver three-phase power to the three-phase wound-rotor driving motors.

AC Commutator Motors Inherently simple ac motors are not adapted to high starting torques and variable speed. There are a number of types of commutating motor that have been developed to meet the requirement of high starting torque and adjustable speed, particularly with single phase. These usually have been accompanied by compensating windings, centrifugal switches, etc., in order to overcome low power factors and commutation difficulties. With proper compensation, commutator motors may be designed to operate at a power factor of nearly unity or even to take leading current.

One of the simplest of the **single-phase commutator motors** is the ac series railway motor such as is used on the erstwhile New York, New Haven, and Hartford Railroad. It is based on the principle that the torque of the dc series motor is in the same direction irrespective of the polarity of its line terminals. This type of motor must be used on low frequency, not over 25 Hz, and is much heavier and more costly than an equivalent dc motor. The torque and speed curves are almost identical with those of the dc series motor. Unlike most ac apparatus the power factor is highest at light load and decreased with increasing load. Such motors operate with direct current even better than with alternating current. For example, the New Haven locomotives also operate from the 600-V dc third-rail system (two motors in series) from the New York City line (238th St.) into Grand Central Station. (See also Sec. 11.)

On account of difficulties inherent in ac operation such as commutation and high reactance drops in the windings, it is economical to construct and operate such motors only in sizes adaptable to locomotives, the ratings being of the order of 300 to 400 hp. **Universal motors** are small simple series motors, usually of fractional horsepower, and will operate on either direct or alternating current, even at 60 Hz. They are used for vacuum cleaners, electric drills, and small utility purposes.

Synchronous Motor Just as dc shunt generators operate as motors, a synchronous generator, connected across a suitable ac power supply, will operate as a motor and deliver mechanical power. Each conductor on the stator must be passed by a pole of alternate polarity every half cycle so that at constant frequency the rpm of the motor is constant and is equal to

$$N = 120f/P \qquad \text{r/min} \tag{15.1.122}$$

and the speed is independent of the load.

There are two types of synchronous motors in general use: the **slip-ring** type and the **brushless** type. The motor field current is transmitted to the motor by brushes and slip rings on the slip-ring type. On the brushless type it is generated by a shaft mounted exciter and rectified and controlled by shaft mounted static devices. Eliminating the slip rings is advantageous in dirty or hazardous areas.

The synchronous motor has the desirable characteristic that its power factor can be varied over a wide range merely by changing the field excitation. With a weak field the motor takes a lagging current. If the load is kept constant and the excitation increased, the current decreases (Fig. 15.1.73) and the phase difference between voltage and current becomes less until the current is in phase with the voltage and the power factor is unity. The current is then at its minimum value such as I_0, and the corresponding field current is called the **normal excitation.** Further increase in field current causes the armature current to lead and the power factor to decrease. Thus **underexcitation** causes the current to **lag; overexcitation** causes the current to **lead.** The effect of varying the field current at constant values of load is shown by the V-curves (Fig. 15.1.73). Unity power factor occurs at the minimum value of armature current, corresponding to normal excitation. The power factor for any point such as P is I_0/I_1, leading current. Because of its adjustable power factor, the motor is frequently run light merely to improve power factor or to control the voltage at some part of a power system. When so used the motor is called a **synchronous condenser.** The motor may, however, deliver mechanical power and at the same time take either leading or lagging current.

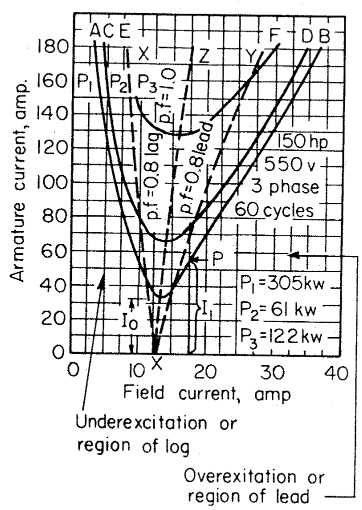


Fig. 15.1.73 V curves of a synchronous motor.

Synchronous motors are used to drive **centrifugal** and **axial compressors,** usually through speed increasers, **pumps, fans,** and other high-horsepower applications where **constant speed, efficiency,** and **power factor** correction are important. **Low-speed** synchronous motors, under 600 r/min, sometimes called **engine** type, are used in driving **reciprocating compressors** and in **ball mills** and in other slow-speed applications. Their low length-to-diameter ratio, because of the need for many poles, gives them a high moment of inertia which is helpful in smoothing the **pulsating torques** of these loads.

If the motor **field current** is separately supported by a **battery** or **constant voltage transformer,** the synchronous motor will maintain speed on a lower voltage dip than will an induction motor because torque is proportional to voltage rather than voltage squared. However, if a synchronous motor drops **out of step,** it will normally not have the ability to reaccelerate the load, unless the driven equipment is automatically unloaded.

The synchronous motor is usually not used in smaller sizes since both the motors and its controls are more expensive than induction motors, and the ability of a small motor to supply VARs to correct power factor is limited.

If situated near an inductive load the motor may be overexcited, and its leading current will neutralize entirely or in part the lagging quadrature current of the load. This reduces the I^2R loss in the transmission lines and also increases the kilowatt ratings of the system apparatus. The **synchronous condenser and motor** can also be used to control voltage and to stabilize power lines. If the condenser or motor is overexcited, its leading current flowing through the line reactance causes a rise in voltage at the motor; if it is underexcited, the lagging current flowing through the line reactance causes a drop in voltage at the motor. Thus within limits it becomes possible to control the voltage at the end of a transmission line by regulating the fields of synchronous condensers or motors. Long 220-kV lines and the 287-kV Hoover Dam–Los Angeles line require several thousand kVa in synchronous condensers floating at their load ends merely for voltage control. If the load becomes small, the voltage would rise to very high values if the synchronous condensers were not underexcited, thus maintaining nearly constant voltage.

A **salient-pole synchronous motor** may be started as an induction motor. In **laminated-pole** machines conducting bars of copper, copper alloy, or aluminum, **damper** or **amortisseur windings** are inserted in the pole face and short-circuited at the ends, exactly as a squirrel-cage winding in the induction motor is connected. The bars can be designed only for starting purposes since they carry no current at synchronous speed and have no effect on efficiency. In solid-pole motors a block of steel is bolted to the pole and performs the current-carrying function of the damper winding in the laminated-pole motor. At times the pole faces are interconnected to minimize starting-pulsating torques. When the synchronous motor reaches 95 to 98 percent speed as an induction motor, the motor field is applied by a **timer** or **slip frequency** control circuit, and the motor pulls into step at 100 percent speed. While accelerating, the motor field is connected to resistances to minimize induced voltages and currents.

Two-pole motors are built as **turbine** type or **round-rotor motors** for mechanical strength and do not have the thermal capacity or space for starting windings, so they must be started by supplementary means.

One such supplementary starting means is the use of a variable-frequency source, either a **variable-speed generator** or more commonly a static **converter-inverter.** The motor is brought up to speed in synchronism with a slowly increasing frequency. One common application is the starting of the large motor-generators used in **pump-storage** utility systems.

Variable frequency may be used to start salient-pole machines also. Requirements for a start without high torques and pulsations or high voltage drops on small electrical systems may dictate the use of something other than full voltage starting.

The **synchronous reluctance motor** is similar to an induction machine with salient poles machined in the rotors. Under light loads the motor will synchronize on reluctance torque and lock in step with the rotating field at synchronous speed. These motors are used in small sizes with variable-frequency inverters for speed control in the paper and textile industry.

The **synchronous-induction motor** is fundamentally a wound-rotor slip-ring induction motor with an air gap greater than normal, and the rotor slots are larger and fewer. On starting, resistance is inserted in the rotor circuit to produce high torque, and this is cut out as the speed increases. As synchronism is approached, the rotor windings are connected to a dc power source and the motor operates synchronously.

Timing or **clock motors** operate synchronously from ac power systems. Figure 15.1.74*a* illustrates the Warren Telechron motor which operates on the hysteresis principle. The stator consists of a laminated element with an exciting coil, and each pole piece is divided, a short-circuited shading turn being placed on each of the half poles so formed. The rotor consists of two or more hard-steel disks of the shape shown, mounted on a small shaft. The shaded poles produce a 3,600 r/min rotating magnetic field (at 60 Hz), and because of hysteresis

Table 15.1.14 Performance Data for Coupled Synchronous Motors

hp	Poles	r/min	A	Excitation, kW	Efficiencies, percent ½ load	¾ load	Full load	Weight, lb
Unity power factor, 3 phase, 60 Hz, 2,300 V								
500	4	1,800	100	3	94.5	95.2	95.3	5,000
2,000	4	1,800	385	9	96.5	97.1	97.2	15,000
5,000	4	1,800	960	13	96.5	97.3	97.5	27,000
10,000	6	1,200	1,912	40	97.5	97.9	98.0	45,000
500	18	400	99.3	5	92.9	93.9	94.3	7,150
1,000	24	300	197	8.4	93.7	94.6	95.0	15,650
4,000	48	150	781	25	94.9	95.6	95.6	54,000
80% power factor, 3 phase, 60 Hz, 2,300 V								
500	4	1,800	127	4.5	93.3	94.0	94.1	6,500
2,000	4	1,800	486	13	95.5	96.1	96.2	24,000
5,000	4	1,800	1,212	21	95.5	96.3	96.5	37,000
10,000	6	1,200	2,405	50	96.8	97.3	97.4	70,000
500	18	400	125	7.2	92.4	93.4	93.6	9,500
1,000	24	300	248	11.6	93.3	94.2	94.4	17,500
4,000	48	150	982	40	94.6	95.3	95.5	11,500

SOURCE: Westinghouse Electric Corp.

loss, the disk follows the field just as the rotor of an induction motor does. When the rotor approaches the synchronous speed of 3,600 r/min, the rotating magnetic field takes a path along the two rotor bars and locks the rotor in with it. The rotor and the necessary train of reducing gears rotate in oil sealed in a small metal can. Figure 15.1.74*b* shows a subsynchronous motor. Six squirrel-cage bars are inserted in six slots of a solid cylindrical iron rotor, and the spaces between the slots form six salient poles. The motor, because of the squirrel cage, starts as an induction motor, attempting to attain the speed of the rotating field, or 3,600 r/min (at 60 Hz). However, when the rotor reaches 1,200 r/min, one-third synchronous speed, the salient poles of the rotor lock in with the poles of the stator and hold the rotor at 1,200 r/min.

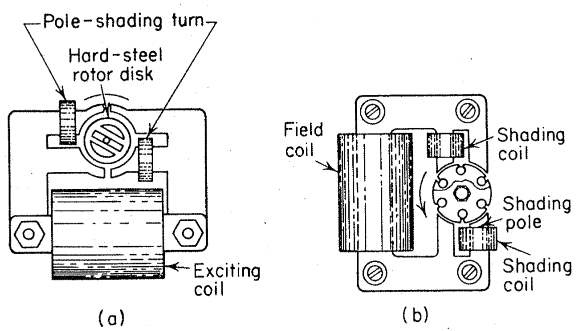


Fig. 15.1.74 Synchronous motors for timing. (*a*) Warren Telechron motor; (*b*) Holtz induction-reluctance subsynchronous motor.

AD-DC CONVERSION

Static Rectifiers

Silicon devices, and to a lesser extent gas tubes, are the primary means of ac to dc or dc to ac conversion in modern installations. They are advantageous when compared to synchronous converters or motor generators because of efficiency, cost, size, weight, and reliability. Various bridge configurations for single-phase and three-phase applications are shown in Fig. 15.1.75*a*. Table 15.1.15 shows the relative outputs of rectifier circuits. The use of two three-phase bridges fed from an ac source consisting of a three-winding transformer with both a Δ and Y secondary winding so that output voltages are 30° out of phase will reduce dc ripple to approximately 1 percent.

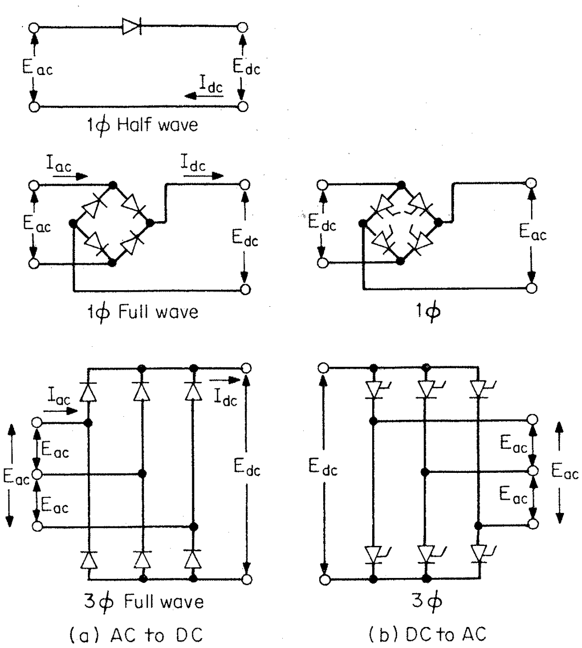


Fig. 15.1.75 AC-DC conversion with static devices.

Table 15.1.15 Relationships for AC-DC Conversion Static Devices

Device description	Voltages, %		Currents, %		Ripple, %
	Eac	Edc	Iac	Idc	
1ϕ-half wave	100	45	100	100	121
1ϕ-full wave	100	90	100	90	48
3ϕ-full wave	100	135	100	123	4.2

The use of **silicon-controlled rectifiers** (SCRs) to replace rectifiers in the various bridge configurations allows the output voltages to be varied from rated output voltage to zero. The output voltage wave will not be a sine wave but a series of square waves, which may not be suitable for some applications. Dc to ac conversions are shown in Fig. 15.1.75*b*.

SYNCHRONOUS CONVERTERS

The synchronous converter is essentially a dc generator with slip rings connected by taps to equidistant points in the armature winding. Alternating current may also be taken from and delivered to the armature. The machine may be single-phase, in which case there are two slip rings and two slip-ring taps per pair of poles; it may be three-phase, in which case there are three slip rings and three slip-ring taps per pair of poles, etc. Converters are usually used to convert alternating to direct current, in which case they are said to be operating **direct;** they may equally well convert direct to alternating current, in which case they are said to be operating **inverted.** A converter will operate satisfactorily as a dc motor, a synchronous motor, a dc generator, a synchronous generator, or it may deliver direct and alternating current simultaneously, when it is called a **double-current generator.**

The rating of a converter increases very rapidly with increase in the number of phases owing, in part, to better utilization of the armature copper and also because of more uniform distribution of armature heating.

Because of the materially increased rating, converters are nearly all operated six-phase. The rating decreases rapidly with decrease in power factor, and hence the converter should operate near unity power factor. The diametrical ac voltage is the ac voltage between two slip-ring taps 180 electrical degrees apart. With a two-pole closed winding, i.e., a winding that closes on itself when the winding is completed, the diametrical ac voltage is the voltage between any two slip-ring taps diametrically opposite each other.

With a sine-voltage wave, the dc voltage is the peak of the diametrical ac voltage wave. The voltage relations for sine waves are as follows: dc volts, 141; single phase, diametrical, 100; three-phase, 87; four-phase, diametrical, 100; four phase, adjacent taps, 71; six phase, diametrical, 100; six phase, adjacent taps, 50. These relations are obtained from the sides of polygons inscribed in a circle having a diameter of 100 V, as shown in Fig. 15.1.76.

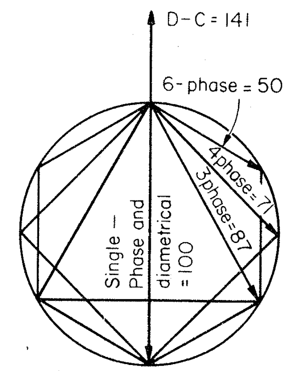


Fig. 15.1.76 EMF relations in converter.

Selsyns The word **selsyn** is an abbreviation of self-synchronizing and is applied to devices which are connected electrically, and in which an angular displacement of the rotating member of one device produces an equal angular displacement in the rotating member of the second device. There are several types of selsyns and they may be dc or ac, single-phase or polyphase. A simple and common type is shown in Fig. 15.1.77. The two stators S_1, S_2 are phase-wound stators, identical electrically with synchronous-generator or induction-motor stators. For simplicity Gramme-ring windings are shown in Fig. 15.1.77. The two stators are connected three-phase and in parallel. There are also two bobbin-type rotors R_1, R_2, with single-phase windings, each connected to a single-phase supply such as 115 Volts, 60 Hz. When R_1 and R_2 are in the same angular

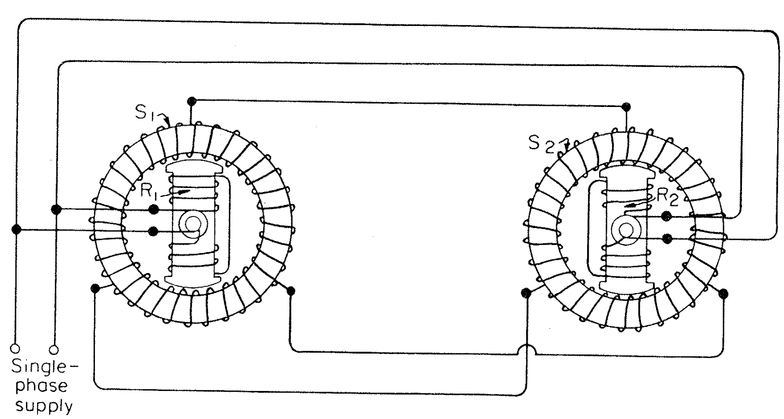


Fig. 15.1.77 Selsyn system.

positions, the emfs induced in the two stators by the ac flux of the rotors are equal and opposite, there are no interchange currents between stators, and the system is in equilibrium. However, if the angular displacement of R_1, for example, is changed, the magnitudes of the emfs induced in the stator winding of S_1 are correspondingly changed. The emfs of the two stators then become unbalanced, currents flow from S_1 to S_2, producing torque on R_2. When R_2 attains the same angular position as R_1, the emfs in the two rotors again become equal and opposite, and the system is again in equilibrium.

If there is torque load on either rotor, a resultant current is necessary to sustain the torque, so that there must be an angular displacement between rotors. However, by the use of an auxiliary selsyn a current may be fed into the system which is proportional to the angular difference of the two rotors. This current will continue until the error is corrected. This is called **feedback.** There may be a master selsyn, controlling several secondary units.

Selsyns are used for position indicators, e.g., in bridge–engine-room signal systems. They are also widely used for fire control so that from any desired position all the turrets and guns on battleships can be turned and elevated simultaneously through any desired angle with a high degree of accuracy. The selsyn itself rarely has sufficient power to perform these operations, but it actuates control through power multipliers such as amplidynes.

RATING OF ELECTRICAL APPARATUS

The **rating** of electrical apparatus is almost always determined by the maximum temperature at which the materials in the machine, especially the insulation and lubricant, may be operated for long periods without deterioration. It is permissible, as far as temperature is concerned, to overload the apparatus so long as the safe temperature is not exceeded. The ANSI/IEEE Standard 100-1977 classifies **insulating materials** in seven different **classes:**

1. *Class 90 insulation.* Materials or combinations of materials such as cotton, silk, and paper without impregnation which will have suitable thermal endurance if operated continually at 90°C.
2. *Class 105 insulation.* Materials or combinations of materials such as cotton, silk, and paper when suitably impregnated or coated or when immersed in a dielectric liquid. This class has sufficient thermal endurance at 105°C.
3. *Class 130 insulation.* Materials or combinations of materials such as mica, glass fiber, asbestos, etc., with suitable bonding substances. This class has sufficient thermal endurance at 130°C.
4. *Class 155 insulation.* Same materials as class 130 but with bonding substances suitable for continuous operation at 155°C.
5. *Class 180 insulation.* Materials or combinations of materials such as silicone elastomer, mica, glass fiber, asbestos, etc., with suitable bonding substances such as appropriate silicone resins. This class has sufficient thermal life at 180°C.
6. *Class 220 insulation.* Materials suitable for continuous operation at 220°C.
7. *Class over-220 insulation.* Materials consisting entirely of mica, porcelain, glass, quartz, and similar inorganic materials which have suitable thermal life at temperatures over 220°C.

NOTE: In all cases, other materials or combinations of materials other than those mentioned above may be used in a given class if from experience or accepted tests they can be shown to have comparable thermal life. It is common practice also to specify **insulation systems in electrical machinery** by letter. For example, integral horsepower ac motors may have a maximum temperature rise in the winding (determined by winding resistance) of 60°C for **class A** insulation, 80°C for **class B**, 105°C for **class F**, and 125°C for **class H**—all based on a 40°C ambient.

The recommended methods of measurement are: (1) the thermometer method is preferred for uninsulated windings, exposed metal parts, gases and liquids, or surface methods generally; thermocouples are preferred for rapidly changing surface temperatures; (2) the applied-thermocouple method is suitable for making surface temperature measurements when it is desired to measure the temperature of surfaces that are accessible to thermocouples but not to liquid-in-glass thermometers; (3) the contact-thermocouple method is suitable for measuring temperatures of bare metal surfaces such as those of commutator bars and slip rings; (4) the resistance method is suitable for insulated windings, except for windings of such low resistance that measurements cannot be accurately made due to uncontrollable resistance in contacts or where it is impracticable to make connections to obtain measurements before an undesirable drop in temperature occurs; (5) the embedded-detector method is suitable for interior measurements at designated locations as specified in the standards for certain kinds of equipment, such as large rotating machines.

Efficiency of Electrical Motors

Methods of determining efficiency are by **direct measurement** or by **segregated losses.** Methods are outlined in Standard Test Procedure for Polyphase Induction Motors and Generators, ANSI/IEEE Std. 112-1978; Standard Test Code for DC Machines, IEEE Std. 113-1973; Test Procedure for Single-Phase Induction Motors, ANSI/IEEE Std. 114-1982; and Test Procedures for Synchronous Machines, IEEE Std. 115-1965.

Direct measurements can be made by using calibrated **motors, generators,** or **dynamometers** for input to generators and output from motors, and precision electrical motors for input to motors and output from generators.

$$\text{Efficiencies} = \frac{\text{output}}{\text{input}} \tag{15.1.123}$$

The segregated **losses** in motors are classified as follows: (1) Stator I^2R (shunt and series field I^2R for dc); (2) rotor I^2R (armature I^2R for dc); (3) core loss; (4) stray-load loss; (5) friction and windage loss; (6) brush-contact loss (wound rotor and dc); (7) brush-friction loss (wound rotor and dc); (8) exciter loss (synchronous and dc); and (9) ventilating loss (dc). Losses are calculated separately and totaled.

Measure the electrical output of the generator; then

$$\text{Efficiency} = \frac{\text{output}}{\text{output} + \text{losses}} \tag{15.1.124a}$$

Measure the electrical input of the motors; then

$$\text{Efficiency} = \frac{\text{input} - \text{losses}}{\text{input}} \tag{15.1.124b}$$

When testing dc motors motors, compensation should be made for the **harmonics** associated with rectified ac used to provide the variable dc voltage to the motors. Instrumentation should be chosen to accurately reflect the rms value of currents.

Temperature rise under full-load conditions may be measured by tests as outlined in the IEEE Standards referred to above. Methods of loading are: (1) Load motor with **dynamometer** or **generator** of similar capacity and run until temperatures stabilize; (2) load generator with **motor-generator** set or **plant** load and run until temperature stabilizes; (3) alternately apply **dual frequences** to motor until it reaches rated temperature; (4) synchronous motor may be operated as synchronous **condenser** at no load with **zero power factor** at rated current, voltage, and frequency until temperatures stabilize.

Industrial Applications of Motors

Alternating or Direct Current The induction motor, particularly the squirrel-cage type, is preferable to the dc motor for constant-speed work, for the initial cost is less and the absence of a commutator reduces maintenance. Also there is less fire hazard in many industries, such as sawmills, flour mills, textile mills, and powder mills. The use of the induction motor in such places as cement mills is advantageous since with dc motors the grit makes the maintenance of commutators difficult.

For variable-speed work like cranes, hoists, elevators, and for adjustable speeds, the dc motor characteristics are superior to induction-motor characteristics. Even then, it may be desirable to use induction motors since their less desirable characteristics are more than balanced by their simplicity and the fact that ac power is available, and to obtain dc power conversion apparatus is usually necessary. Where both lights and motors are to be supplied from the same ac system, the 208/120-V four-wire three-phase system is now in common use. This gives 208 V three-phase for the motors, and 120 V to neutral for the lights.

Full-load speed, temperature rise, efficiency, and power factor as well as breakdown torque and starting torque have long been parameters of concern in the application and purchase of motors. Another qualification is service factor. The service factor of an alternating current motor is a multiplier applicable to the horsepower rating. When so applied, the result is a permissible horsepower loading under the conditions specified for the service factor. When operated at service factor load with 1.15 or higher service factor, the permissible temperature rise by resistance is as follows: class A insulation 70°C, class B, 90°C, and class F, 115°C.

Special enclosures, fittings, seals, ventilation systems, electromagnetic design, etc., are required when the motor is to be operated under unusual service conditions, such as exposure to (1) combustible, explosive, abrasive, or conducting dusts, (2) lint or very dirty conditions where the accumulation of dirt might impede the ventilation, (3) chemical fumes or flammable or explosive gases, (4) nuclear radiation, (5) steam, salt laden air, or oil vapor, (6) damp or very dry locations, radiant heat, vermin infestation, or atmosphere conducive to the growth of fungus, (7) abnormal shock, vibration, or external mechanical loading, (8) abnormal axial thrust or side forces on the motor shaft, (9) excessive departure from rated voltage, (10) deviation factors of the line voltage exceeding 10 percent, (11) line voltage unbalance exceeding 1 percent, (12) situations where low noise levels are required, (13) speeds higher than the highest rated speed, (14) operation in a poorly ventilated room, in a pit, or in an inclined attitude, (15) torsional impact loads, repeated abnormal overloads, reversing or electric braking, (16) operation at standstill with any winding continuously energized, and (17) operation with extremely low structureborne and airborne noise. For dc machines, a further unusual service condition occurs when the average load is less than 50 percent over a 24-h period or the continuous load is less than 50 percent over a 4-h period.

The standard direction of rotation for all nonreversing dc motors, ac single-phase motors, synchronous motors, and universal motors is counterclockwise when facing the end of the machine opposite the drive end. For dc and ac generators, the rotation is clockwise.

Further information may be found in Publication No. MG-1 of the National Electrical Manufacturers Association.

It must be recognized that heat is conducted by electrical conductors. Windings in motors operating in a 40°C ambient at class F temperature rises are running at temperatures 90°C higher than the maximum allowable temperature (75°C) of cable ordinarily used in interior wiring. Heat conducted by the motor leads in such a situation could cause a failure of the branch circuit cable in the terminal box. See Tables 15.1.21 and 15.1.22.

ELECTRIC DRIVES

Cranes and Hoists The dc series motor is best adapted to cranes and hoists. When the load is heavy the motor slows down automatically and develops increased torque thus reducing the peaks on the electrical system. With light loads, the speed increases rapidly, thus giving a lively crane. The series motor is also well adapted to moving the bridge itself and also the trolley along the bridge. Where alternating current only is available and it is not economical to convert it, the slip-ring type of induction motor, with external-resistance speed control, is the best type of ac motor. Squirrel-cage motors with high resistance end rings to give high starting torque (design D) are used (design D motors; also see Ilgner system).

Constant-Torque Applications **Piston pumps, mills, extruders,** and **agitators** may require constant torque over their complete speed range. They may require high starting torque design C or D squirrel-cage motors to bring them up to speed. Where speed is to be varied while running, a variable armature voltage **dc motor** or a **variable-frequency** squirrel-cage induction-motor drive system may be used.

Centrifugal Pumps Low WK^2 and low starting torques make design B general-purpose squirrel-cage motors the preference for this application. When **variable flow** is required, the use of a **variable-frequency** power supply to vary motor speed will be energy efficient when compared to changing flow by **control-valve** closure to increase head.

Centrifugal Fans High WK^2 may require high starting torque design C or D squirrel-cage motors to bring the fan up to speed in a reasonable period of time. When **variable flow** is required, the use of a **variable-frequency** power supply or a multispeed motor to vary fan speed will be energy efficient when compared to closing **louvers.** For large fans, **synchronous-motor** drives may be considered for high efficiency and improved power factor.

Axial or Centrifugal Compressors For smaller compressors, say, up to 100 hp, the squirrel-cage induction motor is the drive of choice. When the WK^2 is high, a design C or D high-torque motor may be required. For larger compressors, the synchronous motor is more efficient and improves power factor. Where **variable flow** is required, the **variable-frequency** power supply to vary motor speed is more efficient than controlling by **valve** and in some applications may eliminate a **gear box** by allowing the motor to run at compressor operating speed.

Pulsating-Torque Applications **Reciprocating compressors, rock crushers,** and **hammer mills** experience widely varying torque pulsations during each revolution. They usually have a **flywheel** to store energy, so a high-torque, high-slip design D motor will accelerate the high WK^2 rapidly and allow energy recovery from the flywheel when high torque is demanded. On larger drives a slow-speed, engine-type synchronous motor can be directly connected. The motor itself supplies significant WK^2 to smooth out the torque and current pulsations of the system.

SWITCHBOARDS

Switchboards may, in general, be divided into four classes: direct-control panel type; remote mechanical-control panel type; direct-control truck type; electrically operated. With **direct-control panel-type boards** the switches, rheostats, bus bars, meters, and other apparatus are mounted on or near the board and the switches and rheostats are operated directly, or by operating handles if they are mounted in back of the board. The voltages, for both direct current and alternating current, are usually limited to 600 V and less but may operate up to 2,500 V ac if oil circuit breakers are used. Such panels are not recommended for capacities greater than 3,000 kVA. **Remote mechanical-control panel-type boards** are ac switchboards with the bus bars and connections removed from the panels and mounted separately away from the load. The oil circuit breakers are operated by levers and rods. This type of board is designed for heavier duty than the direct-control type and is used up to 25,000 kVA. **Direct-control truck-type switchboards** for 15,000 V or less consist of equipment enclosed in steel compartments completely assembled by the manufacturers. The high-voltage parts are enclosed, and the equipment is interlocked to prevent mistakes in operation. This equipment is designed for low- and medium-capacity plants and auxiliary power in large generating stations. **Electrically operated switchboards** employ solenoid or motor-operated circuit breakers, rheostats, etc., controlled by small switches mounted on the panels. This makes it possible to locate the high-voltage and other equipment independently of the location of switchboard.

In all large stations the switching equipment and buses are always mounted entirely either in separate buildings or in outdoor enclosures. Such equipment is termed **bus structures** and is electrically operated from the main control board.

Marble has high dielectric qualities and was formerly used exclusively for the panels. It is now used occasionally where its appearance is desired for architectural purposes. Slate is used extensively and is finished in black enamel, marine, and natural black. Ebony asbestos is also used frequently, is lighter than marble or slate, has high dielectric strength and insulation resistivity, and can be readily cut, drilled, and machined. Steel panels, usually ⅛ in thick, are light, economical in construction and erection, and at the present time are favored over other types.

Switchboards should be erected at least 3 or 4 ft from the wall. Switchboard frames and structures should be grounded. The only exceptions are effectively insulated frames of single-polarity dc switchboards. For low-potential work, the conductors on the rear of the switchboard are usually made up of flat copper strip, known as **bus-bar** copper. The size required is based upon a current density of about 1,000 A/in^2. Figure 15.1.78 gives the approximate continuous dc carrying capacity of copper bus bars for different arrangements and spacings for 35°C temperature rise.

Switchboards must be individually adapted for each specific electrical system. Space permits the showing of the diagrams of only three boards each for a typical electrical system (Fig. 15.1.79). Aluminum bus bars are also frequently used.

Equipment of Standard Panels Following are enumerated the various parts required in the equipment of standard panels for varying services:

Generator or synchronous-converter panel, dc two-wire system: 1 circuit breaker; 1 ammeter; 1 handwheel for rheostat; 1 voltmeter; 1 main switch (three-pole single throw or double throw) or 2 single-pole switches.

Generator or synchronous-converter panel, dc three-wire system: 2 circuit breakers; 2 ammeters; 2 handwheels for field rheostats; 2 field switches; 2 potential receptacles for use with voltmeter; 3 switches; 1 four-point starting switch.

Generator or synchronous-motor panel, three-phase three-wire system: 3 ammeters; 1 three-phase wattmeter; 1 voltmeter; 1 field ammeter; 1 double-pole field switch; 1 handwheel for field rheostats; 1 synchronizing receptacle (four-point); 1 potential receptacle (eight-point); 1 field rheostat; 1 triple-pole oil switch; 1 power-factor indicator; 1 synchronizer; 2 series transformers; 1 governor control switch.

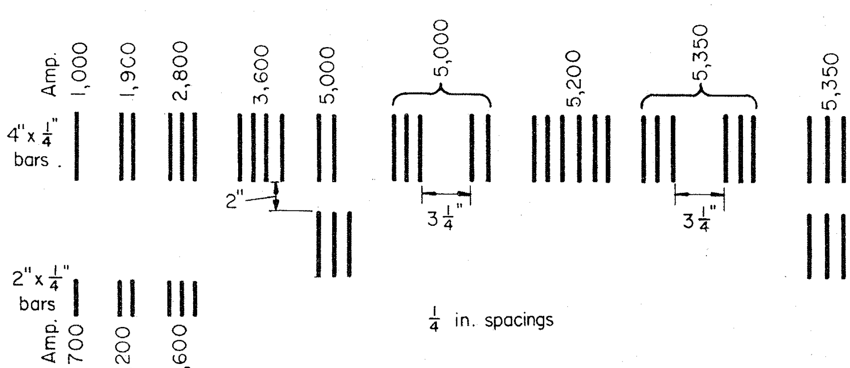


Fig. 15.1.78 Current-carrying capacity of copper bus bars.

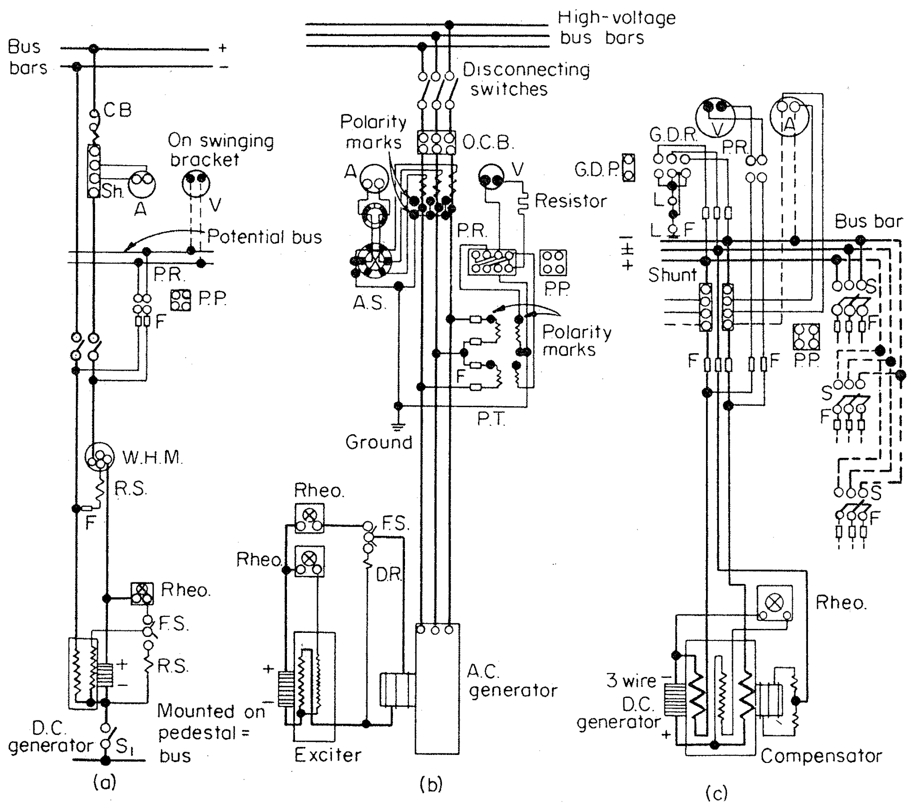


Fig. 15.1.79 Switchboard wiring diagrams for generators. (*a*) 125-V or 250-V dc generator; (*b*) three-phase, synchronous generator and exciter for small or isolated plant; (*c*, three-wire dc generator for small or isolated plant. A, ammeter; AS, three-way ammeter switch; CB, circuit breaker; CT, current transformer; DR, ground detector receptacle; L, ground detector lamp; OC, overload coil; OCB, oil circuit breaker; PP, potential ring; PR, potential receptacle; PT, potential transformer; Rheo, rheostat; RS, resistor; S, switch; Sh, shunt; V, voltmeter; WHM, watthour meter.

Synchronous-converter panel, three phase: 1 ammeter; 1 power-factor indicator; 1 synchronizing receptacle; 1 triple-pole oil circuit breaker; 2 current transformers; 1 potential transformer; 1 watthour meter (polyphase); 1 governor control switch.

Induction motor panel, three-phase: 1 ammeter; series transformers; 1 oil switch.

Feeder panel, dc, two-wire and three-wire: 1 single-pole circuit breaker; 1 ammeter; 2 single-pole main switches; potential receptacles (1 four-point for two-wire panel; 1 four-point and 1 eight-point for three-wire panel).

Feeder panel, three-wire, three-phase and single-phase: 3 ammeters; 1 automatic oil switch (three-pole for three-phase, two-pole for single-phase); 2 series transformers; 1 shunt transformer; 1 wattmeter; 1 voltmeter; 1 watthour meter; 1 handwheel for control of potential regulator.

Exciter panel (for 1 or 2 exciters): 1 ammeter (2 for 2 exciters); 1 field rheostat (2 for 2 exciters); 1 four-point receptacle (2 for 2 exciters); 1 equalizing rheostat for regulator.

Switches The current-carrying parts of switches are usually designed for a current density of 1,000 A/in^2. At contact surfaces, the current density should be kept down to about 50 A/in^2.

Circuit Breakers Switches equipped with a tripping device constitutes an elementary load interrupter switch. The difference between a load interrupter switch and a circuit breaker lies in the interrupting capacity. A circuit breaker must open the circuit successfully under short circuit conditions when the current through the contacts may be several orders of magnitude greater than the rated current. As the circuit is being opened, the device must withstand the accompanying mechanical forces and the heat of the ensuing arc until the current is permanently reduced to zero.

The opening of a metallic circuit while carrying electric current causes an electric arc to form between the parting contacts. If the action takes place in air, the air is ionized (a plasma is formed) by the passage of current. When ionized, air becomes an electric conductor. The space between the parting contacts thus has relatively low voltage drop and the region close to the surface of the contacts has relatively high voltage drop. The thermal input to the contact surfaces (VI) is therefore relatively large and can be highly destructive. A major aim in circuit breaker design is to quench the arc rapidly enough to keep the contacts in a reusable state. This is done in several ways: (1) lengthening the arc mechanically, (2) lengthening the arc magnetically by driving the current-carrying plasma sideways with a magnetic field, (3) placing barriers in the arc path to cool the plasma and increase its length, (4) displacing and cooling the plasma by means of a jet of compressed air or inert gas, and (5) separating the contacts in a vacuum chamber.

By a combination of shunt and series coils the circuit breaker can be made to trip when the energy reverses. Circuit breakers may trip unnecessarily when the difficulty has been immediately cleared by a local breaker or fuse. In order that service shall not be thus interrupted unnecessarily, **automatically reclosing breakers** are used. After tripping, an automatic mechanism operates to reclose the breaker. If the short circuit still exists, the breaker cannot reclose. The breaker attempts to reclose two or three times and then if the short circuit still exists it remains permanently locked out.

Metal-clad switch gears are highly developed pieces of equipment that combine buses, circuit breakers, disconnecting devices, controlling devices, current and potential transformers, instruments, meters, and interlocking devices, all assembled at the factory as a single unit in a compact steel enclosing structure. Such equipment may comprise truck-type circuit breakers, assembled as a unit, each housed in a separate steel compartment and mounted on a small truck to facilitate removal for inspection and servicing. The equipment is interlocked to prevent mistakes in operation and in the removal of the unit; the removal of the unit breaks all electrical connections by suitable disconnecting switches in the rear of the compartment, and all metal parts are grounded. This design provides compactness, simplicity, ease of inspection, and safety to the operator.

High-voltage circuit breakers can be oil type, in which the contacts open under **oil, air-blast** type, in which the arc is extinguished by a powerful blast of air directed through an orifice across the arc and into an arc chute, H_2S type, or vacuum contact type. The tripping of high-voltage circuit breakers is initiated by an abnormal current acting through the secondary of a current transformer on an inverse-time relay in which the time of closing the relay contacts is an inverse time function of the current; i.e., the greater the current the shorter the time of closing. The breaker is tripped by a dc tripping coil, the dc circuit being closed by the relay contacts. Modern circuit breakers should open the circuit within 6 cycles from the time of the closing of the relay contacts.

Air-blast circuit breakers have received wide acceptance in all fields in recent years, both for indoor work and for outdoor applications. Indoor breakers are available up to 40 kV and interrupting capacities up to 2.5 GVA. Outdoor breakers are available in ratings up to that of the EHV (extra-high voltage) 765-kV three-pole breaker capable of interrupting 55 GVA, or 40,000-A symmetrical current. Its operating rating is 3,000 A, 765 kV. The arc is extinguished by a blast of air. Switching stations, gas-insulated and operating at 550 kV, are also in use.

POWER TRANSMISSION

Power for long-distance transmission is usually generated at 6,600, 13,200 and 18,000 V and is stepped up to the transmission voltage by Δ-Y-connected transformers. The transmission voltage is roughly 1,000 V/mi. Preferred or standard transmission voltages are 22, 33, 44, 66, 110, 132, 154, 220, 287, 330, 500, and 765 kV. High-voltage lines across country are located on private rights of way. When they reach urban areas, the power must be carried underground to the substations which must be located near the load centers in the thickly settled districts. In many cases it is possible to go directly to underground cables since these are now practicable up to 345 kV between three-phase line conductors (200 kV to ground). High-voltage cables are expensive in both first cost and maintenance, and it may be more economical to step down the voltage before transmitting the power by underground cables. Within a city, alternating current may be distributed from a substation at 13,200, 6,600, or 2,300 V, being stepped down to 600, 480, and 240 V, three-phase for power and 240 to 120 V single-phase three-wire for lights, by transformers at the consumers' premises. **Direct current** at 1,200 or 600 V for railways, 230 to 115 V for lighting and power, is supplied by motor-generator sets, synchronous converters, and rectifiers. **Constant current** for series street-lighting systems is obtained through constant-current transformers.

Transmission Systems

Power is almost always transmitted three-phase. The following fundamental relations apply to any transmission system. The weight of conductor required to transmit power by any given system with a given percentage power loss varies directly with the power, directly as the square of the distance, and inversely as the square of the voltage. The cross-sectional area of the conductors with a given percentage power loss varies directly with the power, directly with the distance, and inversely as the square of the voltage.

For two systems of the same length transmitting the same power at different voltages and with the same power loss for both systems, the cross-sectional area and weight of the conductors will vary inversely as the square of the voltages. The foregoing relations between the cross section or weight of the conductor and transmission distance and voltage hold for all systems, whether dc, single-phase, three-phase, or four-phase. With the power, distance, and power loss fixed, all symmetrical systems having equal voltages to neutral require equal weights of conductor. Thus, the three symmetrical systems shown in Fig. 15.1.80 all deliver the same power, have the same power loss and equal voltages to neutral, and the transmission distances are all assumed to be equal. They all require the same weight of conductor since the weights are inversely proportional to all resistances. (No actual neutral conductor is used.) The respective power losses are (1) $2I^2R$ W; (2) $3(2I/3)^2(3R/2) = 2I^2R$ W; (3) $4(I/2)^2(2R) = 2I^2R$ W, which are all equal.

Size of Transmission Conductor Kelvin's law states, "The most economical area of conductor is that for which the annual cost of energy wasted is equal to the interest on that portion of the capital outlay which can be considered proportional to the weight of copper used." In Fig. 15.1.81 are shown the annual interest cost, the annual cost of I^2R loss, and the total cost as functions of circular mils cross section for both typical overhead conductors and three-conductor cables. Note that the total-cost curves have very flat minimums, and usually other factors such as the character of the load and the voltage regulation, are taken into consideration.

In addition to resistance, overhead power lines have inductive reactance to alternating currents. The inductive reactance

$$X = 2\pi f\{80 + 741.1 \log [(D - r)/r]\} 10^{-6} \quad \Omega/\text{conductor mile} \qquad \textbf{(15.1.126)}$$

where f = frequency, D = distance between centers of conductors (in), and r their radius (in). Table 15.1.16 gives the

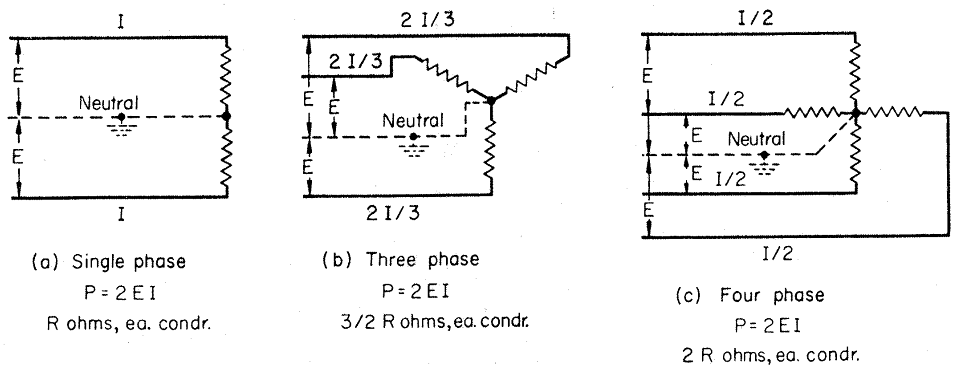


Fig. 15.1.80 Three equivalent symmetrical transmission, or distribution, systems. (*a*) Single-phase (*b*) three-phase; (*c*) four-phase.

inductive reactance per mile at 60 Hz and the resistance of stranded and solid copper conductor. (See Table 15.1.20.)

Any **symmetrical system** having n conductors can be divided into n equal single-phase systems, each consisting of one wire

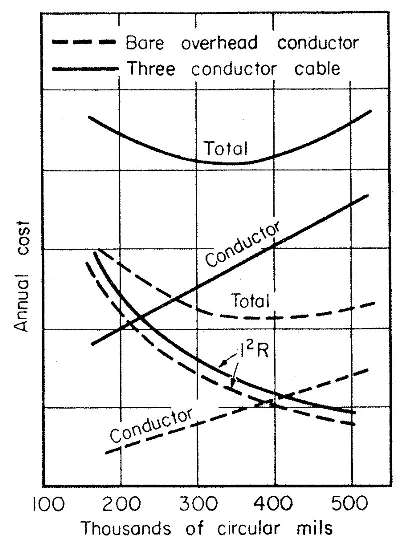


Fig. 15.1.81 Most economical sizes of overhead and underground conductors.

and a return circuit of zero impedance and each having as its voltage the system voltage to neutral.

Figure 15.1.82 shows a symmetrical three-phase system, with one phase detached. The load or received voltage between line conductors is E'_R so that the receiver voltage to *neutral* is $E_R = E'_R/\sqrt{3}$V. The current is I A, the load power factor is $\cos\theta$, and the line resistance and reactance are R and X Ω per wire, and the sending-end voltage is E_S. The phasor diagram is shown in Fig. 15.1.83 (compare with Fig. 15.1.63). Its solution is

$$E_S = \sqrt{(E_R\cos\theta + IR)^2 + (E_R\sin\theta + IX)^2} \qquad \textbf{(15.1.127)}$$

[see Eq. **(15.1.101)**].

Figure 15.1.84 (Mershon diagram) shows the right-hand portion of Fig. 15.1.83 plotted to large scale, the arc 00 corresponding to the arc *ab* (Fig. 15.1.83). The abscissa 0 (Fig. 15.1.84) corresponds to point *b* (Fig. 15.1.83) and is the load voltage E_R taken as 100 percent. The concentric circular arcs 0–40 are given in percentage of E_R. To find the sending-end voltage E_S for any power factor $\cos\theta$, compute first the resistance drop IR and the reactance drop IX in percentage of E_R. Then follow the ordinate corresponding to the load power factor to the inner arc 00 (*a*, Fig. 15.1.83). Lay off the percentage IR drop horizontally to the right, and the percentage IX drop vertically upward. The arc at which the IX drop terminates (*c*, Fig. 15.1.83) when added to 100 percent gives the sending-end voltage E_S in percent of the load voltage E_R.

Example. Let it be desired to transmit 20,000 kW three-phase 80 percent power factor lagging current, a distance of 60 mi. The voltage at the receiving end is 66,000 V, 60 Hz and the line loss must not exceed 10 percent of the power delivered. The conductor spacing must be 7 ft (84 in). Determine the sending-end voltage and the actual efficiency. $I = 20{,}000{,}000/(66{,}000 \times 0.80 \times \sqrt{3}) = 218.8$ A. $3 \times 218.8^2 \times R' = 0.10 \times 20{,}000{,}000$. $R' = 13.9\ \Omega = 0.232\ \Omega$/mi. By referring to Table 13.1.16, 250,000 cir mils copper having a resistance of 0.2278 Ω/mi may be used. The total resistance $R = 60 \times 0.2278 = 13.67\ \Omega$. The reactance $X = 60 \times 0.723 = 43.38\ \Omega$. The volts to neutral at the load, $E_R = 66{,}000/\sqrt{3} = 38{,}100$ V. $\cos\theta = 0.80$; $\sin\theta = 0.60$. Using Eq. **15.1.127**, $E_S = \{[(38{,}100 \times 0.80) + (218.8 \times 13.67)]^2 + [(38{,}100 \times 0.60) + (218.8 \times 43.38)]^2\}^{1/2} = 46{,}500$ V to neutral or $\sqrt{3} \times 46{,}500 = 80{,}500$ between lines at the sending end. The line loss is $3(218.8)^2 \times 13.67 = 1963$ kW. The efficiency $\eta = 20{,}000/(20{,}000 + 1963) = 0.911$, or 91.1 percent. This same line is solved by means of the Mershon diagram as follows. Let $E_R = 38{,}100\ \text{V} = 100$ percent. $IR = 218.8 \times 13.67 = 2{,}991\ \text{V} = 7.85$ percent. $IX = 218.8 \times 43.38 = 9{,}490\ \text{V} = 24.9$ percent. Follow the 0.80 power-factor ordinate (Fig. 15.1.84) to its intersection with the arc 00; from this point go 7.85 percent horizontally to the right and then 24.9 percent vertically. (These percentages are measured on the horizontal scale.) This last distance terminates on the 22.5 percent arc. The sending-end voltage to neutral is then $1.225 \times 38{,}100 = 46{,}500$ V, so that the sending-end voltage between line conductors is $E'_S = 46{,}500\sqrt{3} = 80{,}530$ V.

In Table 15.1.16 the spacing is the distance between the centers of the two conductors of a single-phase system or the distance between the centers of each pair of conductors of a three-phase system if they are equally spaced. If they are not equally spaced, the geometric mean distance *GMD* is used, where $GMD = \sqrt[3]{D_1D_2D_3}$ (Fig. 15.1.85*a*.) With the flat horizontal spacing shown in Fig. 15.1.85*b*, $GMD = \sqrt[3]{2D^3} = 1.26D$.

In addition to copper, aluminum cable steel-reinforced (ACSR), Table 15.1.17, is used for transmission conductor. For the same resistance it is lighter than copper, and with high voltages the larger diameter reduces corona loss.

Table 15.1.16 Resistance and Inductive Reactance per Single Conductor

Size, cir mils or AWG	No. of strands	O D, in	Ohms per mile	60 Hz Spacing, ft 1	2	3	4	5	6	7	8	10	12	15	20	30
Hard-drawn Copper, Stranded																
500,000	37	0.814	0.1130	0.443	0.527	0.576	0.611	0.638	0.660	0.679	0.695	0.722	0.745	0.772	0.807	0.856
400,000	19	0.725	0.1426	0.458	0.542	0.591	0.626	0.653	0.675	0.694	0.710	0.737	0.760	0.787	0.822	0.871
300,000	19	0.628	0.1900	0.476	0.560	0.609	0.644	0.671	0.693	0.712	0.728	0.755	0.778	0.805	0.840	0.889
250,000	19	0.574	0.2278	0.487	0.571	0.620	0.655	0.682	0.704	0.723	0.739	0.766	0.789	0.816	0.851	0.900
0000	19	0.528	0.2690	0.497	0.581	0.630	0.665	0.692	0.714	0.733	0.749	0.776	0.799	0.826	0.861	0.917
000	7	0.464	0.339	0.518	0.602	0.651	0.686	0.713	0.735	0.754	0.770	0.797	0.820	0.847	0.882	0.931
00	7	0.414	0.428	0.532	0.616	0.665	0.700	0.727	0.749	0.768	0.784	0.811	0.834	0.861	0.896	0.945
0	7	0.368	0.538	0.546	0.630	0.679	0.714	0.741	0.763	0.782	0.798	0.825	0.848	0.875	0.910	0.959
Hard-drawn Copper, Solid																
0000	..	0.4600	0.264	0.510	0.594	0.643	0.678	0.705	0.727	0.746	0.762	0.789	0.812	0.839	0.874	0.923
000	..	0.4096	0.333	0.524	0.608	0.657	0.692	0.719	0.741	0.760	0.776	0.803	0.826	0.853	0.888	0.937
00	..	0.3648	0.420	0.538	0.622	0.671	0.706	0.733	0.755	0.774	0.790	0.817	0.840	0.867	0.902	0.951
0	..	0.3249	0.528	0.552	0.636	0.685	0.720	0.747	0.769	0.788	0.804	0.831	0.854	0.881	0.916	0.965
1	..	0.2893	0.665	0.566	0.650	0.699	0.734	0.761	0.783	0.802	0.818	0.845	0.868	0.895	0.930	0.979

Until 1966, 345 kV was the highest operating voltage in the United States. The first 500 kV system put into operation (1966) was a 350-mi transmission loop of the Virginia Electric and Power Company; the longest transmission distance was

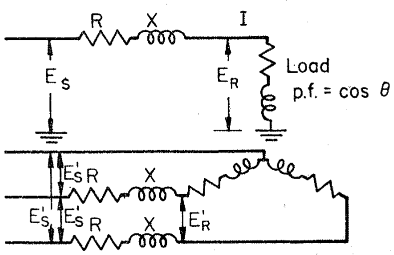


Fig. 15.1.82 Three-phase power system.

170 mi. The towers, about 94 ft high, are of corrosion-resistant steel, and the conductors are 61-strand cables of aluminum alloy, rather than the usual aluminum cable with a steel core (ACSR). The conductor diameter is 1.65 in with two "bundled" conductors per phase and 18-in spacing. The standard

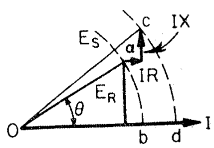


Fig. 15.1.83 Phasor diagram for power line.

span is 1,600 ft, the conductor spacing is flat with 30-ft spacing between phase-conductor centers, and the minimum clearance to ground is 34 to 39 ft. To maintain a minimum clearance of 11 ft to the towers and 30 ft spacing between phases, vee insu-

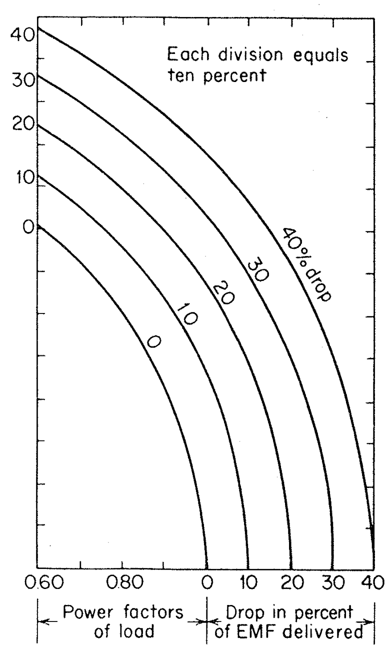


Fig. 15.1.84 Mershon diagram for determining voltage drop in ac power lines.

lator strings, each consisting of twenty-four 10-in disks, are used with each phase. The highest EHV system in North America is the 765-kV system in the midwest region of the United States. A dc transmission line on the west coast of the United States operating at ± 450 kV is transmitting power in bulk more than 800 miles.

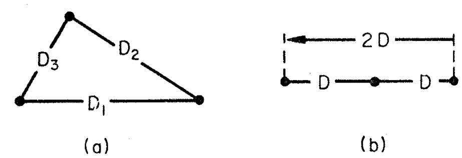


Fig. 15.1.85 Unequal spacing of three-phase conductors. (*a*) GMD $= \sqrt[3]{D_1 D_2 D_3}$; (*b*) flat horizontal spacing; GMD $= 1.26\ D$.

High-voltage dc transmission has a greater potential for savings and a greater ability to transmit large blocks of power longer distances than has three-phase transmission. For the same crest voltage there is a saving of 50 percent in the weight of the conductor. Because of the power stability limit due to inductive and capactive effects (inherent with ac transmission), the ability to transmit large blocks of power long distances has not kept pace with power developments, even at the present highest ac transmission voltage of 765 kV. With direct current there is no such power stability limit.

Where cables are necessary, as under water, the capacitive charging current may, with alternating current become so large that it absorbs a large proportion, if not all, of the cable-carrying capability. For example, at 132 kV, three-phase (76 kV to ground), with a 500 MCM cable, at 36 mi, the charging current at 60 Hz is equal to the entire cable capability so that no capability remains for the load current. With direct current there is no charging current, only the negligible leakage current, and there are no ac dielectric losses. Furthermore, the dc voltage at which a given cable can operate is twice the ac voltage.

The high dc transmission voltage is obtained by converting the ac power voltage to direct current by means of mercury-arc rectifiers; at the receiving end of the line the dc voltage is inverted back to a power-frequency voltage by means of mercury-arc inverters.

Alternating to direct to alternating current is nonsynchronous transmission of electric power. It can be overhead or under the surface. In the early 1970s the problem of bulk electric-power transmission over high-voltage transmission lines above the surface developed the insistent discussion of land use and environmental cost.

While the maximum ac transmission voltage in use in the United States (1977) is 765 kV, ultra-high voltage (UHV) is under consideration. Transmission voltages of 1200 to 2550 kV are being studied. The right-of-way requirements for power transmission are significantly reduced at higher voltages. For example, in one study the transmission of 7500 MVA at 345 kV ac was found to require 14 circuits on a corridor 725 ft (221.5 m) wide, whereas a single 1200-kV ac circuit of 7500-MVA capacity would need a corridor 310 ft (91.5 m) wide.

Continuing research and development efforts may result in the development of more economic high-voltage underground transmission links. Sufficient bulk power transmission capability would permit **power wheeling,** i.e., the use of generating capacity to the east and west to serve a given locality as the

Table 15.1.17 Properties of Aluminum Cable Steel-reinforced (ACSR)

Cir mils or AWG		No. of wires		OD, in	Cross section, in²		Total lb/mi	Ω/mi of single conductor at 25°C				
									200 A		600 A	
Aluminum	Copper equivalent	Aluminum	Steel		Aluminum	Total		0 amp dc	25 Hz	60 Hz	25 Hz	60 Hz
1,590,000	1,000,000	54	19	1.545	1.249	1.4071	10,777	0.0587	0.0589	0.0594	0.0592	0.0607
1,431,000	900,000	54	19	1.465	1.124	1.2664	9,699	0.0652	0.0654	0.0659	0.0657	0.0671
1,272,000	800,000	54	19	1.382	0.9990	1.1256	8,621	0.0734	0.0736	0.0742	0.0738	0.0752
1,192,500	750,000	54	19	1.338	0.9366	1.0553	8,082	0.0783	0.0785	0.0791	0.0787	0.0801
1,113,000	700,000	54	19	1.293	0.8741	0.9850	7,544	0.0839	0.0841	0.0848	0.0843	0.0857
1,033,500	650,000	54	7	1.246	0.8117	0.9170	7,019	0.0903	0.0906	0.0913	0.0908	0.0922
954,000	600,000	54	7	1.196	0.7493	0.8464	6,479	0.0979	0.0980	0.0985	0.0983	0.0997
874,500	550,000	54	7	1.146	0.6868	0.7759	5,940	0.107	0.107	0.108	0.107	0.109
795,000	500,000	26	7	1.108	0.6244	0.7261	5,770	0.117	0.117	0.117	0.117	0.117
715,500	450,000	54	7	1.036	0.5620	0.6348	4,859	0.131	0.131	0.133	0.131	0.133
636,000	400,000	54	7	0.977	0.4995	0.5642	4,319	0.147	0.147	0.149	0.147	0.149
556,500	350,000	26	7	0.927	0.4371	0.5083	4,039	0.168	0.168	0.168	0.168	0.168
477,000	300,000	26	7	0.858	0.3746	0.4357	3,462	0.196	0.196	0.196	0.196	0.196
397,500	250,000	26	7	0.783	0.3122	0.3630	2,885	0.235	0.235	0.235	0.235	0.235
336,400	0000	26	7	0.721	0.2642	0.3073	2,442	0.278	0.278	0.278	0.278	0.278
266,800	000	26	7	0.642	0.2095	0.2367	1,936	0.350	0.350	0.350	0.350	0.350
0000	00	6	1	0.563	0.1662	0.1939	1,542	0.441	0.443	0.446	0.447	0.464
000	0	6	1	0.502	0.1318	0.1537	1,223	0.556	0.557	0.561	0.562	0.579
00	1	6	1	0.447	0.1045	0.1219	970	0.702	0.703	0.707	0.706	0.718
0	2	6	1	0.398	0.0829	0.0967	769	0.885	0.885	0.889	0.887	0.893

SOURCE: Aluminum Co. of America.

earth revolves and the area of peak demand glides across the countryside.

Corona is a reddish-blue electrical discharge which occurs when the voltage-gradient in air exceeds 30 kV peak, 21.1 kV rms, at 76 cm pressure. This electrical discharge is caused by ionization of the air and becomes more or less concentrated at irregularities on the conductor surface and on the outer strands of stranded conductors. Corona is accompanied by a hissing sound; it produces ozone and, in the presence of moisture, nitrous acid. On high-voltage lines corona produces a substantial power loss, corrosion of the conductors, and radio and television interference. The fair-weather loss increases as the square of the voltage above a critical value e_0 and is greatly increased by fog, smoke, rainstorms, sleet, and snow (see Fig.

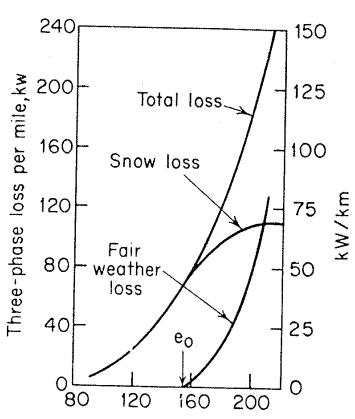


Fig. 15.1.86 Corona loss with snowstorm.

15.1.86). To reduce corona, the diameter of high-voltage conductors is increased to values much greater than would be required for the necessary conductance cross section. This is accomplished by the use of hollow, segmented conductors and by the use of aluminum cable, steel-reinforced (ACSR), which often has inner layers of jute to increase the diameter. In extra-high-voltage lines (400 kV and greater), corona is reduced by the use of bundled conductors in which each phase consists of two or three conductors spaced about 16 in (0.41 m) from one another.

Underground Power Cables

Insulations for power cables include heat-resisting, low-water-absorptive synthetic rubber compounds, varnished cloth, impregnated paper, cross-linked polyethylene thermosetting compounds, and thermoplastics such as polyvinyl chloride (PVC) and polyethylene (PE) compounds (see Sec. 6).

Properly chosen **rubber-insulated cables** may be used in wet locations with a nonmetallic jacket for protective covering instead of a metallic sheath. Commonly used jackets are flame-resisting, such as neoprene and PVC. Such cables are relatively light in weight, easy to train in ducts and manholes, and easily spliced. When distribution voltages exceed 2,000 V phase to phase, an ozone-resisting type of compound is required. Such rubber insulation may be used in cables carrying up to 28,000 V between lines in three-phase grounded systems. The insulation wall will be thicker than with varnished cloth, polyethylene, or paper.

Varnished-cloth cables are made by applying varnish-treated closely woven cloth in the form of tapes, helically, to the metallic conductor. Simultaneously a viscous compound is applied between layers which fills in any voids at laps in the taping and imparts flexibility when the cable is bent by permitting move-

Table 15.1.18 Ampacities of Insulated Cables in Underground Raceways

[Based on conductor temperature of 90°C, ambient earth temperature of 20°C, 100 percent load factor, thermal resistance (RHO) of 90, and three circuits in group]

Conductor size, AWG or MCM	Copper				Aluminum			
	3-1/C cables per raceway		1-3/C cable per raceway		3-1/C cables per raceway		1-3/C cable per raceway	
	2,001–5,000-V ampacity	5,001–35,000-V ampacity	2,001–5,000-V ampacity	5,001–35,000-V ampacity	2,001–5,000-V ampacity	5,001–35,000-V ampacity	2,001–5,000-V ampacity	5,001–35,000-V ampacity
8	56		53		44		41	
6	73	77	69	75	57	60	54	59
4	95	99	89	97	74	77	70	75
2	125	130	115	125	96	100	90	100
1	140	145	135	140	110	110	105	110
1/0	160	165	150	160	125	125	120	125
2/0	185	185	170	185	145	145	135	140
3/0	210	210	195	205	160	165	155	160
4/0	235	240	225	230	185	185	175	180
250	260	260	245	255	205	200	190	200
350	315	310	295	305	245	245	230	240
500	375	370	355	360	295	290	280	285
750	460	440	430	430	370	355	345	350
1,000	525	495	485	485	425	405	400	400

NOTE: This is a general table. For other temperatures and installation conditions, see NFPA 70, 1984, National Electrical Code, Tables 77 through 80 and associated notes.

ment of one tape upon another. This type of insulation has higher dielectric loss than impregnated paper but is suitable for the transmission of power up to 28,000 V between phases over short distances. Such insulated cables may be used in dry locations with flame-resisting fibrous braid, reinforced neoprene tape, or PVC jacket and are often further protected with an interlocked metallic tape armor; but in wet locations these cables should be protected by a continuous metallic sheath such as lead or aluminum. Since varnish-cloth-insulated cable has high ozone resistance, heat resistance, and impulse strength, it is well adapted for station or powerhouse wiring or for any service where the temperature is high or where there are sudden increases in voltage for short periods. Since the varnish is not affected by mineral oils, such cables make excellent leads for transformers and oil switches.

PVC is readily available in several fast, bright colors and is often chosen for color-coded multiconductor control cables. It has inherent flame and oil resistance, and as single conductor wire and cable with the proper wall thickness for a particular application, it usually does not need any outside protective covering. On account of its high dielectric constant and high power factor, its use is limited to low voltages, i.e., under 1,000 V, except for series lighting circuits.

Polyethylene, because of its excellent electrical characteristics, first found use when it was adapted especially for high-frequency cables used in radio and radar circuits; for certain telephone, communication, and signal cables; and for submarine cables. Submarine telephone cables with built-in repeaters laid first in the Atlantic Ocean and then in the Pacific are insulated with polyethylene. Because of polyethylene's thermal characteristics, the standard maximum conductor operating temperature is 75°C. It is commonly used for power cables (including large use for underground residential distribution), with transmissions up to 15,000 V. Successful installations have been in service at 46 kV and some at 69 kV. The upper limit has not been reached, inasmuch as work is in progress on higher-voltage polyethylene power cables as a result of advancements in the art of compounding.

Cross-linked polyethylene is another insulation which is gaining in favor in the process field. For power cable insulations, the cross-linking process is most commonly obtained chemically. It converts polyethylene from a thermoplastic into a thermosetting material; the result is a compound with a unique combination of properties, including resistance to heat and oxidation, thus permitting an increase in maximum conductor operating temperature to 90°C. Since the service record with this compound has been good at voltages which have been gradually increased to 15 kV, it is expected that its voltage range also will be extended in the future.

Impregnated-paper insulation is used for very-high-voltage cables whose range has been extended to 345 kV. To eliminate the detrimental effects of moisture and to maintain proper impregnation of the paper, such cables must have a continuous metallic sheath such as lead or aluminum or be enclosed within a steel pipe; the operation of the cable depends absolutely on the integrity of that enclosure. In three-conductor belted-type cables the individual insulated conductors are surrounded by a belt or wall of impregnated paper over which the lead sheath is applied. When all three conductors are within one sheath, their inductive effects practically neutralize one another and eddy-current loss in the sheath is negligible. In the type-H cable, each of the individual conductors is surrounded with a perforated metallic covering, either aluminum foil backed with a paper tape or thin perforated metal tapes wound over the paper. All three conductors are then enclosed within the metal sheath. The metallic coverings being grounded electrically, each conductor acts as a single-conductor cable. This construction eliminates "tangential" stresses within the insulation and reduces pockets or voids. When paper tapes are wound on the conductor, impregnated with an oil or a petrolatum compound, and covered with a lead sheath, they are called **solid type.**

Three-conductor cables are now operating at 33,000 V, and single-conductor cables at 66,000 V between phases (38,000 V to ground). In New York and Chicago, special hollow-conductor oil-filled single-conductor cables are operating successfully at 132,000 V (76,000 V to ground). In France, cables are operating at 345 kV between conductors.

Other methods of installing underground cables are to draw them into steel pipes, usually without the sheaths, and to fill the pipes with oil under pressure **(oilstatic)** or **nitrogen** under 200 lb pressure. The ordinary medium-high-voltage underground cables are usually drawn into duct lines. With a straight run and ample clearance the length of cable between manholes may reach 600 to 1,000 ft. Ordinarily, the distance is more nearly 400 to 500 ft. With bends of small radius the distance must be further reduced.

Cable ratings are based on the permissible operating temperatures of the insulation and environmental installation conditions.

POWER DISTRIBUTION

Distribution Systems The choice of the system of power distribution is determined by the type of power that is available and by the nature of the load. To transmit a given power over a given distance with a given power loss (I^2R), the weight of conductor varies inversely as the square of the voltage. Incandescent lamps will not operate economically at voltages much higher than 120 V; the most suitable voltages for dc motors are 230, 500, and 550 V; for ac motors, standard voltages are 230, 460, and 575 V, three-phase. When power for lighting is to be distributed in a district where the consumers are relatively far apart, alternating current is used, being distributed at high voltage (2,400, 4,160, 4,800, 6,900, and 13,800 V) and transformed at the consumer's premises, or by transformers on poles or located in manholes or vaults under the street or sidewalks, to 240/120 V three-wire for lighting and domestic customers, and to 208, 240, 480, and 600 volts, three-phase, for power.

The first central station power systems were built with dc generation and distribution. The economical transmission distance was short. Densely populated, downtown areas of cities were therefore the first sections to be served. Growth of electric service in the United States was phenomenal in the last two decades of the nineteenth century. After 1895 when ac generation was selected for the development of power from Niagara Falls, the expansion of dc distribution diminished. The economics overwhelmingly favored the new ac system.

Direct current service is still available in small pockets in some cities. In those cases, ac power is generated, transmitted, and distributed. The conversion to dc takes place in rectifiers installed in manholes near the load or in the building to be served. Some dc customers resist the change to ac service because of their need for motor speed control. Elevator and printing press drives and some cloth-cutting knives are examples of such needs. (See Low-Voltage AC Network.)

Series Circuits These **constant-current** circuits were widely used for **street lighting.** The voltage was automatically adjusted to match the number of lamps in series and maintain a constant current. With the advent of **HID lamps** and individual **photocells** on street lighting fixtures which require parallel circuitry, this system fell into disuse and is rarely seen.

Parallel Circuits **Power** is usually distributed at **constant potential,** and all the devices or receivers in the circuit are connected in parallel, giving a constant-potential system, Fig. 15.1.87*a*. If conductors of constant cross section are used and

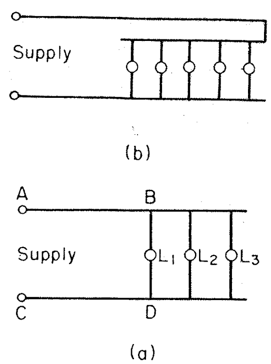


Fig. 15.1.87 (*a*) Parallel circuit; (*b*) loop circuit.

all the loads, L_1, L_2 etc., are operating, there will be a greater voltage *IR* drop per unit length of wire in the portion of the circuit *AB* and *CD* than in the other portions; also the voltage will not be the same for the different lamps but will decrease along the mains with distance from the generating end.

Loop Circuits A more nearly equal voltage for each load is obtained in the loop system, Fig. 15.1.87*b*. The electrical distance from one generator terminal to the other through any receiver is the same as that through any other receiver, and the voltage at the receivers may be maintained more nearly equal, but at the expense of additional conductor material.

Series-Parallel Circuit For incandescent lamps the power must be at low voltage (115 V) and the voltage variations must be small. If the transmission distance is considerable or the loads are large, a large or perhaps prohibitive investment in conductor material would be necessary. In some special cases, lamps may be operated in groups of two in series as shown in Fig. 15.1.88. The transmitting voltage is thus doubled, and, for

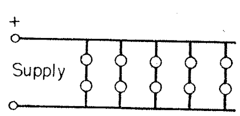


Fig. 15.1.88 Series-parallel system.

a given number of lamps, the current is halved, the permissible voltage drop (*IR*) in conductors doubled, the conductor resistance quadrupled, the weight of conductor material thus being reduced to 25 percent of that necessary for simple parallel operation.

Three-Wire System In the series-parallel system the loads must be used in pairs and both units of the pair must have the same power rating. To overcome these objections and at the same time to obtain the economy in conductor material of operating at higher voltage, the three-wire system is used. It consists merely of adding a third wire or neutral to the system of Fig. 15.1.88 as shown in Fig. 15.1.89.

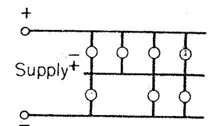


Fig. 15.1.89 Three-wire system.

If the neutral wire is of the same cross section as the two outer wires, this system requires only 37.5 percent the copper required by an equivalent two-wire system. Since the neutral ordinarily carries less current than the outers, it is usually smaller and the ratio of copper to that of the two-wire system is even less than 37.5 percent (see Table 15.1.19).

When the loads on each half of the system are equal, there will be no current in the middle or neutral wire, and the condition is the same as that shown in Fig. 15.1.88. When the loads on the two sides are unequal, there will be a current in the neutral wire equal to the difference of the currents in the outside wires.

AC Three-Wire Distribution Practically all energy for lighting and small motor work is distributed at 1,150, 2,300, or 4,160 V ac to transformers which step down the voltage to 240 and 120 V for three-wire domestic and lighting systems as well as 208, 240, 480, and 600 volts, three-phase, for power. For the three-wire systems the transformers are so designed that the secondary or low-voltage winding will deliver power at 240 V, and the middle or neutral wire is obtained by connecting to the center or midpoint of this winding (see Fig. 15.1.91).

Grounding The neutral wire of the secondary circuit of the transformer should be grounded on the pole (or in the manhole) and at the service switch in the building supplied. If, as a result of a lightning stroke or a fault in the transformer insulation, the transformer primary circuit becomes grounded at *a* (Fig. 15.1.91) and the transformer insulation between primary and secondary windings is broken down at *b* and if there were no permanent ground connection in the secondary neutral wire, the potential of wire 1 would be raised 2,300 V above ground potential. This constitutes a very serious hazard to life for persons coming in contact with the 120 V system. The National Electrical Code requires the use of a ground wire not smaller than 8 AWG copper. With the neutral grounded (Fig. 15.1.91), voltages to ground on the secondary system cannot exceed 120 V. (See National Electrical Code 1984, Art. 250.)

Feeders and Mains Where power is supplied to a large district, improved voltage regulation is obtained by having **centers of distribution.** Power is supplied from the station bus at high voltage to the centers of distribution by large cables known as **feeders.** Power is distributed from the distribution centers to the consumers through the mains and transformed to a usable voltage at the user's site. As there are no loads connected to the feeders between the generating station and centers of distribution, the voltage at the latter points may be maintained constant. **Pilot wires** from the centers of distribution often run back to the station, allowing the operator or automatic controls to

Table 15.1.19 Resistance and 60-Hz Reactance for Wires with Small Spacings, Ω, at 20°C
(See also Table 15.1.16)

AWG and size of wire, cir mils	Resistance in 1,000 ft of line (2,000 ft of wire), copper	Reactance in 1,000 ft of line (2,000 ft of wire) at 60 Hz for the distance given in inches between centers of conductors										
		1/2	1	2	3	4	5	6	9	12	18	24
14– 4,107	5.06	0.138	0.178	0.218	0.220	0.233	0.244	0.252	0.271	0.284	0.302	
12– 6,530	3.18	0.127	0.159	0.190	0.210	0.223	0.233	0.241	0.260	0.273	0.292	
10– 10,380	2.00	0.116	0.148	0.180	0.199	0.212	0.223	0.231	0.249	0.262	0.281	
8– 16,510	1.26	0.106	0.138	0.169	0.188	0.201	0.212	0.220	0.238	0.252	0.270	0.284
6– 26,250	0.790	0.095	0.127	0.158	0.178	0.190	0.201	0.209	0.228	0.241	0.260	0.272
4– 41,740	0.498	0.085	0.117	0.149	0.167	0.180	0.190	0.199	0.217	0.230	0.249	0.262
2– 66,370	0.312	0.074	0.106	0.138	0.156	0.169	0.180	0.188	0.206	0.220	0.238	0.252
1– 83,690	0.248	0.068	0.101	0.132	0.151	0.164	0.174	0.183	0.201	0.214	0.233	0.246
0–105,500	0.196	0.063	0.095	0.127	0.145	0.159	0.169	0.177	0.196	0.209	0.228	0.241
00–133,100	0.156	0.057	0.090	0.121	0.140	0.153	0.164	0.172	0.190	0.204	0.222	0.236
000–167,800	0.122	0.052	0.085	0.116	0.135	0.148	0.158	0.167	0.185	0.199	0.217	0.230
0000–211,600	0.098	0.046	0.079	0.111	0.130	0.143	0.153	0.161	0.180	0.193	0.212	0.225
250,000	0.085	.....	0.075	0.106	0.125	0.139	0.148	0.157	0.175	0.189	0.207	0.220
300,000	0.075	.....	0.071	0.103	0.120	0.134	0.144	0.153	0.171	0.185	0.203	0.217
350,000	0.061	.....	0.067	0.099	0.188	0.128	0.141	0.149	0.168	0.182	0.200	0.213
400,000	0.052	.....	0.064	0.096	0.114	0.127	0.138	0.146	0.165	0.178	0.197	0.209
500,000	0.042	.....	.....	0.090	0.109	0.122	0.133	0.141	0.160	0.172	0.192	0.202
600,000	0.035	.....	.....	0.087	0.106	0.118	0.128	0.137	0.155	0.169	0.187	0.200
700,000	0.030	.....	.....	0.083	0.102	0.114	0.125	0.133	0.152	0.165	0.184	0.197
800,000	0.026	.....	.....	0.080	0.099	0.112	0.122	0.130	0.148	0.162	0.181	0.194
900,000	0.024	.....	.....	0.077	0.096	0.109	0.119	0.127	0.146	0.159	0.178	0.191
1,000,000	0.022	.....	.....	0.075	0.094	0.106	0.117	0.125	0.144	0.158	0.176	0.188

NOTE: For other frequencies the reactance will be in direct proportion to the frequency.

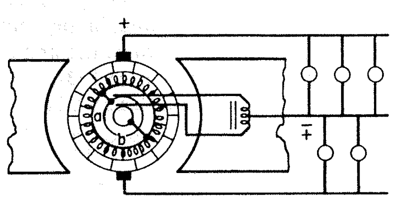


Fig. 15.1.90 Three-wire generator.

maintain constant voltage at the centers of distribution. This system provides a means of maintaining very close **voltage regulation** at the consumer's premises.

A common and economical method of supplying business and thickly settled districts with high load densities is to employ a 208/120-V, three-phase, four-wire **low-voltage ac network.** The network operates with 208 V between outer wires giving 120 V to neutral (Fig. 15.1.92). Motors are connected across the three outer wires operating at 208 V, three-phase. Lamp loads are connected between outer wires and the grounded neutral. The network is supplied directly from 13,800-V feeders by 13,800/208-V three-phase transformer

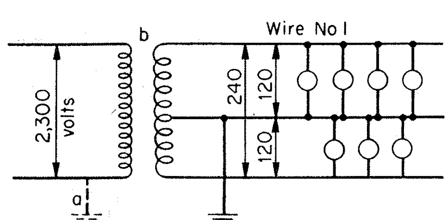


Fig. 15.1.91 Three-wire 230/115-V ac system.

units, usually located in manholes, vaults, or outdoor enclosures. This system thus eliminates the necessity for transformation in the substation. A large number of such units feed the network, so that the secondaries are all in parallel. Each transformer is provided with an overload reverse-energy circuit breaker (network protector), so that a feeder and its transformer are isolated if trouble develops in either. This system is flexible since units can be easily added or removed in accordance with the rapid changes in local loads that occur particularly in downtown business districts.

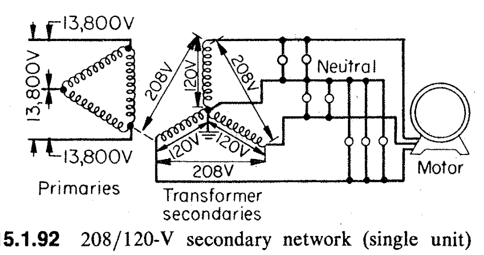


Fig. 15.1.92 208/120-V secondary network (single unit) showing voltages.

Voltage Drops In ac distribution systems the voltage drop from transformer to consumer in lighting mains should not exceed 2 percent in first-class systems, so that the lamps along the mains can all operate at nearly the same voltage and the annoying flicker of lamps may not occur with the switching of appliances. This may require a much larger conductor than the most economical size. In transmission lines and in feeders where there are no intermediate loads and where means of regulating the voltage are provided, the drop is not limited to the low values that are necessary with mains and the matter of economy may be given consideration.

WIRING CALCULATIONS

These calculations can be used for dc, and for ac if the reactance can be neglected. The determination of the proper size of conductor is influenced by a number of factors. Except for short distances, the minimum size of conductor shown in Table 15.1.21, which is based on the maximum permissible current for each type of insulation, cannot be used; the size of conductor must be larger so that the voltage drop IR shall not be too great. With branch circuits supplying an incandescent-lamp load, this drop should not be more than a small percentage of the voltage between wires. The National Electrical Code 1984 requires that conductors for feeders, i.e., from the service equipment to the final branch circuit overcurrent device, be sized to prevent (1) a voltage drop of more than 3 percent at the farthest outlet of power, heating, and lighting loads or combinations thereof, and (2) a maximum voltage drop on combined feeders and branch circuits to the farthest outlet of more than 5 percent.

The resistance of 1 cir mil·ft of commercial copper may be taken as 10.8 Ω. The resistance of a copper conductor may be expressed as $R = 10.8l/A$, where l = length, ft and A = area, cir mils. If the length is expressed in terms of the transmission distance d (since the two wires are usually run parallel), the voltage drop IR to the end of the circuit is

$$e = 21.6Id/A \qquad (15.1.128)$$

and the size of conductor in circular mils necessary to give the permissible voltage drop e is

$$A = 21.6Id/e \qquad (15.1.129)$$

If e is expressed as a percentage x of the voltage E between conductors, then

$$A = 2{,}160Id/xE \qquad (15.1.130)$$

EXAMPLE. Determine the size of conductor to supply power to a 10-hp, 220-V dc motor 500 ft from the switchboard with 5 V drop. Assume a motor efficiency of 86 percent. The motor will then require a current of $(10 \times 746)/(0.86 \times 220) = 39.4$ A. From Eq. (15.1.129), $A = 21.6 \times 39.4 \times 500/5 = 85{,}100$ cir mils. The next largest wire is no. 0 AWG.

The calculation of the size conductor for three-wire circuits is made in practically the same manner. With a balanced circuit there is no current in the neutral wire, and the current in each outside wire will be equal to one-half the sum of the currents taken by all the receiving devices connected between neutral and outside wires plus the sum of the currents taken by the receivers connected between the outside wires. Using this total current and neglecting the neutral wire, make calculations for the size of the outside wires by means of Eq. (15.1.129). The neutral wire should have the same cross section as the outside wires in interior wiring.

EXAMPLE. Determine the size wire which should be used for the three-wire main of Fig. 15.1.93. Allowable drop is 3 V and the distance to the load center 40 ft; circuit loaded with two groups of receivers each taking 60 A connected between the neutral and the outside wires, and one group of receivers taking 20 A connected across the outside wires.

Solution: load $= (60 + 60)/2 + 20 = 80$ A. Substituting in Eq. (15.1.129), cir mils $= 21.6Id/e = 21.6 \times 80 \times 40/3 = 23{,}030$ cir mils.

From Tables 15.1.19 and 15.1.21, no. 6 wire, which has a cross section of 26,250 cir mils, is the next size larger. This size of wire would satisfy the voltage-drop requirements, but rubber-insulated no. 6 has a safe carrying capacity of but 55 A. The current in the circuit is 80 A. Therefore, rubber-insulated wire no. 3, which has a carrying capacity of 80 A, should be used. The neutral wire should be the same size as the outside wires.

See also examples in the National Electrical Code 1984.

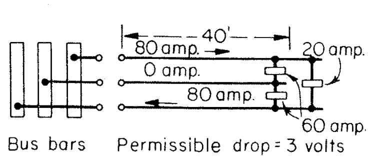


Fig. 15.1.93 Three-wire 230/115-V main.

Wiring calculations for ac circuits require some consideration of power factor, reactance, and skin effect. Skin effect becomes pronounced only when very large conductors are used for alternating current. For interior wiring, conductors larger than 700,000 cir mils should not be used, and many prefer not to use conductors larger than 300,000 cir mils. Should the required copper cross section exceed these values, a number of conductors may be operated in parallel.

For voltages under 5,000 the effect of line capacitance may be neglected. With ordinary single-phase interior wiring, where the effect of the line reactance may be neglected and where the power factor of the load (incandescent lamps) is nearly 100 percent, the calculations are made the same as for dc circuits. Three-wire ac circuits of ordinary length with incandescent lamp loads are also determined in the same manner. When the load is other than incandescent lamps, it is necessary to know the power factor of the load in order to make calculations. When the exact power factor cannot be accurately determined, the following approximate values may be used: incandescent lamps, 0.95 to 1.00; lamps and motors, 0.75 to 0.85; motors 0.5 to 0.80. Equation (15.1.131) gives the value of current in a single-phase circuit. See also Table 15.1.25.

$$I = (P \times 1{,}000)/(E \times \text{pf}) \qquad (15.1.131)$$

where I = current, A; P = kW; E = load voltage; and pf = power factor of the load. The size of conductor is then determined by substituting this value of I in Eq. (15.1.129) or (15.1.130).

For three-phase three-wire ac circuits the current per wire

$$I = 1{,}000P/\sqrt{3}E\text{pf} = 580P/E\text{pf} \qquad (15.1.132)$$

Computations are usually made of voltage drop **per wire** (see Fig. 15.1.94). Hence, if reactance can be neglected, the conductor cross section in cir mils is one-half that given by Eq. (15.1.129). That is,

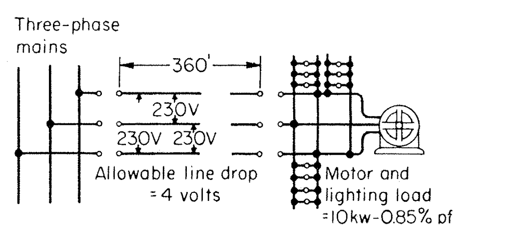


Fig. 15.1.94 Three-phase lamp and induction motor load.

$$A = 10.8Id/e \text{ cir mils} \qquad (15.1.133)$$

where e in Eq. (15.1.133) is the voltage drop per wire. The voltage drop between any two wires is $\sqrt{3}e$. The percent voltage drop should be in terms of the voltage to **neutral.** That is, percent drop $= [e/(E/\sqrt{3})]100 = [\sqrt{3}e/E]100$ (see Fig. 15.1.82).

EXAMPLE. In Fig. 15.1.94, load 10 kW; voltage of circuit 230; power factor 0.85; distance 360 ft; allowable drop per wire 4 V. Substituting in Eq. (15.1.132) $I = (580 \times 10)/(230 \times 0.85) = 29.7$ A. Substituting in Eq. (15.1.133), $A = 10.8 \times 29.7 \times 360/4 = 28{,}900$ cir mils.

The next larger commercially available standard-size wire (see Table 15.1.19) is 41,700 cir mils corresponding to AWG no. 4. From Table 15.1.21 this will carry 70 A with rubber insulation, and is therefore ample in section for 29.7 A. Three no. 4 wires would be used for this circuit.

From Table 15.1.19 the resistance of 1,000 ft of no. 4 copper wire is 0.249 Ω. Hence, the voltage drop per conductor, $e = 29.7 \times (360/1{,}000)0.249 = 2.66$ V. Percent voltage drop $= \sqrt{3} \times 2.66/230 = 2.00$ percent.

Where all the wires of a circuit, two wires for a single-phase circuit, four wires for a four-phase circuit (see Fig. 15.1.32 and Fig. 15.1.80*c*), and three wires for a three-phase circuit, are carried in the same conduit or where the wires are separated less than 1 in between centers, the effect of line (inductive) reactance may ordinarily be neglected. Where circuit conductors are large and widely separated from one another and the circuits are long, the inductive reactance may increase the voltage drop by a considerable amount over that due to resistance alone. Such problems are treated using *IR* and *IX* phasors. Line reactance decreases somewhat as the size of wire increases and decreases as the distance between wires decreases.

EXAMPLE. Determine the size of wire necessary for the branch to the 50-hp, 60-Hz, 250-V single-phase induction motor of Fig. 15.1.95.

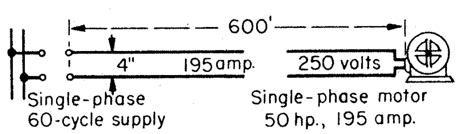


Fig. 15.1.95 Single-phase induction motor load on branch circuit.

The name-plate rating of the motor is 195 A, and its full-load power factor is 0.85. The wires are run open and separated 4 in; length of circuit, 600 ft. Assume the line drop must not exceed 7 percent, or 0.07 × 250 = 17.5 V. The point made by this example is emphasized by the assumption of an outsize motor.

Solution. To ascertain approximately the size of conductor, substitute in Eq. (15.1.129) giving cir mils $= 21.6 \times 195 \times 600/17.5 = 144{,}400$. Referring to Table 15.1.19, the next larger size wire is no. 000 or 167,800 cir mils. This size would be ample if there were no line reactance. In order to allow for reactance drop, a larger conductor is selected and the corresponding voltage drop determined. Inasmuch as this is a motor branch, the code rules require that the carrying capacity be sufficient for a 25 percent overload. Therefore the conductor should be capable of carrying 195 × 1.25 = 244 A. From Table 15.1.21, a 350,000-cir mil conductor rubber-insulated cable would be required to carry 244 A. Resistance drop (see Table 15.1.19), $IR = 195 \times 0.061 \times 0.6 = 7.14$ V. 7.14/250 = 2.86 percent. From Table 15.1.19, $X = 0.128 \times 0.6 = 0.0768$ Ω. $IX = 195 \times 0.768 = 14.98$ V. 14.98/250 = 5.99 percent. Using the Mershon diagram (Fig. 15.1.84), follow the ordinate corresponding to power factor, 0.85, until it intersects the smallest circle. From this point, lay off horizontally the percentage resistance drop, 2.86. From this last point, lay off vertically the percentage reactance drop 5.99. This last point lies about on the 6.0 percent circle, showing that with 195 amp the difference between the sending-end and receiving-end voltages is 0.06 × 250 = 15.0 V, which is within the specified limits.

Also Eq. (15.1.127), may be used. $\cos\theta = 0.85$; $\sin\theta = 0.527$.

$$E_s = \{[(250 \times 0.85) + 7.14]^2 + [(250 \times 0.527) + 14.98]^2\}^{1/2} = 264.3\text{V}$$

$$264.3/250 = 105.7 \text{ percent}$$

In the calculation of three-phase three-wire circuits where line reactance must be considered, the method found above under Power Transmission may be used. The system is considered as being three single-phase systems having a ground return the resistance and inductance of which are zero, and the voltages are equal to the line voltages divided by $\sqrt{3}$. When the three conductors are spaced unequally, the value of GMD given in Fig. 15.1.85 should be used in Tables 15.1.16 and 15.1.19. (When the value of resistance or reactance per 1,000 ft of **conductor** is desired, the values in Table 15.1.19 should be divided by 2.)

The National Electrical Code of 1984 specifies that the size of conductors for branch circuits should be such that the voltage drop will not exceed 3 percent to the farthest outlet for power, heating, lighting, or combination thereof, requiring further that the total voltage drop for feeders and branch circuits should not exceed 5 percent overall. For examples of calculations for interior wiring, see National Electrical Code of 1984 (Chap. 9).

INTERIOR WIRING

Interior wiring requirements are based, for the most part, on the **National Electrical Code** (NEC), which has been adopted by the

Table 15.1.20 Wire Table for Standard Annealed Copper at 20°C in SI Units

AWG size	Diameter, mm	Kgf/km	m/Ω	Area, mm²
14	1.628	18.50	120.7	2.08
12	2.053	29.42	191.9	3.31
10	2.588	46.77	305.1	5.261
8	3.264	74.37	485.2	8.367
6	4.115	118.2	771.5	13.30
4	5.189	188.0	1227	21.15
2	6.544	299.0	1951	33.62
1	7.348	377.0	2460	42.41
0	8.252	475.4	3102	53.49
00	9.266	599.5	3911	67.43
000	10.40	755.9	4932	85.01
0000	11.68	935.2	6219	107.2

National Fire Protection Association, American National Standards Institute (ANSI), and the Occupational Safety and Health Act (OSHA).

The Occupational Safety and Health Act of 1970 (OSHA) made the National Electrical Code a national standard. Conformance with the NEC became a requirement in most commercial, industrial, agricultural, etc., establishments in the United States. Some localities may not accept NEC standards. In those cases, local rules must be followed.

NEC authority starts at the point where the connections are made to the conductor of the **service drop** (overhead) or **lateral** (underground) from the electricity supply system. The **service equipment** must have a rating not less than the load to be carried (computed according to NEC methods). Service equipment is defined as the necessary equipment, such as circuit breakers or fused switches and accompanying accessories. This equipment must be located near the point of entrance of supply conductors to a building or other structure or an otherwise defined area. Service equipment is intended to be the main control and means of cutoff of the supply.

Service-entrance conductors connect the electricity supply to the service equipment. Service-entrance conductors running along the exterior or entering a building or other structure may be installed (1) as separate conductors, (2) in approved cables, (3) as cable bus, or (4) enclosed in rigid conduit. Also, for voltages less than 600 V, the conductors may be installed in electrical metallic tubing, wireways, auxiliary gutters, or busways. Service-entrance cables which are exposed to physical damage from awnings, swinging signs, coal chutes, etc., must be of the protected type or be protected by conduit, electrical metallic tubing, etc. Service heads must be raintight. Thermoplastic or rubber insulation is required in overhead services. A grounded conductor may be bare. If exposed to the weather or embedded in masonry, raceways must be raintight and arranged to drain. Underground service raceway or duct entering from an underground distribution system must be sealed with a suitable compound (spare ducts, also).

NEC rules permit **multiple services** to a building for various reasons, such as: (1) fire pumps, (2) emergency light and power, (3) multiple occupancy, (4) when the calculated load is greater than 3,000 A, (5) when the building extends over a large area, and (6) where different voltages, frequencies, number of phases, or classes of use are required.

Ordinary service drops (overhead) and lateral (underground) must be large enough to carry the load but not smaller than no. 8 copper or no. 6 aluminum. As an exception, for installations to supply only limited loads of a single branch circuit, such as small polyphase power, etc., service drops must not be smaller than no. 12 hard-drawn copper or equivalent, and service laterals must be not smaller than no. 12 copper or no. 10 aluminum.

The phrase **large enough to carry the load** requires elaboration. The various conductors of public-utility electric-supply systems are sized according to the calculations and decisions of the personnel of the specific public utility supplying the service drop or lateral. At the load end of the drop or lateral, the NEC rules apply, and from that point on into the consumer's premises, NEC rules are the governing authority. There is a discontinuity at this point in the calculation of combined load demand for electricity and allowable current (ampacity) of conductors, cables, etc. This discontinuity in calculations results from the fact that the utility company operates locally, whereas the NEC is a set of national standards and therefore cannot readily allow for regional differences in electrical coincident demand, ambient temperature, etc. The NEC's aim is the assurance of an electrically "safe" human environment. This will be fostered by following the NEC rules.

Service-entrance cables are conductor assemblies which bear the type codes **SE** (for overhead services) and **USE** (for underground services). Under specified conditions, these cables may also be used for interior feeder and branch-circuit wiring.

The service-entrance equipment must have the **capability of safely interrupting** the current resulting from a short circuit at its terminals. **Available short-circuit current** is the term given to the maximum current that the power system can deliver through a given circuit to any negligible-impedance short circuit applied at a given point. (This value can be in terms of symmetrical or asymmetrical, momentary or clearing current, as specified.)

In most instances, the available short-circuit current is limited by the impedance of the last transformer in the supply system. **Large power users,** however, must become aware of changes in the electricity supply system which, because of growth of system capacity or any other reason, would increase the short-circuit current available to their service-entrance equipment. If this current is too great, explosive failure can result.

Kilowatthour and sometimes demand-metering equipment are connected to the service-entrance conductors. Proceeding toward the utilization equipment, the power-supply system fans out into feeders and branch circuits (see Fig. 15.1.96).

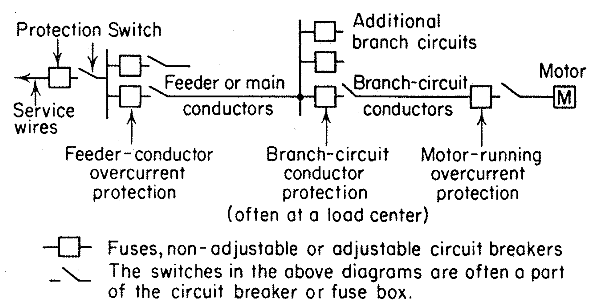


Fig. 15.1.96 Motor and wiring protection.

Each of the feeders, i.e., a run of untapped conductor or cable, is connected to the supply through a switch and fuses or a circuit breaker. At a point, usually near that portion of the electrical loads which are to be supplied, a **panel box** or perhaps a **load-center assembly** of switching and/or control equipment is installed. From this panel box or load-center assembly, circuits radiate; i.e., circuits are installed to extend into the area being served to connect electrical machinery or devices or make available electric receptacles connected to the source of electric power.

Each feeder and each branch circuit will have its own **overcurrent protection** and **disconnect** means in the form of a fuse and switch combination or a circuit breaker.

There is a provision in the NEC 1984 rules for the following types of feeder and branch circuit wiring:

1. Open Wiring on Insulators (NEC 1984, Art. 320). This wiring method uses approved cleats, knobs, tubes, and flexible tubing for the protection and support of insulated conductors run on or in buildings and not concealed by the building structure. It is permitted only in industrial or agricultural establishments.

2. Concealed Knob-and-Tube Work (NEC 1984, Art. 324). Concealed knob-and-tube work may be used in the hollow spaces of walls and ceilings. It may be used only for the extension of existing facilities.

3. Flat Conductor Cable, Type FCC (NEC 1984, Art 328). Type FCC cable may be installed under carpet squares. It may not be used outdoors or in wet locations, in corrosive or hazardous areas, or in residential, school, or hospital buildings.

4. Mineral-Insulated Metal-Sheathed Cable, Type MI (NEC 1984, Art. 330). Type MI Cable contains one or more electrical conductors insulated with a highly compressed refractory mineral insulation and enclosed in a liquid- and gastight metallic tube sheath. Appropriate approved fittings must be used with it. It may be used for services, feeders, and branch circuits either exposed or concealed and dry or wet. It may be used in Class I, II, or III hazardous locations. It may be used for under-plaster extensions and embedded in plaster finish or brick or other masonry. It may be used where exposed to weather or continuous moisture, for underground runs and embedded in masonry, concrete or fill, in buildings in the course of construction or where exposed to oil, gasoline, or other conditions. If the environment would cause destruction of the sheath, it must be protected by suitable materials.

5. Power and Control Tray Cable (NEC 1984, Art 340). Type TC cable is a factory assembly of two or more insulated conductors with or without associated bare or covered-grounding conductors under a nonmetallic sheath, approved for installation in cable trays, in raceways, or where supported by messenger wire.

6. Metal-Clad Cable, Type MC and AC Series (NEC 1984, Art. 333 and 334). These are metal-clad cables, i.e., an assembly of insulated conductors in a flexible metal enclosure. Type MC are power cables and in the range up to 600 V are made in conductor sizes of no. 4 and larger for copper and no. 2 and larger for aluminum. Type AC are branch and feeder cables with armor of flexible metal tape. All AC types except ACL have an internal bonding strip of copper or aluminum in intimate contact with the armor for its entire length. Metal-clad branch circuit cable was formerly called BX. Metal-clad cables may generally be installed where not subject to physical damage, for feeders and branch circuits in exposed or concealed work, with qualifications for wet locations, direct burial in concrete, etc. The use of Type AC cable is prohibited (1) in motion-picture studios, (2) in theaters and assembly halls, (3) in hazardous locations, (4) where exposed to corrosive fumes or vapors, (5) on cranes or hoists except where flexible connections to motors, etc., are required, (6) in storage-battery rooms, (7) in hoistways or on elevators except (*i*) between risers and limit switches, interlocks, operating buttons, and similar devices in hoistways and in escalators and moving walkways and (*ii*) short runs on elevator cars, where free from oil, and if securely fastened in place, or (8) in commercial garages in hazardous locations. Type ACL (lead-covered) shall not be used for direct burial in the earth.

7. Nonmetallic-Sheathed Cable, Types NM and NMC (NEC 1984, Art. 336). These are assemblies of two more insulated conductors (nos. 14 through 2 for copper, nos. 12 through 2 for aluminum) having an outer sheath of moisture-resistant, flame-retardant, nonmetallic material. In addition to the insulated conductors, the cable may have an approved size of uninsulated or bare conductor for grounding purposes only. The outer covering of NMC cable is flame retardant and corrosion resistant. The use of this type of cable, commonly called **Nomex,** is permitted in one- or two-family dwellings, multifamily dwellings, and other structures provided that such structures do not exceed three floors above grade.

8. Shielded Nonmetallic-Sheathed Cable, Type SNM, (NEC 1984, Art. 337). Type SNM is a factory-assembled cable consisting of two or more insulated conductors (nos. 14 through 2 copper and nos. 12 through 2 aluminum) in an extruded core of moisture-resistant material, covered with an overlapping spiral metal tape and wire shield and jacketed with an extruded moisture-, flame-, corrosion-, fungus-, and sunlight-resistant material. This cable is to be used (1) under appropriate ambient-temperature conditions and (2) in continuous rigid-cable support or in raceways. It can be used in some hazardous locations as defined by the NEC.

9. Service Entrance Cable, Types SE and USE (NEC 1984, Art. 338). These cables, containing one or more individually insulated conductors, are primarily used for electric services. Type SE has a flame-retardant, moisture-resistant covering and is not required to have built-in protection against mechanical abuse. Type USE is recognized for use underground. It has a moisture-resistant covering, but not necessarily a flame-retardant one. Like the SE cable, USE cable is not required to have inherent protection against mechanical abuse. Under specified conditions, SE and USE cables can be used for feeders and branch circuits.

10. Underground Feeder and Branch-Circuit Cable, Type UF (NEC 1984, Art. 339). This cable is made in sizes 14 through 4/0, and the insulated conductors are Types TW, RHW, and others approved for the purpose. As in the NM cable, the UF type may contain an approved size of uninsulated or bare conductor for grounding purposes only. The outer jacket of this cable shall be flame-retardant, moisture-resistant, fungus-resistant, corrosive-resistant, and suitable for direct burial in the ground.

11. Other Installation Practices. The NEC details rules for nonmetallic circuit extensions and underplaster extensions. It also provides detailed rules for installation of electrical wiring in (*a*) rigid metal conduit (which may be used for all atmospheric conditions and locations with due regard to corrosion protection and choice of fittings), (*b*) rigid nonmetallic conduit (which is essentially corrosion-proof), in electrical metallic tubing (which is lighter-weight than rigid metal conduit), (*c*) flexible metal conduit, (*d*) liquidtight flexible metal conduit, (*e*) surface raceways, (*f*) underfloor raceways, (*g*) multioutlet assembles, (*h*) cellular metal floor raceways, (*i*) structural raceways, (*j*) cellular concrete floor raceways, (*k*) wireways (sheet-metal troughs with hinged or removable covers), (*l*) flat, Type FC, cable assemblies installed in a surface metal raceway (Type FC cable contains three or four no. 10 special stranded copper wires), (*m*) busways, and (*n*) cable-bus. Busways and cable-bus installations are permitted for exposed work only.

In all installation work, only approved outlets, switch and junction boxes, fittings, terminal strips, and dead-end caps shall be used, and they are to be used in an approved fashion (see NEC 1984, Art. 370).

Table 15.1.21 lists the allowable ampacities of copper and aluminum conductors. Table 15.1.22 lists various conductor insulation systems approved by the 1984 NEC for conductors used in interior wiring. Table 15.1.20 relates AWG wire sizes to metric units. Dimensions and allowable fill of conduit and tubing are listed in Table 15.1.23. Table 15.1.24 lists the cross-

Table 15.1.21 Allowable Ampacities of Insulated Conductors
(Not more than three conductors in conduit. Based on ambient air temperature of 40°C.)

	Temperature rating of insulation, °C					
	Copper			Aluminum		
	60	75	90	60	75	90
Conductor size: AWG or MCM	Types RUW, T, TW, UF	Types RH, RHW, RUH, THW, THWN, XHHW, USE, ZW	Types SA, AVB, FEP, FEPB, THHN, RHH, XHHW*	Types RUW, T, TW, UF	Types RH, RHW, RUH, THW, THWN, XHHW, USE	Types SA, AVB, THHN, RHH, XHHW*
14	18†	22†	25†	—	—	—
12	23†	28†	32†	18†	22†	26†
10	29†	37†	42†	23†	29†	34†
8	36	48	55	28	37	43
6	50	64	75	37	50	58
4	65	83	97	50	65	76
3	76	98	114	59	76	89
2	87	112	130	68	87	102
1	104	134	156	81	104	122
0	119	153	179	93	119	139
00	135	175	204	106	137	159
000	160	207	242	125	162	189
0000	184	238	278	144	186	217
250	210	271	317	165	213	249
300	232	300	351	183	236	276
350	254	328	384	201	259	303
400	274	354	415	218	281	329
500	314	407	477	252	326	381
600	345	448	525	280	362	424
700	376	489	574	308	399	467
750	392	509	598	322	417	488
800	403	524	616	334	432	506
900	426	555	653	357	463	542
1,000	499	585	689	380	493	578
Ambient temp., °C	For ambient temperatures other than 40°C, multiply the ampacities shown above by the appropriate factors shown below.					
21–35	1.32	1.20	1.14	1.32	1.20	1.14
26–30	1.22	1.13	1.10	1.22	1.13	1.10
31–35	1.12	1.07	1.05	1.12	1.07	1.05
36–40	1.00	1.00	1.00	1.00	1.00	1.00
41–45	0.87	0.93	0.95	0.87	0.93	0.95
46–50	0.71	0.85	0.89	0.71	0.85	0.89
51–55	0.50	0.76	0.84	0.50	0.76	0.84
56–60	—	0.65	0.77	—	0.65	0.77
61–70	—	0.38	0.63	—	0.38	0.63
71–80	—	—	0.45	—	—	0.45

SOURCE: NEC® 1984, Table 310-23. Reprinted with permission from NFPA 70-1984, National Electrical Code®, Copyright© 1983, National Fire Protection Association, Quincy, Massachusetts 02269. This reprinted material is not the complete and official position of the NFPA on the referenced subject, which is represented only by the standard in its entirety.
*For dry locations only, rated 75°C for wet locations.
†Overcurrent protection shall not exceed 15 A for no. 14 copper and no. 12 aluminum, 20 A for no. 12 copper, 25 A for no. 10 aluminum, and 30 A for no. 10 copper.
NOTE: This is a general table. For other installation conditions, see NFPA 70 - 1984 National Electric Code®, Article 310. This table is effective January 1, 1987.

sectional area of various insulated conductors. For installations not covered by the tables presented here, review the 1984 NEC. Estimated full-load currents of motors can be taken from Table 15.1.25.

Switching Arrangements Small quick-break switches must be set in or on a metal box or fitting and may be of the push, tumbler, or rotary type. The following types of switches are used to control lighting circuits: (1) single-pole, (2) double-pole, (3) three-point or three-way, (4) four-way, in combination with three-way switches to control lights from three or more stations, (5) electrolier.

In all metallic protecting systems, such as conduit, armored cable, or metal raceways, joints and splices in conductors must be made only in junction boxes or other proper fittings; therefore, these fittings can be located only in accessible places and never concealed in partitions. Splices or joints in the wire

Table 15.1.22 Conductor Type and Application

Type letter	Insulation, trade name (see NEC 1984 Art. 310, for complete information)	Environment	Outer covering[a]
	Max operating temperature = 60°C (140°F)		
RUW	Moisture-resistant latex rubber	Dry and wet	1
T	Thermoplastic	Dry	None
TW	Moisture-resistant thermoplastic	Dry and wet	None
TF	Thermoplastic-covered, solid or 7-strand	c,d	None
TFF	Thermoplastic-covered, flexible stranding	c,d	None
MTW	Moisture-, heat-, and oil-resistant thermoplastic machine-tool wiring (NFPA Stand. 79, NEC 1975, Art. 670)	Wet	None or nylon
UF	Moisture-resistant, underground feeder	Dry and wet	None
	Max operating temperature = 75°C (167°F)		
RH	Heat-resistant rubber	Dry	1,2
RHW	Moisture- and heat-resistant rubber[e]	Dry and wet	1,2
RUH	Heat-resistant latex rubber	Dry	1
THW	Moisture- and heat-resistant thermoplastic	Dry and wet	None
THWN	Moisture- and heat-resistant thermoplastic	Dry and wet	Nylon
XHHW	Moisture- and heat-resistant cross-linked polymer	Wet	None
RFH-1 & 2	Heat-resistant rubber-covered solid or 7-strand	b–d	None
FFH-1 & 2	Heat-resistant rubber-covered flexible stranding	b–d	None
UF	Moisture-resistant and heat-resistant	Dry and wet	None
USE	Heat- and moisture-resistant	Dry and wet	4
ZW	Modified ethylene tetrafluorethylene	Wet	None
	Max operating temperature = 85°C (185°F)		
MI	Mineral-insulated (metal-sheathed)	Dry and wet	Copper
	Max operating temperature = 90°C (194°F)		
RHH	Heat-resistant rubber	Dry	1,2
THHN	Heat-resistant thermoplastic	Dry	Nylon
THW	Moisture- and heat-resistant thermoplastic	f	None
XHHW	Moisture- and heat-resistant cross-linked synthetic polymer	Dry	None
FEP	Fluorinated ethylene propylene	Dry	None
FEPB	Fluorinated ethylene propylene	Dry	3
TFN	Heat-resistant thermoplastic covered, solid or 7-strand	c,d	Nylon
TFFN	Heat-resistant thermplastic flexible stranding		Nylon
MTW	Moisture-, heat-, and oil-resistant thermoplastic machine-tool wiring (NFPA Stand. 79, NEC 1975, Art. 670)	Dry	None or nylon
SA	Silicone asbestos	Dry	Asbestos or Glass
	Max operating temperature = 150°C (302°F)		
Z, ZW	Modified ethylene tetrafluorethylene	Dry	None
	Max operating temperature = 200°C (392°F)		
FEP, FEPB	Fluorinated ethylene propylene	Dry	None
	Special applications		3
PF, PGF	Fluorinated ethylene propylene	c,d	None or glass braid
PFA	Perfluoroalkoxy	Dry	None
SF-2	Silicone rubber, solid or 7-strand	c,d	Nonmetallic

Table 15.1.22 Conductor Type and Application (*Continued*)

Type letter	Insulation, trade name (see NEC 1984 Art. 310, for complete information)	Environment	Outer covering[a]
	Max operating temperature = 250°C (482°F)		
MI	Mineral-insulated (metal-sheathed), for special applications	Dry and wet	Copper
TFE	Extruded polytetrafluoroethylene, only for leads within apparatus or within raceways connected to apparatus, or as open wiring (silver or nickel-coated copper only)	Dry	None
PFAH	Perfluoroalkoxy (special application)	Dry	None
PTF	Extruded polytetrafluorethylene, solid or 7-strand (silver or nickel-coated copper only)	[c,d]	None

SOURCE: NEC 1984 Tables 310–13 and 402-3. Reprinted with permission from NFPA 70-1984, National Electrical Code®, Copyright© 1983. National Fire Protection Association, Quincy, Massachusetts 02269. This reprinted material is not the complete and official position of the NFPA on the referenced subject, which is represented only by the standard in its entirety.

[a]1: Moisture-resistant, flame-retardent nonmetallic; 2: outer covering not required when rubber insulation has been specifically approved for the purpose; 3: no. 14-8 glass braid, no. 6-2 asbestos braid; 4: moisture-resistant nonmetallic.

[b]Limited to 300 V.

[c]No. 18 and no. 16 conductor for remote controls, low-energy power, low-voltage power, and signal circuits; NEC 1975 Sec. 725-16, Sec. 760-16.

[d]Fixutre wire no. 18-16.

[e]For over 2,000 V, the insulation shall be ozone-resistant.

[f]Special applications within electric discharge lighting equipment. Limited to 1,000 V open-circuit volts or less.

must never be in the conduit piping, raceway, or metallic tubing itself, for the splices may become a source of trouble as a result of corrosion or grounding if water should enter the conduit.

All conductors of an ac system must be placed in the same metallic casing so that their resultant magnetic field is nearly zero. If this is not done, eddy currents are set up causing heating and excessive loss. With single conductors in a casing, an excessive reactance drop may result.

Service wires are the conductors that bring the electric power into a building and should enter the building as near as possible to the service switch, so that when the switch is open all the electrical conductors and equipment inside the building will be dead. The service wires must be rubber- or thermoplastic-covered from the point of support on the outside of the building to the service switch or cutout and must be no. 6 wire or larger except for installations consisting of two-wire branch circuits where no. 8 wire may be permitted. A minimum of 100-A three-wire service is recommended for all single-family residences.

Generally, when the conductors from overhead lines enter a building, the wires are encased in rigid conduit equipped with a weather cap or a service entrance cable (type ASE armored or SE type, unarmored) may be attached directly to the building wall. The inner end of the service enters a metal service cabinet in which the service fuses and switch are located. Ser-

Table 15.1.23 Dimensions and Allowable Fill of Conduit and Tubing

			Allowable fill, in² of conductors (not lead-covered)		
Trade size, in	ID, in	Area, in²	One conductor, 53% fill	Two conductors, 31% fill	Over two conductors,* 40% fill
½	0.622	0.30	0.16	0.09	0.12
¾	0.824	0.53	0.28	0.16	0.21
1	1.049	0.86	0.46	0.27	0.34
1¼	1.380	1.50	0.80	0.47	0.60
1½	1.610	2.04	1.08	0.63	0.82
2	2.067	3.36	1.78	1.04	1.34
2½	2.469	4.79	2.54	1.43	1.92
3	3.068	7.38	3.91	2.29	2.95
3½	3.548	9.90	5.25	3.07	3.96
4	4.026	12.72	6.74	3.94	5.09
4½	4.506	15.94	8.45	4.94	6.38
5	5.047	20.00	10.60	6.20	8.00
6	6.065	28.89	15.31	8.96	11.56

SOURCE: NEC 1984, p. 70-684. Reprinted with permission from NFPA 70-1984, National Electrical Code®, Copyright © 1983, National Fire Protection Association, Quincy, Massachusetts 02269. This reprinted material is not the complete and official position of the NFPA on the referenced subject, which is represented only by the standard in its entirety.

*For conductor derating with more than three conductors see NEC 1984, Art. 310.

Table 15.1.24 Nominal Cross-Sectional Area of Rubber-Covered and Plastic-Covered Conductors

	Size, AWG					
	18	16	14	12	10	8
Type	Cross-sectional area, in^2					
RFH-2	0.0167	0.0196	(fixture wire)			
SF-2	0.0167	0.0196	0.0230	(fixture wire)		
RH	...	...	0.0230	0.0278	0.0460	0.0854
RHH, RHW	...	...	0.0327	0.0384	0.0460	0.0854
RHH, RHW (without outer covering)	...	...	0.0206	0.0251	0.0311	0.0598
THW	...	...	0.0206	0.0251	0.0311	0.0598
T, TW, RUH, RUW	...	...	0.0135	0.0172	0.0224	0.0471
TF	0.0088	0.0109	...	...	...	...
THWN, THHN	...	...	0.0087	0.0117	0.0184	0.0373
TFN	0.0064	0.0079	...	...	...	...
FEPB, Z, ZF, ZFF	...	...	0.0087	0.0115	0.0159	0.0272
FEP, TFE	...	...	0.0087	0.0115	0.0159	0.0333
PTF	0.0052	0.0066	0.0087	...	...	...
XHHW	...	...	0.0131	0.0167	0.0216	0.0456

	Size, AWG								
	6	4	3	2	1	1/0	2/0	3/0	4/0
	Cross-sectional area, in^2								
RH, RHH,* RHW*	0.1238	0.1605	0.1817	0.2067	0.2715	0.3107	0.3578	0.4151	0.4840
RUH, RUW	0.0819	0.1087	0.1263	0.1473	...	...	...	...	...
TW, T, THW	0.0819	0.1087	0.1263	0.1473	0.2027	0.2367	0.2781	0.3288	0.3904
TFE	0.0467	0.0669	0.0803	0.0973	0.1385	0.1676	0.1974	0.2436	0.2999
FEPB, ZF, ZFF	0.0716	0.0962	0.1122	0.1316	...	...	...	...	...
FEP	0.0467	0.0669	0.0803	0.0973	...	...	...	...	...
THHN, THWN	0.0519	0.0845	0.0995	0.1182	0.1590	0.1893	0.2265	0.2715	0.3278
XHHW	0.0625	0.0845	0.0995	0.1182	0.1590	0.1893	0.2265	0.2715	0.3278
Z	0.0716	0.0962	0.1122	0.1320	0.1385	0.1676	0.1948	0.2463	0.3000

	Size, MCM					
	250	500	750	1000	1500	2000
	Cross-sectional area, in^2					
RH, RHH,* RHW*	0.5917	0.9834	1.4082	1.7531	2.5475	3.2079
TW, T, THW	0.4877	0.8316	1.2252	1.5482	2.2748	2.9013
THHN, THWN	0.4026	0.7163	1.0623	1.3623	...	...
XHHW	0.4026	0.7163	1.0936	1.3893	2.0612	2.6590

SOURCE: NEC 1984, p. 70-686. Reprinted with permission from NFPA 70-1984, National Electrical Code®, Copyright© 1983, National Fire Protection Association, Quincy, Massachusetts 02269. This reprinted material is not the complete and official position of the NFPA on the referenced subject, which is represented only by the standard in its entirety.

NOTE: For general branch and feeder circuits the minimum conductor size is 14. Sizes 14 to 8 are siold wire. Sizes 6 and larger are stranded.

*RHH and RHW without covering have the same dimension as THW.

vice conductors may also terminate at an air-break or oil-immersed switch in a metal case or on a panel board which is accessible to qualified persons only.

All underground service wires must be connected to the interior wiring through a blade of the service switch or circuit breaker and be fused or automatically interrupted at the service switch. A service switch controlling a three-wire dc, a single-phase or four-wire three-phase system having a grounded neutral wire does not need to open that conductor.

The single-line diagram, Fig. 15.1.96, indicates a simplified interior arrangement of circuits and the necessary protection of the conductors and terminal load. Where a reduction is made in the wire size a protective device shall be installed to limit the conductor current to a safe value. Large motors and other terminal loads should also have overcurrent protection.

The maximum permissible fuse ratings and the setting of the protective devices for starting and for running protection of motors are given in Table 15.1.26.

Grounding of direct- and alternating-current systems of 300 V and less is usually required. Inside a building the **grounded conductor** (one of the two conductors in two-wire system and the **neutral conductor** in a three-wire or a four-wire system) should have a **white or natural gray covering** throughout to distinguish it from the ungrounded conductors. This identified

Table 15.1.25 Approximate Full-Load Currents of Motors,* A
(See NEC 1984 for more complete information.)

hp	Three-phase ac motors, squirrel-cage and wound-rotor induction types 230 V	460 V	575 V	2,300 V	Synchronous type, unity power factor 230 V	460 V	575 V	2,300 V	Single-phase ac motors 115 V‡	230 V‡	DC motors† 120 V	240 V
½	2	1	0.8	...	...	...	...	...	9.8	4.9	5.4	2.7
¾	2.8	1.4	1.1	...	...	...	...	...	13.8	6.9	7.6	3.8
1	3.6	1.8	1.4	...	...	...	...	...	16	8	9.5	4.7
1½	5.2	2.6	2.1	...	...	...	...	...	20	10	13.2	6.6
2	6.8	3.4	2.7	...	...	...	...	...	24	12	17	8.5
3	9.6	4.8	3.9	...	...	...	...	...	34	17	25	12.2
5	15.2	7.6	6.1	...	...	...	...	...	56	28	40	20
7½	22	11	9	...	...	...	...	...	80	40	58	29
10	28	14	11	...	...	...	...	...	100	50	76	38
15	42	21	17	...	...	...	...	...	...	...	...	55
20	54	27	22	...	...	...	...	...	...	...	...	72
25	68	34	27	...	53	26	21	...	...	...	...	89
30	80	40	32	...	63	32	26	...	...	...	...	106
40	104	52	41	...	83	41	33	...	...	...	...	140
50	130	65	52	...	104	52	42	...	...	...	...	173
60	154	77	62	16	123	61	49	12	...	...	...	206
75	192	96	77	20	155	78	62	15	...	...	...	255
100	248	124	99	26	202	101	81	20	...	...	...	341
125	312	156	125	31	253	126	101	25	...	...	...	425
150	360	180	144	37	302	151	121	30	...	...	...	506
200	480	240	192	49	400	201	161	40	...	...	...	675

*The values of current are for motors running at speeds customary for belted motors and motors having normal torque characteristics. Use nameplate data for low-speed, high-torque, or multispeed motors. For synchronous motors of 0.8 pf multiply the above amperes by 1.25, at 0.9 pf by 1.1. The motor voltages listed are rated voltages. Respective nominal system voltages would be 220 to 240, 440 to 480, and 550 to 600 V. For full-load currents of 208-V motors, multiply the above amperes by 1.10; for 200-V motors, multiply by 1.15.

†Ampere values are for motors running at base speed.

‡Rated voltage. Nominal system voltages are 120 and 240.

grounded conductor must not be fused or be opened unless the other conductors are opened simultaneously. **Green wires** only shall be used for grounding electrical equipment, such as motors, as well as conduits, armor, boxes, and such metallic enclosures. Four-wire circuits have **black, white, red, and blue** conductors or alphanumeric identification.

DC systems need be grounded only at the generating stations because the grounded wire is electrically connected to one of the conductors in all the circuits throughout the system. In ac systems, since one section can be insulated from the other by a transformer, each section of 300 V or under is grounded at the individual services. The conductor grounding the ac system should not be less than no. 8 copper wire and must be without a joint or a splice and run from the supply side of the service switch.

The service conduit that protects the service wires on the outside of the building must be grounded by a wire at least as large as no. 8, run directly to ground.

The entire metallic system surrounding the conductors must be at ground potential. It is only necessary to ground the metallic system, including motors and other equipment, at one point, provided that each section makes a good electrical connection with the next.

Since January 1, 1973, **ground-fault circuit interrupters** have been required by the NEC in some areas for personnel protection from line-to-ground electrical shock. Such circuit interrupters are required by the 1984 NEC in branch circuits supplying certain areas in residences, hotels and motels, health-care facilities, marinas and boat yards, mobile homes and recreational vehicles, swimming pools, and construction sites. The 1984 or subsequent code should be carefully reviewed for exact requirements.

Ground-fault circuit protection may be used at other locations and, if so used, will provide additional protection against line-to-ground shock hazard.

Ground fault-circuit interrupters monitor the current in the two conductors of a circuit. These two currents should be equal. If they are not equal, some current is leaking to ground, indicating a line to ground fault. If the difference between the two currents is 5 mA or more, the ground-fault circuit interrupter will automatically disconnect the faulted circuit in about 0.025 s.

Overload Protection A fuse or circuit breaker must be provided in all ungrounded conductors. Induction motors are usually protected by overload or thermal relays operating as automatic circuit breakers. To protect wiring properly, an automatic cutout must be installed at every point where a change is made in the size of the wire. Fuses or circuit breakers should not be placed either in a grounded line or in a ground wire. (See Fig. 15.1.96.)

Fuses, cutout bases, and **switches** are manufactured and change sizes, as follows: Edison plug (125 V only), 0 to 30 A; spring-clip cartridge (ferrule contact), 0 to 30, 31 to 60 A; knife-blade cartridge type, 61 to 100, 101 to 200, 201 to 400, 401 to 600

Table 15.1.26 Maximum Rating or Setting of Motor Branch-Circuit Protective Devices and Starting-Inrush Code Letters

Type of motor	Percent of full-load current: Non-time delay fuse	Percent of full-load current: Dual-element (time delay) fuse	Percent of full-load current: Instantaneous breaker	Percent of full-load current: Time-limit breaker
Single-phase, all types, no code letter	300	175	700	250
AC motors: single-phase, polyphase squirrel-cage, or synchronous with full-voltage, resistor, or reactor starting:				
No code letter	300	175	700	250
F–V	300	175	700	250
B–E	250	175	700	200
A	150	150	700	150
AC squirrel-cage or synchronous motors with autotransformer starting:				
Not more than 30 A				
No code letter	250	175	700	200
More than 30 A				
No code letter	200	175	700	200
F–V	250	175	700	200
B–E	200	175	700	200
A	150	150	700	150
High reactance squirrel-cage:				
Not more than 30 A				
No code letter	250	175	700	250
More than 30 A				
No code letter	200	175	700	200
Wound rotor, no code letter	150	150	700	150
DC motors:				
No more than 50 hp				
No code letter	150	150	250	150
More than 50 hp				
No code letter	150	150	175	150
Low-torque, low-speed (450 r/min or lower) synchronous motors which start unloaded	200	200	200	200

Code letter	kVA/hp with locked rotor
A	0 – 3.14
B	3.15 – 3.54
C	3.55 – 3.99
D	4.0 – 4.49
E	4.5 – 4.99
F	5.0 – 5.59
G	5.6 – 6.29
H	6.3 – 7.09
J	7.1 – 7.99
K	8.0 – 8.99
L	9.0 – 9.99
M	10.0 – 11.19
N	11.2 – 12.49
P	12.5 – 13.99
R	14.0 – 15.99
S	16.0 – 17.99
T	18.0 – 19.99
U	20.0 – 22.39
V	22.4 and up

SOURCE: NEC 1984, Tables 430-152 and 430-7(*b*). Reprinted with permission from NFPA 70-1984, National Electrical Code®, Copyright© 1983, National Fire Protection Association, Quincy, Massachusetts 02269. This reprinted material is not the complete and official position of the NFPA on the referenced subject, which is represented only by the standard in its entirety.

A; for 601 A and larger, knife-blade cartridge type with equal-size fuses in parallel may be used except for the protection of a branch motor circuit where a circuit breaker can be installed.

Since the rating of a fuse is only about 90 percent of the current that it will carry indefinitely, and since it may also take a few minutes before the heat due to slightly excessive current would be sufficient to melt the fuse wire and hence open the circuit, insulation may be permanently damaged if the fuses are larger than the current-carrying capacities given by Table 15.1.21.

Demand Calculations for Building Feeder Sizes The **demand factor** or **demand** is the ratio of maximum demand to the total connected load. This depends on the type of building, whether hotel, theater, factory, etc. The demand factor for any particular class of installation decreases as the floor area increases. Values of demand factors are found in the 1984 NEC Art. 220.

Load centers are panels or cabinets which act as distribution centers and which are supplied by feeder or main conductors, and from which the current to several branch circuits is taken. In each branch circuit there is usually a small combined switch and circuit breaker. Load centers for widely different types of circuits such as single-phase two-wire, single-phase three-wire, and three-phase are readily obtainable from several manufacturers.

RESISTOR MATERIALS

For use in rheostats, electric furnaces, ovens, heaters, and many electrical appliances, a resistor material with high melting point and high resistivity which does not disintegrate or corrode at high temperatures is necessary. These requirements are met by the nickel-chromium and nickel-chromium-iron alloys. For electrical instruments and measuring apparatus, the resistor material should have high resistivity, low temperature coefficient, and, for many uses, low thermoelectric power against copper. The properties of resistor materials are given in Table 15.1.27. Most of these materials are available in ribbon as well as in wire form. Cast-iron and steel wire are efficient and economical resistor materials for many uses, such as power-absorbing rheostats and motor starters and controllers. (See also Sec. 6.)

Advance has a low temperature coefficient and is useful in many types of measuring instrument and precision equipment. Because of its high thermoelectric power to copper, it is valuable for thermoelements and pyrometers. It is noncorrosive and is used to a large extent in industrial and radio rheostats. **Hytemco** is a nickel-iron alloy characterized by a high temperature coefficient and is used advantageously where self-regulation is required as in immersion heaters and heater pads. **Magno** is a manganese-nickel alloy used in the manufacture of incandescent lamps and radio tubes. **Manganin** is a copper-manganese-nickel alloy which, because of its very low temperature coefficient and its low thermal emf with respect to copper, is very valuable for high-precision electrical measuring apparatus. It is used for the resistance units in bridges, for shunts, multipliers, and similar measuring devices. **Nichrome V** is a nickel-chromium alloy free from iron, is noncorrosive, nonmagnetic, withstands high temperatures, and has high resistivity. It is recommended as material for heating elements in electric furnaces, hot-water heaters, ranges, radiant heaters, and high-grade electrical appliances. **Kanthal** is used for heating applications where higher operating temperatures are required than for Nichrome. Mechanically, it is less workable than Nichrome. **Platinum** is used in specialized heating applications where very high temperatures are required. **Tungsten** may be used in very high temperature ovens with an inert atmosphere. **Pure nickel** is used to satisfy the high requirements in the fabrication of radio tubes, such as the elimination of all gases and impurities in the metal parts. It also has other uses such as in incandescent lamps, for combustion boats, laboratory accessories, and resistance thermometers.

Carbon withstands high temperatures and has high resistance; its temperature coefficient is negative; it will safely carry about 125 A/in^2. Amorphous carbon has a resistivity of 3,800 and 4,100 $\mu\Omega \cdot$cm, retort carbon about 720 $\mu\Omega \cdot$cm, and graphite about 812 $\mu\Omega \cdot$cm. The properties of any particular kind of carbon depend on the temperature at which it was fired. Carbon for rheostats may best be used in the form of compression rheostats. **Silicon carbide** is used to manufacture heating rods that will safely operate at 1,650°C (3,000°F) surface temperature. It has a negative coefficient of resistance up to 650°C, after which it is positive. It must be mechanically protected because of inherent brittleness.

Table 15.1.27 Properties of Metals, Alloys, and Resistor Materials

Material	Composition	Sp gr	Resistivity: Microhms-cm at 20 C	Resistivity: Ohms cir-mil-ft at 20 C	Resistivity: Temp coef of resistance per deg C	Temp range, deg C	Max safe working temp, deg C	Approx melting point, deg C
Advance	Cu 0.55; Ni 0.45	8.9	48.4	294	±0.00002	20–100	500	1210
Comet	Ni 0.30; Cr 0.05; Fe 0.65	8.15	95	570	0.00088	20–500	600	1480
Bronze, commercial	Cu; Zn	8.7	4.2	25	0.0020	0–100	...	1040
Hytemco-Balco	Ni 0.50; Fe 0.50	8.46	20	120	0.0045	20–100	600	1425
Kanthal A	Al 0.055; Cr 0.22; Co 0.055; Fe 0.72	7.1	145	870	0.00002	0–500	1330	1510
Magno	Ni 0.955; Mn 0.045	8.75	20	120	0.0036	20–100	400	1435
Manganin	Cu 0.84; Mn 0.12; Ni 0.04	8.19	48.2	290	±0.000015	15–35	100	1020
Monel metal	Ni 0.67; Cu 0.28	8.9	42.6	256	0.0001	0–100	425	1350
Nichrome	Ni 0.60; Fe 0.25; Cr 0.15	8.247	112	675	0.00017	20–100	930	1350
Nichrome V	Ni 0.80; Cr 0.20	8.412	108	650	0.00013	20–100	1100	1400
Nickel, pure	Ni 0.99	8.9	10	60	0.0050	0–100	400	1450
Platinum	Pt	21.45	10.616	63.80	0.003	0–100	1200	1773
Silver	Ag	10.5	1.622	9.755	.00361	0–100	650	960
Tungsten	W	19.3	5.523	33.22	0.0045	......	*	3410

*Tungsten subject to rapid oxidation in air above 150°C.

MAGNETS

A **permanent magnet** is one that retains a considerable amount of magnetism indefinitely. Permanent magnets are used in electrical instruments, telephone receivers, loudspeakers, magnetos, tachometers, magnetic chucks, motors, and for many purposes where a constant magnetic field or a constant source of magnetism is desired. The magnetic material should have high retentivity, a high remanence, and a high coercive force (see Fig. 15.1.8). These properties are usually found with hardened steel and its alloys and also in ceramic permanent magnet materials.

Since permanent magnets must operate on the molecular mmf imparted to them when magnetized, they must necessarily operate on the portion *CDO* of the hysteresis loop (see Fig. 15.1.8). The area *CDO* is proportional to the **stored energy** within the magnet and is a criterion of its usefulness as a permanent-magnetic material. In the left half of Fig. 15.1.97 are given the *B-H* characteristics of several permanent-magnetic materials; these include 5 to 6 percent tungsten steel (curve 1); 3½ percent chrome magnet steel (curve 2); cobalt magnet steel, containing 16 to 36 percent cobalt and 5 to 9 percent chromium and in some alloys tungsten (curve 3); and the carbon-free aluminum-nickel-cobalt-steel alloys called Alnico. There are many grades of Alnico; the characteristics of three of them are shown by curves 4, 5, 6. Their composition is as follows:

Curve	Alnico no.	Composition, percent Al	Ni	Co	Cu	Ti	Fe
4	5	8	14	24	3	. . .	51
5	6	8	14	24	3	1.25	49.75
6	12	6	18	35	. . .	8.0	33

All the Alnicos can be made by the sand or the precision-casting (lost-wax) process, but the most satisfactory method is by the sintering process.

If the alloys are held in a magnetic field during heat-treatment, a **magnet grain** is established and the magnetic properties in the direction of the field are greatly increased. The alloys are hard, can be formed only by casting or sintering, and cannot be machined except by grinding.

The curves in the right half of Fig. 15.1.97 are "external energy" curves and give the product of *B* and *H*. The optimum point of operation is at the point of maximum energy as is indicated at A_1 on curve 5.

Considering curve 5, if the magnetic circuit remained closed, the magnet would operate at point *B*. To utilize the flux, an air gap must be introduced. The air gap acts as a demagnetizing force, H_1 ($= B_1A_1$), and the magnet operates at point A_1 on the *HB* curve. The line OA_1 is called the *air-gap line* and its slope is given by $\tan \theta_1 = B_1/H_1$ where $H_1 = B_g l_g / l_m$ and $B_1/B_g = A_g/A_m$, where B_g = flux density in gap, T; l_g = length of gap, m; l_m = length of magnet, m; A_g, A_m = areas of the gap and magnet, m².

If the air gap is lengthened, the magnet will operate at A_2 corresponding to a lesser flux density B_3 and the new air-gap line is OA_2. If the gap is now closed to its original value, the magnet will not return to operation at point A_1 but will operate at some point *C* on the line OA_1. If the air gap is varied between the two foregoing values, the magnet will operate along the **minor hysteresis loop** A_2C. Return to point A_1 can be accomplished only by remagnetizing and coming back down the curve from *B* to A_1.

Alnico magnets corresponding to curves 4 and 5 are best adapted to operation with short air gaps, since the introduction of a long air gap will demagnetize the magnet materially. On

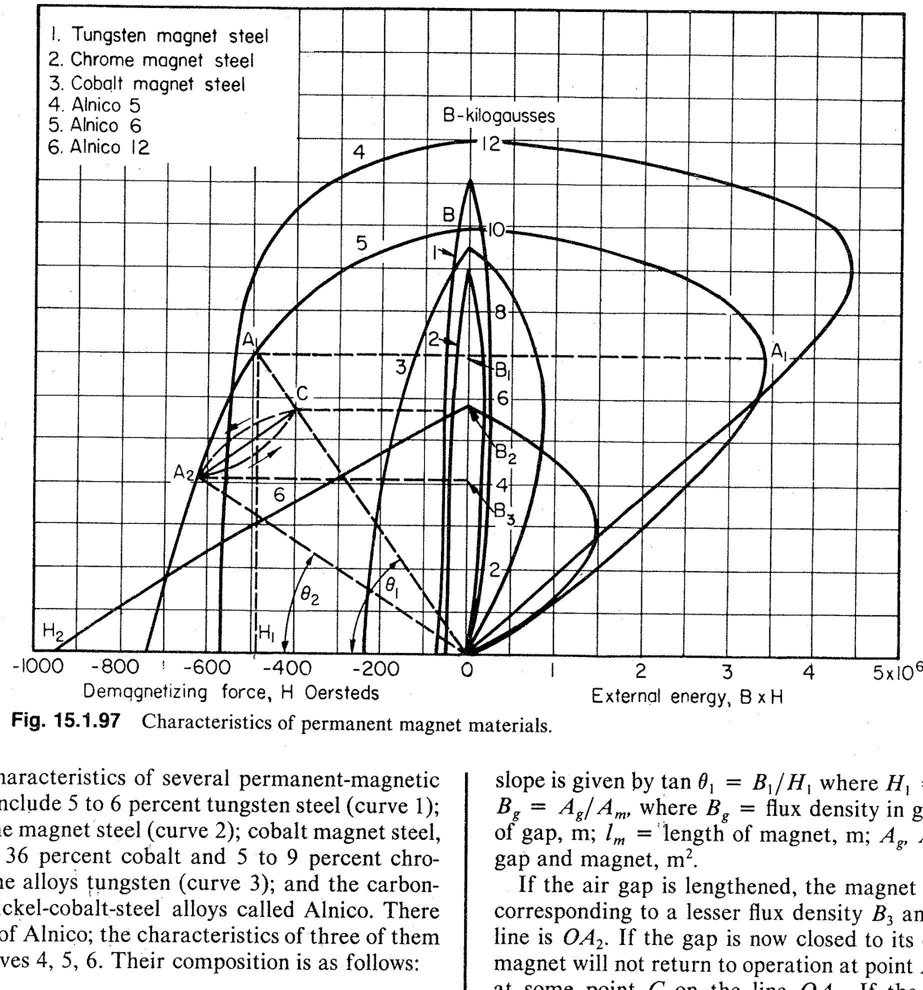


Fig. 15.1.97 Characteristics of permanent magnet materials.

the other hand, a magnet with a long air gap will operate most satisfactorily on curve 6 on account of the high coercive force H_2. With change in the length of the air gap, the operation will be essentially along that curve and the magnet will lose little of its original magnetization.

There are several other grades of Alnico with characteristics between curves 4, 5, and 6. Ceramic PM materials have a very large coercive force.

The steels for permanent magnets are cut in strips, heated to a red-hot temperature, and forged into shape, usually in a "bulldozer." If they are to be machined, they are cooled in mica dust to prevent air hardening. They are then ground, tumbled, and tempered. Alnico and ceramic types are cast and then finish-ground.

Permanent magnets are magnetized either by placing them over a bus bar carrying a large direct current, by placing them across the poles of a powerful electromagnet, or by an ampere-turn pulse.

Unless permanent magnets are subjected to artificial aging, they gradually weaken until after a long period they become stabilized usually at from 85 to 90 percent of their initial strength. With magnets for electrical instruments, where a constant field strength is imperative, artificial aging is accomplished by mechanical vibration or by immersion in oil at 250°F for a period of a few hours.

In an **electromagnet** the magnetic field is produced by an electric current. The core is usually made of soft iron or mild steel because, the permeability being higher, a stronger magnetic field may be obtained. Also since the retentivity is low, there is little trouble due to the sticking of armatures when the circuit is opened. Electromagnets may have the form of simple solenoids, iron-clad solenoids, plunger electromagnets, electromagnets with external armatures, and lifting magnets, which are circular in form with a flat holding surface.

A **solenoid** is a winding of insulated conductor and is wound helically; the direction of winding may be either right or left. A **portative electromagnet** is one designed only for holding material brought in contact with it. A **tractive electromagnet** is one designed to exert a force on the load through some distance and thus do work. The **range** of an electromagnet is the distance through which the plunger will perform work when the winding is energized. For long range of operation, the plunger type of tractive magnet is best suited, for the length of core is governed practically by the range of action desired, and the area of the core is determined by the pull. **Solenoid and plunger** is a solenoid provided with a movable iron rod or bar called a plunger. When the coil is energized, the iron rod becomes magnetized and the mutual action of the field in the solenoid on the poles created on the plunger causes the plunger to move within the solenoid. This force becomes zero only when the magnetic centers of the plunger and solenoid coincide. If the load is attached to the plunger, work will be done until the force to be overcome is equal to the force that the solenoid exerts on the plunger. When the iron of the plunger is not saturated, the strength of magnetic field in the solenoid and the induced poles are both proportional to the exciting current, so that the pull varies as the **current squared.** When the plunger becomes highly saturated, the pull varies almost **directly** with the current.

The **maximum uniform pull** occurs when the end of the plunger is at the center of the solenoid and is equal to

$$F = CAnI/l \text{ lb} \qquad \textbf{(15.1.134)}$$

where A = cross-sectional area of plunger, in^2; n = number of turns; I = current, A; l = length of the solenoid, in; and C = pull, (lb/in^2)/(A-turn/in). C depends on the proportions of the coil, the degree of saturation, the length, and the physical and chemical purity of the plunger. Table 15.1.28 gives values of C for several different solenoids.

Curve 1, Fig. 15.1.98, shows the characteristic pull of an open-magnetic circuit solenoid, 12 in long, having 10,000 A-turns or 833 A-turns/in.

When a **strong pull** is desired at the end of the stroke, a stop may be used as shown in Fig. 15.1.99. Curve 2, Fig. 15.1.98, shows the pull obtained by adding a stop to the plunger. It will be noted that, except when the end of the plunger is near the stop, the stop adds little to the solenoid pull. The pull is made up of two components: one due to the attraction between plunger and winding, the other to the attraction between plunger and stop. The equation for the pull is

$$P = AIn[(In/l_a^2C_1^2) + (C/l)] \qquad \textbf{(15.1.135)}$$

where A = area of the core, in^2; n = number of turns; l_a = length of gap between core and stop; and C, C_1 = constants. At the beginning of the stroke the second member of the equation is predominant, and at the end of the stroke the first member represents practically the entire pull. Approximate values of C and C_1 are C_1 = 2,660 (for l greater than $10d$), C = 0.0096, where d is the diameter of the plunger, in. In SI units

$$P = 1.7512AnI\left(\frac{2.54nI}{l_a^2C_1^2} + \frac{C}{l}\right) \text{ N} \qquad \textbf{(15.1.135a)}$$

where A is in cm^2, l and l_a in cm, and the pull P in N. All other quantities are unchanged.

The **range of uniform** pull can be extended by the use of conical ends of stop and plunger, as shown in Fig. 15.1.100. A

Table 15.1.28 Maximum Pull per Sq. In. of Core for Solenoids with Open Magnetic Circuit

Length of coil l, in.	Length of plunger, in.	Core area A, sq in.	Total ampere-turns $I \times n$	Max pull P, psi	$1{,}000 \times C$	Length of coil l, in.	Length of plunger, in.	Core area A, sq in.	Total ampere-turns $I \times n$	Max pull P, psi	$1{,}000 \times C$
6	Long	1.0	15,900	22.4	9.0	12	Long	1.0	11,200	8.75	9.4
9	Long	1.0	11,330	11.5	9.1	12	Long	1.0	20,500	16.75	9.8
9	Long	1.0	14,200	14.6	9.2	18	36	1.0	18,200	9.8	9.7
10	10	2.76	40,000	40.2	10.0	18	36	1.0	41,000	22.5	9.8
10	10	2.76	60,000	61.6	10.3	18	18	1.0	18,200	9.8	9.7
10	10	2.76	80,000	80.8	10.1	18	18	1.0	41,000	22.5	9.8

SOURCE: From data by Underhill, *Elec. World,* 45, 1906, pp. 796, 881.

stronger magnet mechanically can be obtained by using an iron-clad solenoid, Fig. 15.1.101, in which an iron return path is provided for the flux. Except for low flux densities and short air gaps the dimensions of the iron return path are of no prac-

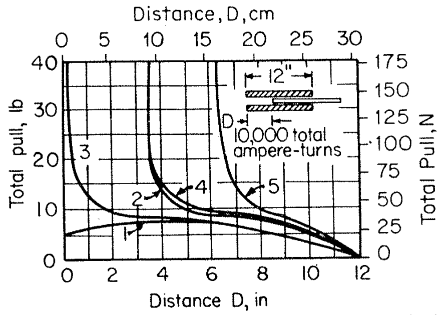


Fig. 15.1.98 Pull on solenoid on plunger. (1) Coil and plunger; (2) coil and plunger with stop; (3) iron-clad coil and plunger; (4) and (5) same as (3) with different lengths of stop.

tical importance, and the fact that an iron return path is used does not affect the pull curve except at short air gaps. This is illustrated in Fig. 15.1.98 where curves 3, 4, and 5 are typical pull curves for this same solenoid when it is made iron-clad, each curve corresponding to a different position of the stop.

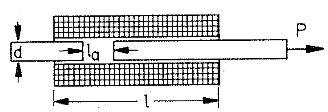


Fig. 15.1.99 Solenoid with stop.

Mechanical jar at the end of the stroke may be prevented by leaving the end of the solenoid open. The plunger then comes to equilibrium when its middle is at the middle of the winding, thus providing a magnetic cushion effect. Electromagnets with

Fig. 15.1.100 Conical plunger and stop.

external armatures are best adapted for **short-range** work, and the best type is the horseshoe magnet. The pull for short-range magnets is expressed by the equation

$$F = B^2A/72{,}134{,}000 \quad \text{lb} \qquad \textbf{(15.1.136)}$$

where B = flux density, Mx/in^2, and A = area of the core, in^2.

In SI units

$$F = 397{,}840B^2A \quad \text{N} \qquad \textbf{(15.1.136a)}$$

where the flux density B is in Wb/m^2; A = area, m^2; and F = force, N.

A greater holding power is obtained if the surfaces of the armature and core are not machined to an absolutely smooth

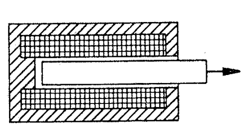


Fig. 15.1.101 Iron-clad solenoid.

contact surface. If the surface is slightly irregular, the area of contact A is reduced but the flux density B is increased approximately in proportion (if the iron is being operated below saturation) and the pull is increased since it varies as the square of the density B. Nonmagnetic stops should be used if it is desired that the armature may be released readily when the current is interrupted.

Lifting magnets are of the portative type in that their function is merely to hold the load. The actual lifting is performed by the hoisting apparatus. The magnet is almost toroidal in shape. The coil shield is of manganese steel which is very hard and thus resists wear and is practically nonmagnetic. The holding power is given by Eq. **(15.1.136)**, where A = area of holding surface, in^2. It is difficult to calculate accurately the holding force of a lifting magnet for it depends on the magnetic characteristics of the load, the area of contact, and the manner in which the load is applied.

Rapid action in a magnet can be obtained by reducing the time constant of the winding and by subdividing the metal parts to reduce induced currents which have a demagnetizing effect when the circuit is closed. The movement of the plunger through the winding causes the winding and its bobbin to be cut by a magnetic field; if the bobbin is of metal and not slotted longitudinally, it is a short-circuited turn linked by a changing magnetic field and hence currents are induced in it. These currents oppose the flux and hence reduce the pull during the transient period. They also cause some heating. Where it is found impossible to reduce the time constant sufficiently, an electromagnet designed for a voltage much lower than normal is often used. A resistor is connected in series which is short-circuited during the stroke of the plunger. At the completion of the stroke the plunger automatically opens the short circuit, reducing the current to a value which will not overheat the magnet under continuous operation. The extremely short time of overload produces very rapid action but does not injure the winding. The solenoids on many automatic motor-starting panels are designed in this manner.

When **slow action** is desired, it can be obtained by using solid cores and yoke and by using a heavy metallic spool or bobbin for the winding. A separate winding short-circuited on itself is also used to some extent.

Sparking at switch terminals may be reduced or eliminated by neutralizing the inductance of the winding. This is accomplished by winding a separate short-circuited coil with its wires parallel to those of the active winding. (This method can be used with dc magnets only.) This is not economical, since one-half the winding space is wasted. By connecting a capacitor across the switch terminals, the energy of the inductive discharge on opening the circuit may be absorbed. For the purpose of neutralizing the inductive discharge and causing a quick release, a small **reverse current** may be sent through the coil winding automatically on opening the circuit. Sleeves of tin, aluminum, or copper foil placed over the various layers of the winding absorb energy when the circuit is broken and reduce the energy dissipated at the switch terminals. This scheme can be used for dc magnets only. **Sticking** of the parts of the magnetic circuit due to residual magnetism may be prevented by the use of nonmagnetic stops. In the case of lifting magnets subjected to rough usage and hard blows (as in a steel works), these stops usually consist of plates of manganese steel, which are extremely hard and nonmagnetic.

AC Tractive Magnets Because of the iron losses due to eddy currents, the magnetic circuits of ac electromagnets should be composed of laminated iron or steel. The magnetic circuit of large magnets is usually built up of thin sheets of sheet metal held together by means of suitable clamps. Small cores of circular cross section usually consist of a bundle of soft iron wires. Since the iron losses increase with the flux density, it is not advisable to operate at as high a density as with direct current. The current instead of being limited by the resistance of the winding is now determined almost entirely by the inductive reactance as the resistance is small. With the removal of the load the current rises to high values. The pull of ac magnets is nearly constant irrespective of the length of air gap.

In a **single-phase magnet** the pull varies from zero to a maximum and back to zero twice every cycle, which may cause considerable chattering of the armature against the stop. This may be prevented by the use of a spring or, in the case of a solenoid coil, by allowing the plunger to seek its position of equilibrium in the coil. **Chattering** may also be prevented by the use of a short-circuited winding or shading coil around one tip of the pole piece or by the use of polyphase. In a **two-phase magnet** the pull is constant and equal to the maximum instantaneous pull produced by one phase so long as the voltage is a sine function. In a **three-phase magnet** under the same conditions the pull is constant and equal to 1.5 times the maximum instantaneous pull of one phase. Should the load become greater than the minimum instantaneous pull, there will be chattering as in a single-phase magnet.

Heating of Magnets The lifting capacity of an electromagnet is limited by the permissible current-carrying capacity of the winding, which, in turn, is dependent on the amount of heat energy that the winding can dissipate per unit time without exceeding a given temperature rise. **Enamels, synthetic varnishes,** and **thermoplastic wire insulations** are available to allow a wide variety of temperature rises.

Design of Exciting Coil Let n = number of turns, l = mean length of turn, in ($l = 2\pi r$, where r is the mean radius, in), A = cross section of wire, cir mils. The resistance of 1 cir mil·ft of copper is practically 12 Ω at 60°C, or 1 Ω/cir mil·in. Hence the resistance, $R = nl/A$ Ω; the current, $I = EA/nl$; the ampere-turns, $nI = EA/l$; the power to be dissipated, $P = E^2A/nl$ W. From the foregoing equations the cross section of wire and the number of turns can be calculated.

Space Factor of Winding Space factor of a coil is the ratio of the space occupied by the conductor to the total volume of the coil or winding. Only in the theoretical case of uninsulated square or rectangular conductor may the space factor be 100 percent. For wire of circular section with insulation of negligible thickness, wound as shown in Fig. 15.1.102*a*, the space

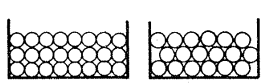


Fig. 15.1.102 Winding space factor.

factor will be 78.5 percent. When the turns of wire are "bedded," as shown in Fig. 15.1.102*b* (the case in most windings, particularly with smaller wires), there is a theoretical gain of about 7 percent in space factor. Experiments have shown that in most cases this gain is about neutralized in practice by the flattening out of the insulation of the wire due to the tension used in winding. When wound in a haphazard manner, the space factors of magnet wires vary according to size, substantially as follows:

	Double cotton covered				*Single cotton covered*			
Size, AWG	0	5	10	15	20	25	30	35
Space factor, percent	60	53.8	45.5	35.1	32.2	32	25.7	16

Magnet wire is usually a soft insulated annealed copper wire of high conductivity. It can be obtained in square, rectangular, and circular section, but the round or cylindrical wire is used almost entirely in the smaller sizes. Ribbons are frequently used in the larger sizes. **Aluminum** has been used at times. A number of different varnish or enamel insulating compounds are available in different temperature classes, up to 220°C, the **temperature** rating and thickness for **dielectric strength** being dependent on the usage. Where **mechanical strength** is important and space is not at a premium, textile or paper insulation is used on the wires and later varnish impregnated for high dielectric strength. **Asbestos-covered** magnet wire is used where the temperature is high. It combines resistance to heat and abrasion, can resist mild acids, has good dielectric strength, and is fireproof.

Table 15.1.29 gives the diameters of magnet wire with the different types of insulation. For further data on electrical insulating materials, see Sec. 6.

AUTOMOBILE SYSTEMS

Automobile Ignition Systems

The ignition system in an automobile produces the spark which ignites the combustible mixture in the engine cylinders. This is accomplished by a high-voltage, or high-tension, spark between metal points in a spark plug. (A spark plug is an insulated bushing screwed into the cylinder head.) Spark plugs usually have porcelain insulation, but for some special uses, such as in airplane engines, mica may be used. There are two general sources for the energy necessary for ignition; one is the electrical system of the car which is maintained by the generator and the battery **(battery ignition),** and the other is a **magneto.** Battery ignition systems have traditionally operated electromechanically, using a spark coil, a high-voltage distributor, and low-voltage breaker points. Electronic ignition systems, working from the battery, became standard on U.S. cars in 1975. These vary in complexity from the use of a single transistor to reduce the current through the points to pointless systems triggered by magnetic pulses or interrupted light beams. Capacitor discharge into a pulse transformer is used in some systems to obtain the high voltage needed to fire the spark plugs.

Battery ignition is most widely used since it is simple, reliable, and low in cost, and the electrical system is a part of the car equipment. The high voltage for the spark is obtained from an ignition coil which consists of a primary coil of relatively few turns and a secondary coil of a large number of turns, both coils being wound on a common magnetic core consisting of either thin strips of iron or small iron wires. In a 6-V system the resistance of the primary coil is from 0.9 to 2 Ω and the inductance is from 5 to 10 mH. The number of secondary turns varies from 9,000 to 25,000, and the ratio of primary to secondary turns varies from 1:40 to 1:100.

The coil operates on the following principle. It stores energy

Table 15.1.29 Magnet-Wire Dimensions, Sizes 14 to 44 AWG

	Bare wire diameter, in			Single		Heavy	
AWG	Minimum	Nominal	Maximum	Minimum increase in diameter, in	Maximum overall diameter, in	Minimum increase in diameter, in	Maximum overall diameter, in.
14	0.0635	0.0641	0.0644	0.0016	0.0666	0.0032	0.0682
15	0.0565	0.0571	0.0574	0.0015	0.0594	0.0030	0.0609
16	0.0503	0.0508	0.0511	0.0014	0.0531	0.0029	0.0545
17	0.0448	0.0453	0.0455	0.0014	0.0475	0.0028	0.0488
18	0.0399	0.0403	0.0405	0.0013	0.0424	0.0026	0.0437
19	0.0355	0.0359	0.0361	0.0012	0.0379	0.0025	0.0391
20	0.0317	0.0320	0.0322	0.0012	0.0339	0.0023	0.0351
21	0.0282	0.0285	0.0286	0.0011	0.0303	0.0022	0.0314
22	0.0250	0.0253	0.0254	0.0011	0.0270	0.0021	0.0281
23	0.0224	0.0226	0.0227	0.0010	0.0243	0.0020	0.0253
24	0.0199	0.0201	0.0202	0.0010	0.0217	0.0019	0.0227
25	0.0177	0.0179	0.0180	0.0009	0.0194	0.0018	0.0203
26	0.0157	0.0159	0.0160	0.0009	0.0173	0.0017	0.0182
27	0.0141	0.0142	0.0143	0.0008	0.0156	0.0016	0.0164
28	0.0125	0.0126	0.0127	0.0008	0.0140	0.0016	0.0147
29	0.0112	0.0113	0.0114	0.0007	0.0126	0.0015	0.0133
30	0.0099	0.0100	0.0101	0.0007	0.0112	0.0014	0.0119
31	0.0088	0.0089	0.0090	0.0006	0.0100	0.0013	0.0108
32	0.0079	0.0080	0.0081	0.0006	0.0091	0.0012	0.0098
33	0.0070	0.0071	0.0072	0.0005	0.0081	0.0011	0.0088
34	0.0062	0.0063	0.0064	0.0005	0.0072	0.0010	0.0078
35	0.0055	0.0056	0.0057	0.0004	0.0064	0.0009	0.0070
36	0.0049	0.0050	0.0051	0.0004	0.0058	0.0008	0.0063
37	0.0044	0.0045	0.0046	0.0003	0.0052	0.0008	0.0057
38	0.0039	0.0040	0.0041	0.0003	0.0047	0.0007	0.0051
39	0.0034	0.0035	0.0036	0.0002	0.0041	0.0006	0.0045
40	0.0030	0.0031	0.0032	0.0002	0.0037	0.0006	0.0040
41	0.0027	0.0028	0.0029	0.0002	0.0033	0.0005	0.0036
42	0.0024	0.0025	0.0026	0.0002	0.0030	0.0004	0.0032
43	0.0021	0.0022	0.0023	0.0002	0.0026	0.0004	0.0029
44	0.0019	0.0020	0.0021	0.0001	0.0024	0.0004	0.0027

SOURCE: "Standard Handbook for Electrical Engineers," Fink and Carrol, McGraw-Hill, New York, 1968.

in a magnetic field relatively slowly and then releases it suddenly. The power developed ($p = dw/dt$) is thus relatively large (w = stored energy). The high emf e_2 which is required for the spark is induced by the sudden change in the flux ϕ in the core of the coil when the primary current is suddenly interrupted, $e_2 = -n_2(d\phi/dt)$, where n_2 is the number of secondary turns. For satisfactory ignition, peak voltages from 10 to 20 kV volts are desirable. Figure 15.1.103 shows the relation

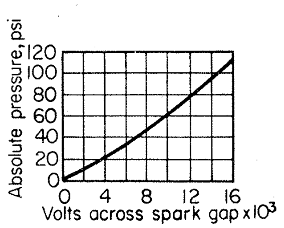


Fig. 15.1.103 Pressure-voltage curve for spark plug.

between the volts required to produce a spark and pressure with compressed air.

A battery ignition system for a four-cylinder engine is shown diagrammatically in Fig. 15.1.104. The primary circuit supplied by the battery consists of the primary coil P and a set of contacts, or "points" operated by a four-lobe cam, in series. In

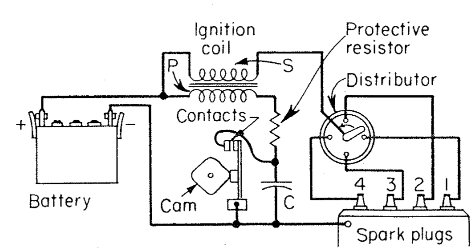


Fig. 15.1.104 Battery-ignition system.

order to reduce arcing and burning of the contacts and to produce a sharp break in the current, a capacitor C is connected across the contacts. The contacts, which are of pure tungsten, are operated by a four-lobe cam which is driven at one-half engine speed. A strong spring tends to keep the contacts closed.

In the Delco-Remy distributor (Fig. 15.1.105) two breaker arms are connected in parallel; one coil and one capacitor are used. One set of contacts is open when the other is just breaking but closes a few degrees after the break occurs. This closes the primary of the ignition coil immediately after the break and increases the time that the primary of the ignition coil is closed and permits the flux in the iron to reach its full value. The interrupter shown in Fig. 15.1.105 is designed for an eight-cylinder engine.

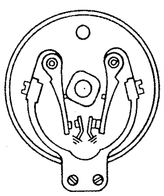


Fig. 15.1.105 Delco-Remy eight-cylinder interrupter.

Electronic ignition systems have no problem operating at the required speed.

The spark should advance with increase in engine speed so as to allow for the time lag in the explosion. To take care of this **automatically** most timers are now equipped with centrifugally operated weights which advance the breaker cam with respect to the engine drive as the speed increases.

Automobile Lighting and Starting Systems

Automobile lighting and starting systems initially operated at 6 V, but at present nearly all cars, except the smaller ones, operate at 12 V because larger engines, particularly V-8s, are now common and require more starting power. With 12 V, for the same power, the starting current is halved, and the effect of resistance in the leads, connections, and brushes is materially reduced. In some systems the positive side of the system is grounded, but more often the negative side is grounded.

A further development is the application of an **ac generator**, or alternator, combined with a rectifier as the generating unit rather than the usual dc generator. One advantage is the elimination of the commutator, made up of segments, which requires some maintenance due to the sparking and wear of the carbon brushes. With the alternator the dc field rotates, the brushes operating on smooth slip rings require almost no maintenance. Also, the system is greatly simplified by the fact that rectifiers are "one-way" devices, and the battery **cannot** deliver current back to the generator when its voltage drops below that of the battery. Thus, no cutout relay, such as is required with dc generators, is necessary. This ac development is the result of the development of reliable, low-cost germanium and silicon semiconductor rectifiers.

Figure 15.1.106 shows a schematic diagram of the Ford system (adapted initially to trucks). The generator stator is wound three-phase Y-connected, and the field is bipolar supplied with direct current through slip rings and brushes. The rectifier diodes are connected full-wave bridge circuit to supply the battery through the ammeter.

Regulator The function of the regulator is to control the generator current so that its value is adapted to the battery voltage which is related to the condition of charge of the battery (see Fig. 15.1.15). Thus, when the battery voltage drops (indicating a lowered condition of charge), the current should be increased, and, conversely, when the battery voltage increases (indicating a high condition of charge), the current should be decreased.

Neglecting for the moment the starting procedure, when the ignition switch is thrown to the normal "on" position at *c*, the coil actuating the field relay is connected to the battery + terminal and causes the relay contacts *a* to close. This energizes the regulator circuits, and, if the two upper voltage regulator contacts *b* are closed as shown, the rotating field of the alternator is connected directly to the battery + terminal, and the field current is then at its maximum value and produces a high generator voltage and large output current. At the same time the voltage regulator coil in series with the 0.3- and 14-Ω resistors is connected between the battery + terminal and ground. If the voltage of the battery rises owing to its higher condition of charge, the current to the voltage regulator coil increases, causing it to open the two upper contacts at *b*. Current from the battery now flows through the 0.3-Ω resistor and divides,

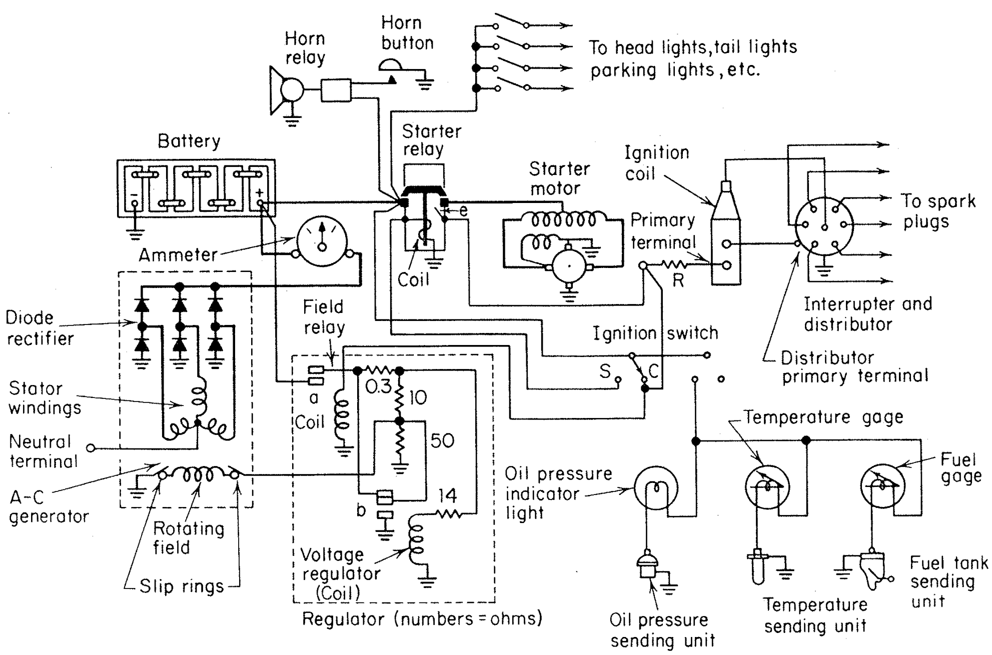


Fig. 15.1.106 Schematic diagram of Ford lighting and starting system.

some going through the 10-Ω resistor and dividing between the field and the 50-Ω resistor, and the remainder going to the voltage regulator coil. The current to the rotating field is thus reduced, causing the alternator output to be reduced. Because of the 0.3 Ω now in circuit, the current to the coil of the voltage regulator is reduced to such a value that it holds the center contact at *b* in the mid, or open, position. If the battery voltage rises to an even higher value, the regulator coil becomes strong enough to close the lower contacts at *b;* this short-circuits the field, reducing its current almost to zero, and thus reducing the alternator output to zero. On the other hand when the battery voltage drops, the foregoing sequence is reversed, and the contacts at *b* operate to increase the current to the alternator field.

As was mentioned earlier, the battery cannot supply current to the alternator because of the "one-way" characteristic of the rectifier. Thus, when the alternator voltage drops below that of the battery and even when the alternator stops running, its current automatically becomes zero. The alternator has a normal rectifier open-circuit voltage of about 14 V and a rating of 20 A.

Starting In most cases, for starting, the ignition key is turned far to the right and held there until the motor starts. Then, when the key is released, the ignition switch contacts assume a normal operating position. Thus, in Fig. 15.1.106, when the ignition-switch contact is in the starting position *S*, the starter relay coil becomes connected by a lead to the battery + terminal and thus becomes energized, closing the relay contacts. The starter motor is then connected to the battery to crank the engine. At the time that the contact closes it makes contact with a small metal brush *e* which connects the battery + terminal to the primary terminal of the ignition coil through the protective resistor *R*. After the motor starts, the ignition switch contacts spring to the normal operating position *C*, and the starter relay switch opens, thereby breaking contact with the small brush *e*. However, when contact *C* is closed, the ignition coil primary terminal is now connected through leads to the battery + terminal. The interrupter, the ignition coil, and the distributor now operate in the manner described earlier (see Fig. 15.1.104); the system shown in Fig. 15.1.106, is that for a six-cylinder engine.

The connection of **accessories** to the electric system is illustrated in Fig. 15.1.106 for the horn, head and other lights, and temperature and fuel gages.

Magneto Ignition

Principle of Magneto A magneto is an electric generator in which the magnetic flux is provided by one or more permanent magnets. It is a self-contained unit and is used advantageously for ignition where a generator and a battery are not needed to supply power to other accessories. The design of magnetos was radically changed when Alnico, with its very high retentivity (Fig. 15.1.97), became available as a permanent-magnet material. One method of utilizing Alnico magnets is to insert the bar magnets in the frame of the magneto (Fig. 15.1.107). The rotor is a soft-iron bobbin. A primary winding of relatively few turns and a secondary winding of a relatively large number of turns are wound over the laminated yoke *Y*. The position of the rotor shown in (*a*) provides a low reluctance path for the magnetic flux of the left-hand magnet and a high reluctance path for the magnetic flux of the right-hand magnet so that the flux goes through the yoke from left to right as shown. When the rotor turns one-eighth of a revolution, it becomes horizontal; obviously, each of the magnets acts in opposition relative to the yoke, and the flux therein becomes zero. In (*b*), which also shows the external electrical connections, the rotor is shown as having turned one-fourth of a revolution, or 90°, from its position in (*a*).

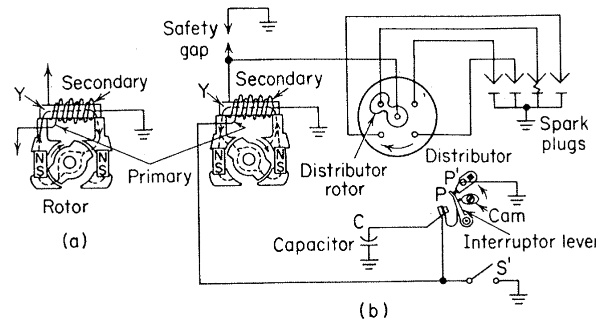


Fig. 15.1.107 Magneto-ignition system.

The rotor now provides a low-reluctance path for the right-hand magnet and a high-reluctance path for the left-hand one. Thus the magnetic flux now goes through the yoke from right to left. It follows that in each 90° interval the flux in the yoke undergoes a complete reversal. It will be recognized that this is an **inductor** type of ac generator.

In the diagram in (*b*), one end each of the primary and of the secondary are grounded together. The other end of the primary is connected to an insulated interrupter lever having a contact point *P*. This makes intermittent contact with the grounded contact point *P′*. The contact point *P* is actuated by a cam which is driven by the same shaft as the magneto rotor and is in a definite relation to it. A switch *S′* is provided to ground and thus short-circuit the secondary when it is desired to stop the engine. A capacitor *C* is connected between the point *P* and ground to absorb the energy of the spark which occurs when the contacts open.

With the contacts closed, the primary is short-circuited, and the varying flux in the core produced by the rotation of the rotor induces an alternating current in the winding which in turn produces an alternating flux in the core. With the rotation of the rotor the current in the secondary rises cyclically to maximum values, and at these instants the cam causes the points *PP′* to open suddenly, interrupting the current in the primary and thus causing a sudden collapse of the flux in the core. This induces a high-impulse emf in the secondary which is transmitted to the distributor and thence to the proper spark plug as shown.

On starting, the speed of the magneto may be so low that the emf is not sufficient to produce a hot spark. This difficulty can be met by impulse starting, in which the rotor is driven through a spring. During cranking the rotor is restrained from turning until the engine comes to the proper firing position, at which time the rotor is suddenly released. The energy stored in the spring produces a high, instantaneous, angular velocity to the rotor, resulting in a high emf and hot spark.

Inductor-type magnetos, having a large number of rotor poles and arranged differently from those shown in Fig. 15.1.107, are used for airplane-engine ignition. In another magneto design the rotor is a solid cylindrical Alnico magnet, permanently magnetized with an N and an S pole diametrically opposite. The frame is laminated, and there is a yoke with primary and secondary windings. When the rotor rotates, its

N and S poles produce an alternating flux in the yoke which induces a short-circuit current in the primary winding and the method of producing the spark is then the same as with Fig. 15.1.107*b*.

Miscellaneous Automobile Electronics Systems

The use of electronics on automobiles is expanding rapidly. **Microprocessors** are being installed to control many functions that were previously not controlled or poorly controlled.

Air-fuel ratio is controlled from sensing oxygen in the exhaust system. **Timing** is based on crankshaft position, acceleration, and engine temperature. **Air injection** into exhaust and recirculation of exhaust reduces emissions. **Knock** is prevented by stopping untimely detonations. **Diagnosis** of problems is performed. Car leveling and load matching are made by shock absorber adjustments. Wheel **spin** is prevented. **Transmission** controls provide improved efficiency. Operator interfaces are by **cathode ray tube, fluid crystal, vacuum fluorescent displays,** and **speech synthesis. Navigation aids** are in a rudimentary state.

15.2 ELECTRONICS

by Byron M. Jones

REFERENCES: "Reference Data for Radio Engineers," Howard Sams & Co. "Transistor Manual," General Electric Co. "SCR Manual," General Electric Co. "The Semiconductor Data Book," Motorola Inc. Fink, "Television Engineering," McGraw-Hill. "Industrial Electronics Reference Book," Wiley. Mano, "Digital and Logic Design," Prentice-Hall. McNamara, "Technical Aspects of Data Communication," Digital Press. Fletcher, "An Engineering Approach to Digital Design," Prentice-Hall. "The TTL Data Book," Texas Instruments, Inc. "1982 MOS Products Catalog," American Microsystems, Inc. "1982 Linear Data Book," National Semiconductor Corp.

The subject of electronics can be approached from the standpoint of the design of devices or the use of devices. For the practicing mechanical engineer, describing devices in terms of their external characteristics seems most likely to be profitable. The approach will be to describe devices as they appear to the outside world.

Components

Resistors, capacitors, reactors, and **transformers** are described earlier in this section, along with basic circuit theory. These explanations are equally applicable to electronic circuits and hence are not repeated here. A description of additional components peculiar to electronic circuits follows.

A **rectifier,** or **diode,** is an electronic device which offers unequal resistance to forward and reverse current flow. Figure 15.2.1 shows the schematic symbol for a diode. The arrow

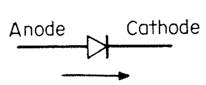


Fig. 15.2.1 Diode schematic symbol.

beside the diode shows the direction of current flow. Current flow is taken to be the flow of positive charges, i.e., the arrow is counter to electron flow. Figure 15.2.2 shows typical forward and reverse volt-ampere characteristics. Notice that the scales for voltage and current are not the same for the first and third quadrants. This has been done so that both the forward and reverse characteristics can be shown on a single plot even though they differ by several orders of magnitude.

Diodes are rated for forward current capacity and reverse voltage breakdown. They are manufactured with maximum current capabilities ranging from 0.05 A to more than 1,000 A. Reverse voltage breakdown varies from 50 V to more than 2,500 V. At rated forward current, the forward voltage drop

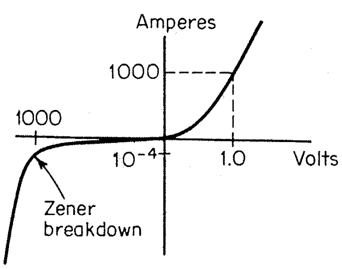


Fig. 15.2.2 Diode forward-reverse characteristic.

varies between 0.7 and 1.5 V for silicon diodes. Although other materials are used for special-purpose devices, by far the most common semiconductor material is silicon. With a forward current of 1000 A and a forward voltage drop of 1 V, there would be a power loss in the diode of 1000 W (more than 1 hp). The basic diode package shown in Fig. 15.2.3 can dissipate about 20 W. To maintain an acceptable temperature in the diode, it is necessary to mount the diode on a **heat sink.** The manufacturer's recommendation should be followed very carefully to ensure good heat transfer and at the same time avoid fracturing the silicon chip inside the diode package.

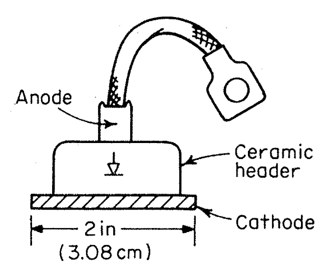


Fig. 15.2.3 Physical diode package.

The selection of **fuses** or **circuit breakers** for the protection of **rectifiers** and rectifier circuitry requires more care than for other electronic devices. Diode failures as a result of circuit faults occur in a fraction of a millisecond. Special semiconductor fuses have been developed specifically for semiconductor circuits. Proper protective circuits must be provided for the protection of not only semiconductors but also the rest of the circuit and nearby personnel. Diodes and diode fuses have a short-circuit rating in amperes-squared-seconds (I^2t). As long as the I^2t rating of the diode exceeds the I^2t rating of its protective fuse, the diode and its associated circuitry will be protected. Circuit breakers may be used to protect diode circuits, but additional line impedance must be provided to limit the current while the circuit breaker clears. Circuit breakers do not interrupt the current when their contacts open. The fault is not cleared until the line voltage reverses at the end of the cycle of the applied voltage. This means that the **clearing time for a circuit breaker** is about ½ cycle of the ac input voltage. Diodes have a 1-cycle overcurrent rating which indicates the fault current the diode can carry for circuit breaker protection schemes. Line inductance is normally provided to limit fault currents for breaker protection. Often this inductance is in the form of leakage reactance in the transformer which supplies power to the diode circuit.

A **thyristor,** often called a **silicon controlled rectifier (SCR),** is a rectifier which blocks current in both the forward and reverse directions. Conduction of current in the forward direction will occur when the anode is positive with respect to the cathode and when the gate is pulsed positive with respect to the cathode. Once the thyristor has begun to conduct, the gate pulse can return to 0 V or even go negative and the thyristor will continue to pass current. To stop the cathode-to-anode current, it is necessary to reverse the cathode-to-anode voltage. The thyristor will again be able to block both forward and reverse voltages until current flow is initiated by a gate pulse. The schematic symbol for an SCR is shown in Fig. 15.2.4. The physical packaging of thyristors is similar to that of rectifiers with similar ratings except, of course, that the thyristor must have an additional gate connection.

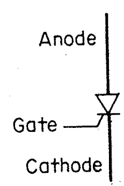


Fig. 15.2.4 Thyristor schematic symbol.

The gate pulse required to fire an SCR is quite small compared with the anode voltage and current. Power gains in the range of 10^6 to 10^9 are easily obtained. In addition, the power loss in the thyristor is very low, compared with the power it controls, so that it is a very efficient power-controlling device. Efficiency in a thyristor power supply is usually 97 to 99 percent. When the thyristor blocks either forward or reverse current, the high voltage drop across the thyristor accompanies low current. When the thyristor is conducting forward current after having been fired by its gate pulse, the high anode current occurs with a forward voltage drop of about 1.5 V. Since high voltage and high current never occur simultaneously, the power dissipation in both the on and off states is low.

The thyristor is rated primarily on the basis of its forward-current capacity and its voltage-blocking capability. Devices are manufactured to have equal forward and reverse voltage-blocking capability. Like diodes, thyristors have I^2t ratings and 1-cycle surge current ratings to allow design of protective circuits. In addition to these ratings, which the SCR shares in common with diodes, the SCR has many additional specifications. Because the thyristor is limited in part by its average current and in part by its rms current, forward-current capacity is a function of the duty cycle to which the device is subjected. Since the thyristor cannot regain its blocking ability until its anode voltage is reversed and remains reversed for a short time, this time must be specified. The time to regain blocking ability after the anode voltage has been reversed is called the **turn-off time.** Specifications are also given for minimum and maximum gate drive. If forward blocking voltage is reapplied too quickly, the SCR may fire with no applied gate voltage pulse. The maximum safe value of rate of reapplied voltage is called the dv/dt rating of the SCR. When the gate pulse is applied, current begins to flow in the area immediately adjacent to the gate junction. Rather quickly, the current spreads across the entire cathode-junction area. In some circuits associated with the thyristor an extremely fast rate of rise of current may occur. In this event localized heating of the cathode may occur with a resulting immediate failure or in less extreme cases a slow degradation of the thyristor. The maximum rate of change of current for a thyristor is given by its di/dt rating. Design for di/dt and dv/dt limits is not normally a problem at power-line frequencies of 50 and 60 Hz. These ratings become a design factor at frequencies of 500 Hz and greater. Table 15.2.1 lists typical thyristor characteristics.

A **triac** is a bilateral SCR. It blocks current in either direc-

Table 15.2.1 Typical Thyristor Characteristics

Voltage	Current, A rms	Current, A avg	$I^2_{\frac{1}{2}}t$, A·s	1-cycle surge, A	di/dt, A/s	dv/dt, V/s	Turn-off time, s
400	35	20	165	180	100	200	10
1,200	35	20	75	150	100	200	10
400	110	70	4,000	1,000	100	200	40
1,200	110	70	4,000	1,000	100	200	40
400	235	160	32,000	3,500	100	200	80
1,200	235	160	32,000	3,500	75	200	80
400	470	300	120,000	5,500	50	100	150
1,200	470	300	120,000	5,500	50	100	150

tion until it receives a gate pulse. It can be used to control in ac circuits. Triacs are widely used for light dimmers and for the control of small universal ac motors. The triac must regain its blocking ability as the line voltage crosses through zero. This fact limits the use of triacs to 60 Hz and below.

A **transistor** is a semiconductor amplifier. The schematic symbol for a transistor is shown in Fig. 15.2.5. There are two

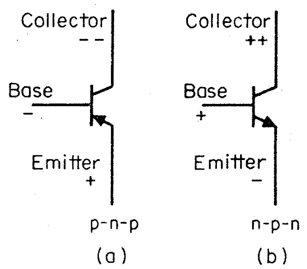


Fig. 15.2.5 Transistor schematic symbol.

types of transistors, *p-n-p* and *n-p-n*. Notice that the polarities of voltage applied to these devices are opposite. In many sizes matched *p-n-p* and *n-p-n* devices are available. The most common transistors have a collector dissipation rating of 150 to 600 mW. Collector to base breakdown voltage is 20 to 50 V. The amplification or gain of a transistor occurs because of two facts: (1) A small change in current in the base circuit causes a large change in current in the collector and emitter leads. This current amplification is designated *hfe* on most transistor specification sheets. (2) A small change in base-to-emitter voltage can cause a large change in either the collector-to-base voltage or the collector-to-emitter voltage. Table 15.2.2 shows basic ratings for some typical transistors. There is a great profusion of transistor types so that the choice of type depends upon availability and cost as well as operating characteristics.

The gain of a transistor is independent of frequency over a wide range. At high frequency, the gain falls off. This cutoff frequency may be as low as 20 kHz for audio transistors or as high as 1 GHz for radio-frequency (rf) transistors.

The schematic symbols for the field effect transistor (FET) is shown Fig. 15.2.6. The flow of current from source to drain is controlled by an electric field established in the device by the voltage applied between the gate and the drain. The effect of this field is to change the resistance of the transistor by altering its internal current path. The FET has an extremely high gate resistance (10^{12} Ω), and as a consequence, it is used for applications requiring high input impedance. Some FETs have been designed for high-frequency characteristics. Other FETs have been designed for high-power applications. The two basic constructions used for FETs are **bipolar junctions** and **metal oxide**

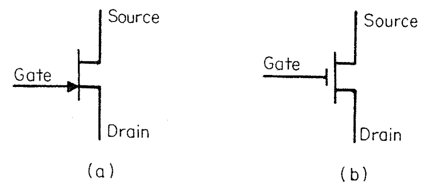


Fig. 15.2.6 Field effect transistor. (*a*) Bipolar junction type (JFET); (*b*) metal oxide semiconductor type (MOSFET).

semiconductors. The schematic symbols for each of these are shown Fig. 15.2.6*a* and 15.2.6*b*. These are called JFETs and MOSFETs to distinguish between them. JFETs and MOSFETs are used as stand-alone devices and are also widely used in integrated circuits. (See below, this section.)

The **unijunction** is a special-purpose semiconductor device. It is a pulse generator and widely used to fire thyristors and triacs as well as in timing circuits and waveshaping circuits. The schematic symbol for a unijunction is shown in Fig. 15.2.7. The device is essentially a silicon resistor. This resistor is connected to base 1 and base 2. The emitter is fastened to this resistor about halfway between bases 1 and 2. If a positive voltage is applied to base 2, and if the emitter and base 1 are at zero, the emitter junction is back-biased and no current flows in the emitter. If the emitter voltage is made increasingly positive, the emitter junction will become forward-biased. When this occurs, the resistance between base 1 and base 2 and between base 2 and the emitter suddenly switches to a very low value. This is a regenerative action, so that very fast and very energetic pulses can be generated with this device.

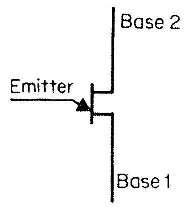


Fig. 15.2.7 Unijunction.

Before the advent of semiconductors, electronic rectifiers and amplifiers were **vacuum tubes** or **gas-filled tubes.** Some use of these devices still remains. If an electrode is heated in a vacuum, it gives up surface electrons. If an electric field is established between this heated electrode and another electrode so that the electrons are attracted to the other electrode, a current will flow through the vacuum. Electrons flow from the heated

Table 15.2.2 Typical Transistor Characteristics

JEDEC number	Type	Collector-emitter volts at breakdown, BV_{CE}	Collector dissipation, P_c (25°C)	Collector current, I_c	Current gain, h_{fe}
2N3904	*n-p-n*	40	310 mW	200 mA	200
2N3906	*p-n-p*	40	310 mW	200 mA	200
2N3055	*n-p-n*	100	115 W	15 A	20
2N6275	*n-p-n*	120	250 W	50 A	30
2N5458	JFET	40	200 mW	9 mA	*
2N5486	JFET	25	200 mW		†

*JFET, current gain is not applicable.
†High-frequency JFET—up to 400 MHz.

cathode to the cold anode. If the polarity is reversed, since there are no free electrons around the anode, no current will flow. This, then, is a vacuum-tube rectifier. If a third electrode, called a **control grid,** is placed between the cathode and the anode, the flow of electrons from the cathode to the anode can be controlled. This is a basic vacuum-tube amplifier. Additional grids have been placed between the cathode and anode to further enhance certain characteristics of the vacuum tube. In addition, multiple anodes and cathodes have been enclosed in a single tube for special applications such as radio signal converters.

If an inert gas, such as neon or argon, is introduced into the vacuum, conduction can be initiated from a cold electrode. The breakdown voltage is relatively stable for given gas and gas pressure and is in the range of 50 to 200 V. The **nixie** display tube is such a device. This tube contains 10 cathodes shaped in the form of the numerals from 0 to 9. If one of these cathodes is made negative with respect to the anode in the tube, the gas in the tube glows around that cathode. In this way each of the 10 numerals can be made to glow when the appropriate electrode is energized.

An **ignitron** is a vapor-filled tube. It has a pool of liquid mercury in the bottom of the tube. Air is exhausted from the enclosure, leaving only mercury vapor, which comes from the pool at the bottom. If no current is flowing, this tube will block voltage whether the anode is plus or minus with respect to the mercury-pool cathode. A small rod called an **ignitor** can form a cathode spot on the pool of mercury when it is withdrawn from the pool. The ignitor is pulled out of the pool by an electromagnet. Once the cathode spot has been formed, electrons will continue to flow from the mercury-pool cathode to the anode until the anode-to-cathode voltage is reversed. The operation of an ignitron is very similar to that of a thyristor. The anode and cathode of each device perform similar functions. The ignitor and gate also perform similar functions. The thyristor is capable of operating at much higher frequencies than the ignitron and is much more efficient since the thyristor has 1.5 V forward drop and the ignitron has 15 V forward drop. The ignitron has an advantage over the thyristor in that it can carry extremely high overload currents without damage. For this reason ignitrons are often used as electronic "crowbars" which discharge electrical energy when a fault occurs in a circuit.

Discrete-Component Circuits

Several common rectifier circuits are shown in Fig. 15.2.8. The waveforms shown in this figure assume no line reactance. The presence of line reactance will make a slight difference in the waveshapes and the conversion factors shown in Fig. 15.2.8. These waveshapes are equally applicable for loads which are pure resistive or resistive and inductive. In a resistive load the current flowing in the load has the same waveshape as the voltage applied to it. For inductive loads, the current waveshape will be smoother than the voltage applied. If the inductance is high enough, the ripple in the current may be indeterminantly small. An approximation of the ripple current can be calculated as follows:

$$I = \frac{E_{dc}\,PCT}{200\pi fNL} \qquad (15.2.1)$$

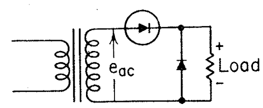
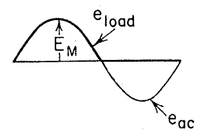
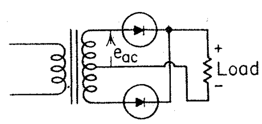
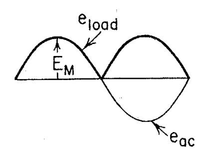
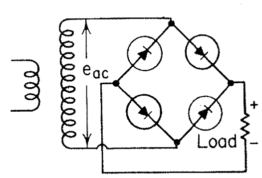
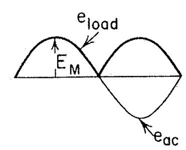
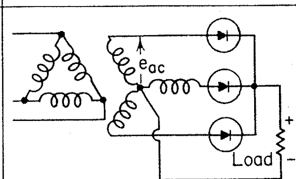
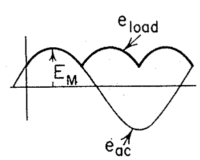
where I = rms ripple current, E_{dc} = dc load voltage, PCT = percent ripple from Fig. 15.2.8, f = line frequency, N = number of cycles of ripple frequency per cycle of line frequency, L = equivalent series inductance in load. Equation (**15.2.1**) will always give a value of ripple higher than that calculated by more exact means, but this value is normally satisfactory for power-supply design.

Capacitance in the load leads to increased regulation. At light loads, the capacitor will tend to charge up to the peak value of the line voltage and remain there. This means that for either the single full-wave circuit or the single-phase bridge the dc output voltage would be 1.414 times the rms input voltage. As the size of the loading resistor is reduced, or as the size of the parallel load capacitor is reduced, the load voltage will more nearly follow the rectified line voltage and so the dc voltage will approach 0.9 times the rms input voltage for very heavy loads or for very small filter capacitors. One can see then that dc voltage may vary between 1.414 and 0.9 times line voltage due only to waveform changes when **capacitor filtering** is used.

Four different **thyristor rectifier circuits** are shown in Fig. 15.2.9. These circuits are equally suitable for resistive or inductive loads. It will be noted that the half-wave circuit for the thyristor has a rectifier across the load, as in Fig. 15.2.8. This diode is called a **freewheeling diode** because it freewheels and carries inductive load current when the thyristor is not conducting. Without this diode, it would not be possible to build up current in an inductive load. The gate-control circuitry is not shown in Fig. 15.2.9 in order to make the power circuit easier to see. Notice the location of the thyristors and rectifiers in the single-phase full-wave circuit. Constructed this way, the two diodes in series perform the function of a freewheeling diode. The circuit can be built with a thyristor and rectifier interchanged. This would work for resistive loads but not for inductive loads. For the full three-phase bridge, a freewheeling diode is not required since the carryover from the firing of one SCR to the next does not carry through a large portion of the negative half cycle and therefore current can be built up in an inductive load.

Capacitance must be used with care **in thyristor circuits.** A capacitor directly across any of the circuits in Fig. 15.2.9 will immediately destroy the thyristors. When an SCR is fired directly into a capacitor with no series resistance, the resulting di/dt in the thyristor causes extreme local heating in the device and a resultant failure. A sufficiently high series resistor prevents failure. An inductance in series with a capacitor must also be used with caution. The series inductance may cause the capacitor to "ring up." Under this condition, the voltage across the capacitor can approach twice peak line voltage or 2.828 times rms line voltage.

The advantage of the thyristor circuits shown in Fig. 15.2.9 over the rectifier circuits is, of course, that the thyristor circuits provide variable output voltage. The output of the thyristor circuits depends upon the magnitude of the incoming line voltage and the phase angle at which the thyristors are fired. The control characteristic for the thyristor power supply is determined by the waveshape of the output voltage and also by the phase-shifting scheme used in the firing-control means for the thyristor. Practical and economic power supplies usually have control characteristics with some degree of nonlinearity. A representative characteristic is shown in Fig. 15.2.10. This control characteristic is usually given for nominal line voltage with the tacit understanding that variations in line voltage will cause approximately proportional changes in output voltage.

Type	Circuit	Output voltage waveform	E_{dc} (avg)	Ripple fundamental frequency	% ripple	Peak inverse voltage
Half-wave 1ϕ			0.318 E_M 0.45 E_{ac}	F	121	3.14 E_{dc}
Full-wave 1ϕ			0.636 E_M 0.9 E_{ac}	2F	48	3.14 E_{dc}
Bridge 1ϕ			0.636 E_M 0.9 E_{ac}	2F	48	1.57 E_{dc}
Half-wave 3ϕ			0.827 E_M 1.17 E_{ac}	3F	18	2.09 E_{dc}

E_M = maximum value of e_{ac}
E_{ac} = effective value of e_{ac}
E_{dc} = average value of d-c load voltage
F = line frequency
% ripple = 100 x rms ripple / E_{dc}

Fig. 15.2.8 Comparison of rectifier circuits.

Transistor amplifiers can take many different forms. A complete discussion is beyond the scope of this handbook. The circuits described here illustrate basic principles. A basic **single-stage amplifier** is shown in Fig. 15.2.11. The transistor can be cut off by making the input terminal sufficiently negative. It can be saturated by making the input terminal sufficiently positive. In the linear range, the base of an *n-p-n* transistor will be 0.5 to 0.7 V positive with respect to the emitter. The collector voltage will vary from about 0.2 V to V_c (20 V, typically). Note that there is a sign inversion of voltage between the base and the collector; i.e., when the base is made more positive, the collector becomes less positive. The resistors in this circuit serve the following functions. Resistor $R1$ limits the input current to the base of the transistor so that it is not harmed when the input signal overdrives. Resistors $R2$ and $R3$ establish the transistor's operating point with no input signal. Resistors $R4$ and $R5$ determine the voltage gain of the amplifier. Resistor $R4$ also serves to stabilize the zero-signal operating point, as established by resistors $R2$ and $R3$. Usual practice is to design single-stage gains of 10 to 20. Much higher gains are possible to achieve, but low gain levels permit the use of less expensive transistors and increase circuit reliability.

Figure 15.2.12 illustrates a basic **two-stage transistor amplifier** using complementary *n-p-n* and *p-n-p* transistors. Note that the first stage is identical to that shown in Fig. 15.2.11. This *n-p-n* stage drives the following *p-n-p* stage. Additional alternate *n-p-n* and *p-n-p* stages can be added until any desired overall amplifier gain is achieved.

Figure 15.2.13 shows the **Darlington** connection of transistors. The amplifier is used to obtain maximum current gain from two transistors. Assuming a base-to-collector current gain of 50 times for each transistor, this circuit will give an input-to-output current gain of 2,500. This high level of gain is not very stable if the ambient temperature changes, but in many cases this drift is tolerable.

Figure 15.2.14 shows a circuit developed specifically to minimize temperature drift and drift due to power supply voltage changes. The **differential amplifier** minimizes drift because of the balanced nature of the circuit. Whatever changes in one transistor tend to increase the output are compensated by reverse

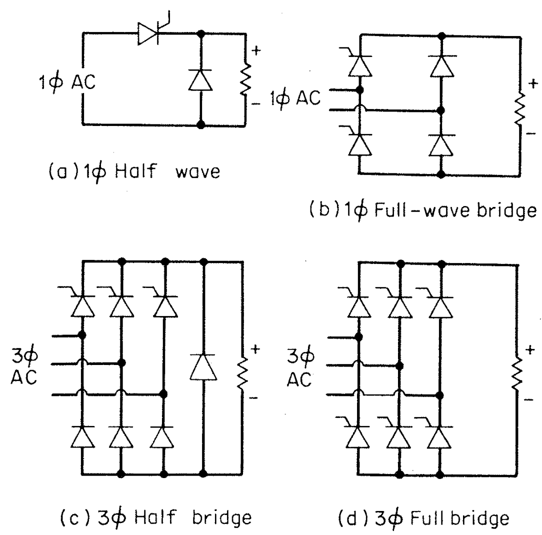


Fig. 15.2.9 Basic thyristor circuits.

trends in the second transistor. The input signal does not affect both transistors in compensatory ways, of course, and so it is amplified. One way to look at a differential amplifier is that twice as many transistors are used for each stage of amplifi-

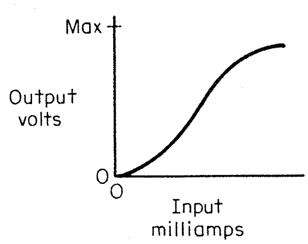


Fig. 15.2.10 Thyristor control characteristic.

cation to achieve compensation. For very low drift requirements, matched transistors are available. For the ultimate in differential amplifier performance, two matched transistors are encapsulated in a single unit. **Operational amplifiers** made with discrete components frequently use differential amplifiers to minimize drift and offset. The operational amplifier is a low-drift, high-gain amplifier designed for a wide range of control and instrumentation uses.

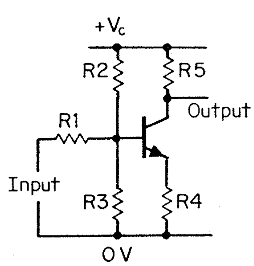


Fig. 15.2.11 Single-stage amplifier.

Oscillators are circuits which provide a frequency output with no signal input. A portion of the collector signal is fed back to the base of the transistor. This feedback is amplified by the transistor and so maintains a sustained oscillation. The frequency of the oscillation is determined by parallel inductance and capacitance. The oscillatory circuit consisting of an inductance and a capacitance in parallel is called an LC **tank circuit.**

This frequency is approximately equal to

$$f = 1/2\pi\sqrt{CL} \tag{15.2.2}$$

where f = frequency, Hz; C = capacitance, F; L = inductance, H. A 1-MHz oscillator might typically be designed with

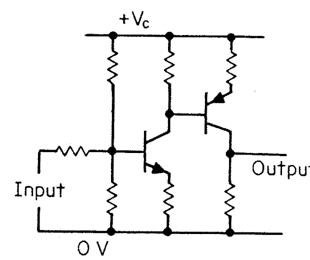


Fig. 15.2.12 Two-stage amplifier.

a 20-μH inductance in parallel with a 0.05-μF capacitor. The exact frequency will vary from the calculated value because of loading effects and stray inductance and capacitance. The **Colpitts** oscillator shown in Fig. 15.2.15 differs from the **Hartley**

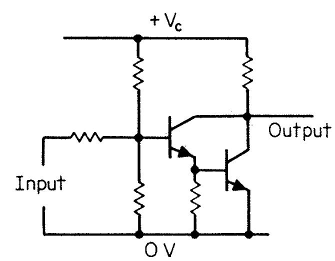


Fig. 15.2.13 Darlington connection.

oscillator shown in Fig. 15.2.16 only in the way energy is fed back to the emitter. The Colpitts oscillator has a capacitive voltage divider in the resonant tank. The Hartley oscillator has an inductive voltage divider in the tank. The **crystal oscillator**

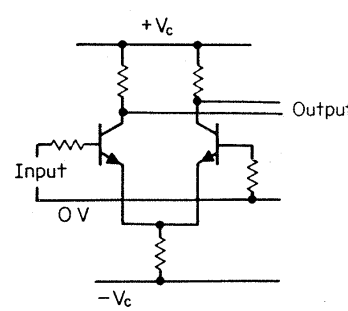


Fig. 15.2.14 Differential amplifier.

shown in Fig. 15.2.17 has much greater frequency stability than the circuits in Figs. 15.2.15 and 15.2.16. Frequency stability of 1 part in 10^7 is easily achieved with a crystal-controlled oscillator. If the oscillator is temperature-controlled by mounting it in a small temperature-controlled oven, the frequency stability can be increased to 1 part in 10^9. The resonant LC tank in the collector circuit is tuned to approximately the crystal frequency. The crystal offers a low impedance at its resonant frequency. This pulls the collector-tank operating frequency to the crystal resonant frequency.

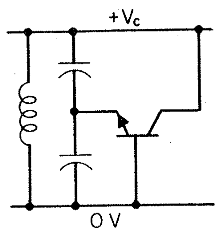


Fig. 15.2.15 Colpitts oscillator.

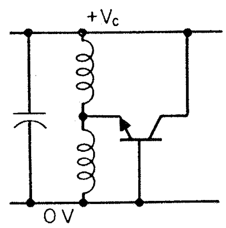


Fig. 15.2.16 Hartley oscillator.

As the desired operating frequency becomes 500 MHz and greater, **resonant cavities** are used as tank circuits instead of discrete capacitors and inductors. A rough guide to the relationship between frequency and resonant-cavity size is the wavelength of the frequency

$$\lambda = 300 \times 10^6/f \tag{15.2.3}$$

where λ = wavelength, m; 300×10^6 = speed of light, m/s; f = frequency, Hz. The resonant cavities will be smaller than

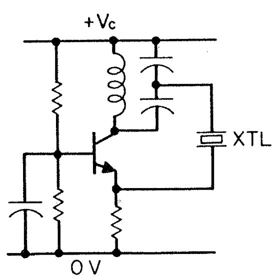


Fig. 15.2.17 Crystal-controlled oscillator.

indicated by Eq. (15.2.3) because in general the cavity is either one-half or one-fourth wavelength and also, in general, the electromagnetic wave velocity is less in a cavity than in free space.

The operating principles of these devices are beyond the scope of this article. There are many different kinds of **microwave tubes** including **klystrons, magnetrons,** and **traveling-wave** tubes. All these tubes employ moving electrons to excite a resonant cavity. These devices serve as either oscillators or amplifiers at microwave frequencies.

Lasers operate at approximately visible-light frequency of 600 THz. This corresponds to a wavelength of 0.5 μm, or, the more usual measure of visible-light wavelength, 5×10^3 Å. Resonant cavities simply cannot be made small enough for these wavelengths. Electronic resonance in the atom serves as the tank circuit for these high frequencies. Quantum mechanics must be employed properly to explain these devices, but a practical understanding can be achieved without delving so deep. Most light is disorganized insofar as the axis of vibration and the frequency of vibration are concerned. When radiation along different axes is attenuated, as with a polarizing screen, the light wave is said to be polarized. When white light is filtered, or when the light source is not white, the light is frequency-limited, or colored. Colored light still has a relatively wide band of frequencies. The laser emits a very narrow band of frequencies, which are extremely stable, many times more stable than a crystal; therefore lasers are used as frequency standards. The narrow frequency band of lasers allows focusing the output into extremely small beams. This feature makes the laser attractive as a cutting tool and as an accurate surveying device. Extremely sharp focus and extremely high frequency make it attractive as a high-density communications carrier. Experimental work is being done with **phase-locked lasers,** which not only have a single frequency of output but have output oscillations in phase with each other. This degree of organization promises further commercial development of laser devices.

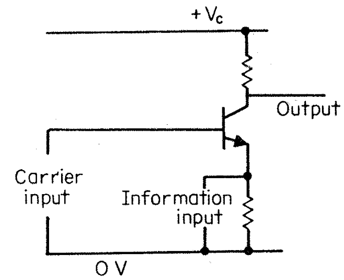


Fig. 15.2.18 AM modulator.

A radio wave consists of two parts, a **carrier,** and an information signal. The carrier is a steady high frequency. The information signal may be a voice signal, a video signal, or telemetry information. The carrier wave can be modulated by varying its amplitude or by varying its frequency. **Modulators** are circuits which impress the information signal onto the carrier. A **demodulator** is a circuit in the receiving apparatus which separates the information signal from the carrier. A simple amplitude modulator is shown in Fig. 15.2.18. The transistor is base-driven with the carrier input and emitter driven with the information signal. The modulated carrier wave appears at the collector of the transistor. An FM modulator is shown in Fig. 15.2.19. The carrier must be changed in frequency in

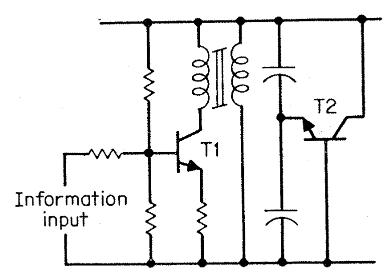


Fig. 15.2.19 FM modulator.

response to the information signal input. This is accomplished by using a saturable ferrite core in the inductance of a Colpitts oscillator which is tuned to the carrier frequency. As the collector current in transistor $T1$ varies with the information signal, the saturation level in the ferrite core changes, which in turn varies the inductance of the winding in the tank circuit and alters the operating frequency of the oscillator.

The **demodulator** for an AM signal is shown in Fig. 15.2.20. The diode rectifies the carrier plus information signal so that the filtered voltage appearing across the capacitor is the information signal. Resistor $R2$ blocks the carrier signal so that the

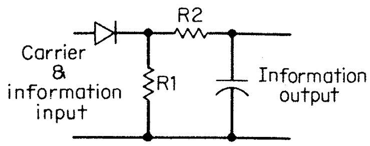


Fig. 15.2.20 AM demodulator.

output contains only the information signal. An FM **demodulator** is shown in Fig. 15.2.21. In this circuit, the carrier plus information signal has a constant amplitude. The information is in the form of varying frequency in the carrier wave. If inductor $L1$ and capacitor $C1$ are tuned to near the carrier frequency but not exactly at resonance, the current through resistor $R1$ will vary as the carrier frequency shifts up and down. This will create an AM signal across resistor $R1$. The diode, resistors $R2$ and $R3$, and capacitor $C2$ demodulate this signal as in the circuit in Fig. 15.2.20.

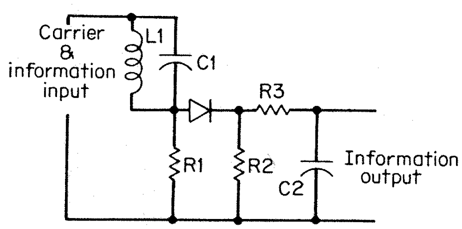


Fig. 15.2.21 FM discriminator.

The waveform of the basic **electronic timing circuit** is shown in Fig. 15.2.22 along with a basic timing circuit. Switch $S1$ is closed from time $t1$ until time $t3$. During this time, the transistor shorts the capacitor and holds the capacitor at 0.2 V. When switch $S1$ is opened at time $t3$, the transistor ceases to conduct and the capacitor charges exponentially due to the current flow through resistor $R1$. Delay time can be measured

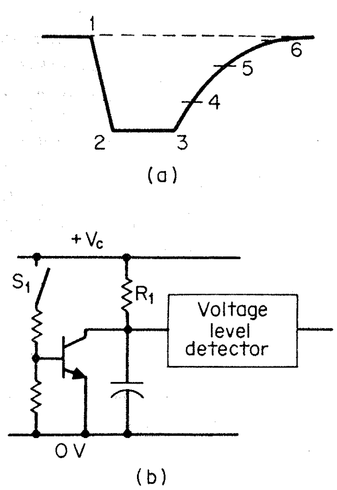


Fig. 15.2.22 Basic timing circuit.

to any point along this exponential charge. If the time is measured until time $t6$, the timing may vary due to small shifts in supply voltage or slight changes in the voltage-level detecting circuit. If time is measured until time $t4$, the voltage level will be easy to detect, but the obtainable time delay from time $t3$ to time $t4$ may not be large enough compared with the reset time $t1$ to $t2$. Considerations like these usually dictate detecting at time $t4$. If this time is at a voltage level with is 63 percent of V_c, the time from $t3$ to $t4$ is one time constant of $R1$ and C. This time can be calculated by

$$t = RC \tag{15.2.4}$$

where t = time, s; R = resistance, Ω; C = capacitance, F. A timing circuit with a 0.1-s delay can be constructed using a 0.1-μF capacitor and a 1.0-MΩ resistor.

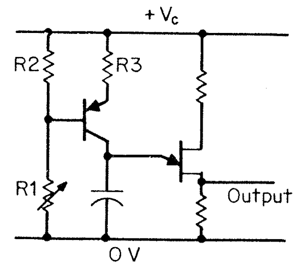


Fig. 15.2.23 Improved timing circuit.

An improved timing circuit is shown in Fig. 15.2.23. In this circuit, the unijunction is used as a level detector, a pulse generator, and a reset means for the capacitor. The transistor is used as a constant current source for charging the timing capacitor. The current through the transistor is determined by resistors $R1$, $R2$, and $R3$. This current is adjustable by means of $R1$. When the charge on the capacitor reaches approximately 50 percent of V_c, the unijunction fires, discharging the capacitor and generating a pulse at the output. The discharged capacitor is then recharged by the transistor, and the cycle continues to repeat. The pulse rate of this circuit can be varied from one pulse per minute to many thousands of pulses per second.

Integrated Circuits

Table 15.2.3 lists some of the more common physical packages for discrete component and integrated semiconductor devices. Although discrete components are still used for electronic

Table 15.2.3 Semiconductor Physical Packaging

Signal devices	
Plastic	TO92
Metal can	TO5, TO18, TO39
Power devices	
Tab mount	TO127, TO218, TO220
Diamond case	TO3, TO66
Stud mount	
Flat base	
Flat pak (Hockey puck)	
Integrated circuits	
Dip (dual in-line pins)	(See Fig. 15.2.24.)
Flat pack	
Chip carrier (50-mil centers)	

design, **integrated circuits (ICs)** are becoming predominant in almost all types of electronic equipment. Dimensions of common dual in-line pin (DIP) integrated-circuit devices are shown in Fig. 15.2.24. An IC costs far less than circuits made

Number of pins	*W*	*P*	*L*	Approximate height
64	0.8	0.1	3.3	0.24
40	0.6	0.1	2.0	0.2
28	0.6	0.1	1.4	0.2
24	0.6	0.1	1.3	0.2
22	0.4	0.1	1.1	0.2
20	0.3	0.1	1.0	0.2
18	0.3	0.1	0.9	0.2
16	0.3	0.1	0.87	0.2
14	0.3	0.1	0.78	0.2
8	0.3	0.1	0.4	0.2

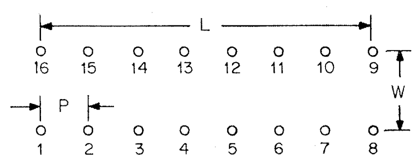


Fig. 15.2.24 Approximate physical dimensions of dual in-line pin (DIP) integrated circuits. All dimensions are in inches. Dual in-line packages are made in three different constructions—molded plastic, cerdip, and ceramic.

with discrete components. Integrated circuits can be classified in several different ways. One way to classify them is by complexity. **Small-scale integration (SSI), medium-scale integration (MSI), large-scale integration (LSI),** and **very large scale integration (VLSI)** refer to this kind of classification. The cost and availability of a particular IC are more dependent upon the size of the market for that device than on the level of its internal complexity. For this reason, the classification by circuit complexity is not as meaningful today as it once was. The literature still refers to these classifications, however. For the purpose of this text, ICs will be separated into two broad classes: linear ICs and digital ICs.

The trend in IC development has been toward greatly increased complexity at significantly reduced cost. Present-day ICs are manufactured with internal spacings as low as 2 μm. The limitation of the contents of a single device is more often controlled by external connections than by internal space. For this reason, more and more complex combinations of circuits are being interconnected within a single device. There is also a tendency to accomplish functions digitally that were formerly done by analog means. Although these digital circuits are much more complex than their analog counterparts, the cost and reliability of ICs make the resulting digital circuit the preferred design. One can expect these trends will continue based on current technology. One can also anticipate further declines in price versus performance. It has been demonstrated again and again that digital IC designs are much more stable and reliable than analog designs.

Linear Integrated Circuits

The basic building block for many linear ICs is the **operational amplifier.** Table 15.2.4 lists the basic characteristics for a few representative IC operational amplifiers. In most instances, an adequate design for an operational amplifier circuit can be made assuming an "ideal" operational amplifier. For an ideal operational amplifier, one assumes that it has infinite gain and no voltage drop across its input terminals. In most designs, feedback is used to limit the gain of each operational amplifier. As long as the resulting closed-loop gain is much less than the open-loop gain of the operational amplifier, this assumption yields results that are within acceptable engineering accuracy. Operational amplifiers use a balanced input circuit which minimizes input voltage offset. Furthermore, specially designed operational amplifiers are available which have extremely low input offset voltage. The input voltage must be kept low because of temperature drift considerations. For these reasons, the assumption of zero input voltage, sometimes called a "virtual ground," is justified. Figure 15.2.25 shows three operational amplifier circuits and the equations which describe their behavior. In this figure S is the **Laplace transform** variable. **Active filters** are designed using operational amplifiers with associated resistors and capacitors in a manner similar to that shown in Fig. 15.2.25.

Table 15.2.5 lists some typical linear ICs, most of which contain operational amplifiers. The **voltage comparator** is an operational amplifier that compares two input voltages and provides an output that indicates which of the two voltages is greater. The **sample-and-hold** circuit samples an analog input at prescribed intervals. Between these sample times, it holds the last value it measured. This circuit is used to convert signals from analog to digital form. The NE 555 timer/oscillator is adaptable for a wide variety of applications. It can be used as a stable, adjustable frequency free-running or **monostable multivibrator.** It can also be used as a **linear ramp generator.** It can be used for time delay or sequential timing applications.

Table 15.2.6 lists linear ICs that are used in audio, radio, and television circuits. The degree of complexity that can be incorporated in a single device is illustrated by the fact that a complete AM-FM radio circuit is available in a single IC device. The **phase-locked loop** is a device that is widely utilized

Table 15.2.4 Operational Amplifiers

Type	Purpose	Input bias current, nA	Input res., Ω	Supply voltage, V	Voltage gain	Unity gain bandwidth MHz
LM741	General purpose	500	2×10^6	+20	25,000	1.0
LM224	Quad gen. purpose	150	2×10^6	3 to 32	50,000	1.0
LM255	FET input	0.1	10^{12}	+22	50,000	2.5
LM444A	Quad FET input	0.005	10^{12}	+22	50,000	1.0

for accurate frequency control. This device produces an output frequency that is set by a digital input. It is a highly accurate and stable circuit. This circuit is often used to demodulate FM radio waves.

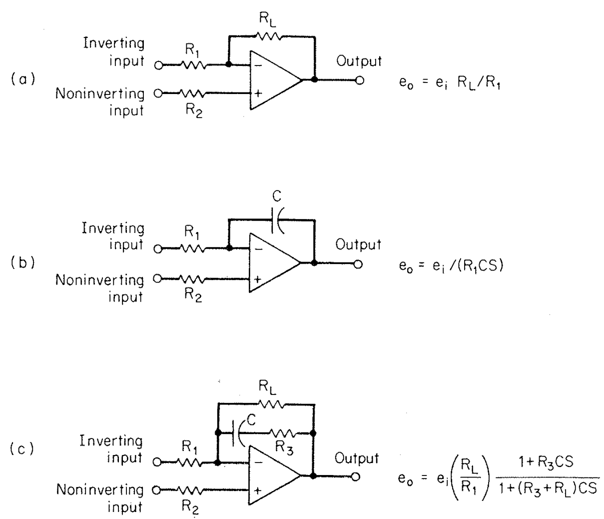


Fig. 15.2.25 Operational amplifier circuits.

Table 15.2.7 lists linear IC circuits that are used in telecommunications. These circuits include digital circuits within them and/or are used with digital devices. Whether these should be classed as linear ICs or digital ICs may be questioned. Several manufacturers include them in their linear device listings and not with their digital devices, and for this reason, they are listed here as linear devices. The radio-control **transmitter-encoder** and **receiver-decoder** provide a means of sending up to four control signals on a single radio-control frequency link. Each of the four channels can be either an on-off channel or a **pulse-width-modulated (PWM)** proportional channel. The **pulse-code modulator–coder-decoder (PCM CODEC)** is typical of a series of IC devices that have been designed to facilitate the design of digital-switched telephone circuits.

Table 15.2.5 Linear Integrated-Circuit Devices

Operational amplifier	Voltage comparator
Sample and hold	Active filters
Analog-to-digital converter	Digital-to-analog converter
Voltage regulator	Voltage reference
Voltage-controlled oscillator	NE 555 timer/oscillator

Table 15.2.6 Audio, Radio, and Television Integrated-Circuit Devices

Audio amplifier	Tone-volume-balance circuit
Dolby filter circuit	Phase-locked loop (PLL)
Intermediate frequency circuit	AM-FM radio
TV chroma demodulator	Digital tuner
Video-IF amplifier-detector	

Table 15.2.7 Telecommunication Integrated-Circuit Devices

Radio-control transmitter-encoder
Radio-control receiver-decoder
Pulse-code modulator–coder-decoder (PCM CODEC)
Single-chip programmable signal processor
Touch-tone generators
Modulator-demodulator (modem)

Digital Integrated Circuits

The basic circuit building block for digital ICs is the gate circuit. A gate is a switching amplifier that is designed to be either on or off. (By contrast, an operational amplifier is a proportional amplifier.) For 5-V logic levels, the gate switches to a 0 whenever its input falls below 0.8 V and to a 1 whenever its input exceeds 2.0 V. This arrangement ensures immunity to spurious noise impulses in both the 0 and the 1 state.

Several representative **transistor-transitor-logic (TTL) gates** are listed in Table 15.2.8. Gates can be combined to form logic devices of two fundamental kinds: combinational and sequential. In **combinational logic,** the output of a device changes whenever its input conditions change. The basic gate exemplifies this behavior.

A number of gates can be interconnected to form a **flip-flop** circuit. This is a bistable circuit that stays in a particular state, a 0 or a 1 state, until its "clock" input goes to a 1. At this time its output will stay in its present state or change to a new state depending upon its input just prior to the clock pulse. Its output will retain this information until the next time the clock goes to a 1. The flip-flop has memory, because it retains its output from one clock pulse to another. By connecting several flip-flops together, several sequential states can be defined permitting the design of a **sequential logic** circuit.

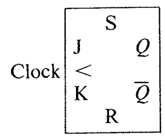
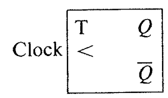
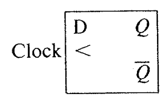
Table 15.2.9 shows three common flip-flops. The **truth table,** sometimes called a **state table,** shows the specification for the behavior of each circuit. The present output state of the flip-

Table 15.2.8 Digital Integrated-Circuit Devices

Type 54/74*	No. circuits per device	No. inputs per device	Function
00	4	2	NAND gate
02	4	2	NOR gate
04	6	1	Inverter
06	6	1	Buffer
08	4	2	AND gate
10	3	3	NAND gate
11	3	3	AND gate
13	2	4	Schmitt trigger
14	6	1	Schmitt trigger
20	2	4	NAND gate
21	2	4	AND gate
30	1	8	NAND gate
74	2		D flip-flop
76	2		JK flip-flop
77	4		Latch
86	4	2	EXCLUSIVE OR gate
174	6		D flip-flop
373	8		Latch
374	8		D flip-flop

NOTE: Example of device numbers are 74LS04, 54L04, 5477, and 74H10. The letters after the series number denote the speed and loading of the device.
*54 series devices are rated for temperatures from −55 to 125°C. 74 series devices are rated for temperatures from 0 to 70°C.

Table 15.2.9 Flip-Flop Sequential Devices

Name	Graphic symbol	Algebraic function	Truth table
JK flip-flop		$Q(t+1) = JQ'(t) + K'Q(t)$	J K $Q(t)$ | $Q(t+1)$ 0 X 0 | 0 1 X 0 | 1 X 0 1 | 1 X 1 1 | 0
T flip-flop		$Q(t+1) = TQ'(t) + T'Q(t)$	T $Q(t)$ | $Q(t+1)$ 0 0 | 0 0 1 | 1 1 0 | 1 1 1 | 0
D flip-flop		$Q(t+1) = D$	D $Q(t)$ | $Q(t+1)$ 0 0 | 0 0 1 | 0 1 0 | 1 1 1 | 1

flop is designated $Q(t)$. The next output state is designated $Q(t+1)$. In addition to the truth table, the Boolean algebra equations in Table 15.2.9 are another way to describe the behavior of the circuits. The **JK flip-flop** is the most versatile of these three flip-flops because of its separate J and K inputs. The **T flip-flop** is called a *toggle*. When its T input is a 1, its output toggles, from 0 to 1 or from 1 to 0, at each clock pulse. The **D flip-flop** is called a *data cell*. The output of the D flip-flop assumes the state of its input at each clock pulse and holds this data until the next clock pulse. The JK flip-flop can be made to function as a T flip-flop by applying the T input to both the J and K input terminals. The JK flip-flop can be made to function as a D flip-flop by applying the data signal to the J input and applying the inverted data signal to the K input. Some common IC flip-flops are listed in Table 15.2.8.

Various types of gates are shown in Table 15.2.10. Combinational logic defined by means of these various gates is used to define the input to flip-flops, which serve as memory devices. At each clock pulse, these flip-flops change state in accordance with their respective inputs. These new states are retained in the flip-flop and also applied to the gates. The output of the gates change (with only a small delay due propagation time), and at the next clock pulse the flip-flops will change to the next state as directed by the gates.

The gates shown in Table 15.2.10 have only two inputs. As was seen in Table 15.2.8, gates may have as many as eight inputs. In the case of an AND **gate,** all its inputs must be 1 in order for its output to be a 1. For an OR gate, if any of its inputs become a 1, then its output will become a 1. The NAND and NOR gates function in a similar way.

Boolean algebra is the branch of mathematics used to analyze logic circuits. Boolean algebra has two operators: $\cdot$, which indicates an AND operation, and $+$, which indicates an OR operation. The $=$ has the same meaning in Boolean algebra as in ordinary algebra. The symbol for "X not" is X' (or sometimes $\overline{X}$). The identity element for the AND operation is 0; the identity element for the OR operation is 1. The rules for Boolean algebra can be derived from set theory applied to a system in which only two numbers exist, i.e., zero and one. These rules are summarized in Huntington's postulates and DeMorgan's theorem and are listed in Table 15.2.11.

To facilitate the analysis of digital circuits and to aid in the application of the rules given in Table 15.2.11, **Karnaugh maps** are used. Typical two-variable and four-variable Karnaugh maps are shown in Fig. 15.2.26 along with the algebraic expressions represented by each map.

The complexity of digital ICs is growing. Where there is a sufficiently large demand for a particular circuit function, LSI and VLSI devices can be designed. Table 15.2.12 lists some of the highly complex circuits that are commercially available. **Read-only memory (ROM)** is a combinational logic device. This device can be programmed to accomplish the same functions that can be achieved with a complex circuit using various types of gates. An extension of the ROM is the **programmed-logic array (PLA).** This device is built specifically as a cost-effective combinational logic device for very complex logic systems.

Gate arrays are yet another means for designing custom IC logic circuits into a single VLSI device. Gate arrays include both combinational and sequential circuit elements. These circuit elements have been standardized so that a custom IC can be developed by specifying the interconnection, within the device, of standard elements to form a specific circuit. The development and tooling cost for gate array devices is much lower than that of a completely new IC. Gate arrays are used for production quantities of 1,000 to 10,000 devices per year. The cost of each gate array device is somewhat higher than a custom IC, so for quantities of more than 10,000 per year, custom devices are usually designed using normal mask techniques.

Computer Integrated Circuits

One of the devices which has become feasible as a result of VLSI is the **microprocessor.** A complete computer can be built in a single IC device. In most cases, however, several devices

Table 15.2.10 Combinational Gate Logic

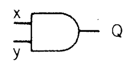
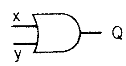
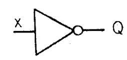
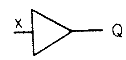
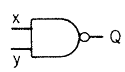
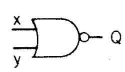
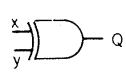
Name	Graphic symbol	Algebraic function	Truth table
AND		$Q = xy$	x y | Q 0 0 | 0 0 1 | 0 1 0 | 0 1 1 | 1
OR		$Q = x + y$	x y | Q 0 0 | 0 0 1 | 1 1 0 | 1 1 1 | 1
Inverter		$Q = x'$	x | Q 0 | 1 1 | 0
Buffer		$Q = x$	x | Q 0 | 0 1 | 1
NAND		$Q = (xy)'$	x y | Q 0 0 | 1 0 1 | 1 1 0 | 1 1 1 | 0
NOR		$Q = (x + y)'$	x y | Q 0 0 | 1 0 1 | 0 1 0 | 0 1 1 | 0
EXCLUSIVE-OR		$Q = xy' + x'y$ $= x + y$	x y | Q 0 0 | 0 0 1 | 1 1 0 | 1 1 1 | 0

Table 15.2.11 Rules for Boolean Algebra

$X + 0 = X$	$X \cdot 1 = X$
$X + 1 = 1$	$X \cdot X' = 0$
$X + X' = 1$	$X \cdot X = X$
$(X')' = X$	$X \cdot 0 = 0$
$X + Y = Y + X$	$X \cdot Y = Y \cdot X$
$X + (Y + Z) = (X + Y) + Z$	$X \cdot (Y Z) = (X \cdot Y)Z$
$X + X \cdot Y = X$	$X \cdot (X + Y) = X$
DeMorgan's theorem: $(X + Y)' = X' \cdot Y'$	$(X \cdot Y)' = X' + Y'$

Y \ X	0	1
0	1	0
1	0	1

$OUT = X'Y' + XY$

(a)

YZ \ WX	00	01	11	10
00	0	0	0	0
01	0	0	0	0
11	1	0	1	1
10	0	0	1	0

$OUT = WXY + X'YZ$

(b)

Fig 15.2.26 Karnaugh maps of typical logic functions.

are employed to build a complete computer system. In most computers, the cost of the microprocessor is negligible compared with the total system cost. Not long ago, the **central processing unit (CPU)** was the most expensive part of a computer. The microprocessor provides the total CPU function at a fraction of the earlier cost.

The power of a microprocessor is a function of its clock rate and the size of its internal registers. Clock rates vary from 1 to 20 MHz. Common register sizes are 8, 16, and 32 bits. Most of the existing personal computers currently use microprocessors which have 8-bit registers. New designs use 16-bit microprocessors. For scientific computing or high-resolution graphics, 32-bit machines are preferred.

As costs become lower, increased register sizes become more common. Increased register size allows a more powerful computer instruction set to be incorporated, and it also allows direct addressing of a larger memory. Both of these capabilities enhance the power of the machine.

Memory is an important part of a computer. Information that is processed by the computer is stored in **random-access memory (RAM).** The CPU can write information into RAM and subsequently read that information. Earlier, RAM was built using many little magnetic cores. Today core memory has been almost entirely replaced by semiconductor memory. Main computer memory is often referred to as "core memory," even though it may, in fact, be a solid-state memory. Semiconductor RAM memory may be dynamic or static. **Static RAM** retains information as long as electric power is applied to the circuit. **Dynamic RAM** retains information as stored charges in capacitors. Since the charge leaks off with time, dynamic memory must be refreshed about every 10 μs. The cost of dynamic memory devices is less than that of static memory. Large memory banks are usually made with dynamic memory because the cost of memory-refresh circuitry is offset by the savings in memory device costs.

The information in both static and dynamic IC memory is lost when power is removed. For this reason, IC memory is

Table 15.2.12 Large-Scale Digital Integrated Circuits

Adder	Accumulator
Arithmetic logic unit	Parity generator-checker
Shift register	Decoder-demultiplexer
Counter	Encoder-multiplexer
Display controller-driver	Custom gate array

termed **volatile memory.** Read-only memory is nonvolatile; the information in ROM is retained even if power is removed. The information in a ROM is masked into the device at the time it is manufactured. **Programmable read-only memory (PROM)** is ROM that can be programmed using a PROM programming machine. Some PROMs may be programmed only once. These ROMs are programmed by fusing links within the device. Once these internal links have been fused, the ROM cannot be changed. Other erasable PROMs, or **EPROMs**, can be erased using high-intensity ultraviolet light. Because these devices can be reprogrammed over and over again, EPROMs are often used during the product development phase and then replaced with less expensive PROMs or ROMs for production units.

Bubble memory is another nonvolatile memory. Information is stored in these devices as "magnetic bubbles." Bubble memory is slower than IC memory, but it is less expensive. Other forms of magnetic memory (disks and tapes) are less expensive than bubble memory and are much slower in response. Bubble memory, unlike tapes and disks, is completely free of moving parts. Other magnetic devices are discussed below with peripheral devices.

In addition to the CPU and memory ICs, other IC devices are required for computer support. These are listed in Table 15.2.13. One general-purpose support IC is called a **peripheral interface adapter (PIA).** The function of the PIA is to provide a programmable interface between the microprocessor and any peripheral device. This IC handles most of the functions needed to interface to the computer bus. A **universal asynchronous receiver-transmitter (UART)** interfaces between the computer and asynchronous devices such as **cathode ray tubes (CRTs)** and modulator-demodulators **(modems).** The **universal synchronous receiver-transmitter (USART)** performs a similar function for synchronous communications devices. The UART and USART perform the interface functions of the PIA and, in addition, perform functions specifically required for synchronous and asynchronous communications. The PIA is designed for broader applications, but the UART and USART perform more functions in their specific applications.

Another specific-purpose device is the **floppy disk controller.** This device has been designed to perform the interface and data reformatting tasks required to interface a floppy disk drive to a microcomputer system.

Other common interface devices are A/D and D/A converters. These devices provide a means of converting from analog to digital information and back again. These devices are usually designed so that several analog signals can be multiplexed through A/D or D/A device.

The interface to the IEEE 488 instrument bus is useful for interfacing a computer into an instrumentation system of this type. This interface provides the hardware requirements so that a program executed in the microprocessor can make it either a master or a slave in such a system.

Table 15.2.13 Digital Computer Integrated-Circuit Devices

Microprocessors	Programmable read-only memory
Static random-access memory	Dynamic random-access memory
Tristate buffer	Tristate transceiver
Programmable timer	Parallel interface adapter
Analog-to-digital converter	Digital-to-analog converter
Floppy disk controller	IEEE 488 bus interface
Universal asynchronous transmitter-receiver (UART)	

Computer peripheral devices consist of disks, tapes, printers, and terminals. There are two main types of disks: floppy disks and hard disks. The most common hard disk used in microcomputers is called a **Winchester disk.** Magnetic disks are nonvolatile memory devices. The access time for a disk is much longer than for IC memory. The cost of disk memory is much lower than for IC memory. For these reasons, disk memory is normally used for permanent storage rather than working storage. The disk is used for working storage when the amount of information to be stored exceeds IC memory capacity.

Magnetic tape is also nonvolatile memory. The cost of tape as a storage medium is quite low. Tape has the disadvantage, compared with disk storage, that data cannot be directly addressed. Tape is inherently a serial stream of data from the beginning of the tape to the end. Disk, on the other hand, is a random-access memory and any part of disk memory can be addressed directly. In general, data can be accessed more quickly from disk than from tape. Since floppy disks and magnetic tapes are both removable media, both of these storage devices are used for off-line storage of computer data.

Printers and terminals are the most common means of communicating between the computer and human users, although significant progress has been made recently with voice input and output to computers. The least expensive printers use a dot matrix to form characters. The printed results are not as pleasing as those from a font-oriented machine. Low-speed font-oriented printers are called **letter-quality printers.**

In addition to hardware, computers require software, or programs, in order to function. Each of the peripheral devices discussed above requires a program called a **driver.** The base program which controls a computer is called a **monitor** or an **operating system.** In addition to the drivers and operating system, most computers include utility programs which allow file maintenance, editing, etc. Application programs are added to complete the software for a computing system.

Programming of microcomputers requires nearly the same amount of time as for larger computers. The cost of microcomputer software has dropped only where there has been a high volume of sales of a given program. Engineering and manufacturing programs are often written specifically for an application. In these cases, software costs may far exceed hardware costs. Business programs have been commercialized much more effectively than technical programs.

Computer Communications

Lower-cost computing has greatly expanded the use of personal computers and the use of interactive graphics for design. **Computer-aided design and manufacturing (CAD/CAM)** is viewed by many to be a significant new development in manufacturing technology. The developments in both personal computing and CAD/CAM have increased the requirements for interprocessor communications.

Communication of information between computers can be accomplished by a number of means. Organizations which have established digital communication standards are listed in Table 15.2.14. Low-speed communications can be accomplished by asynchronous links using a 20-mA current loop or EIA RS-232C Standard. Higher-speed communications require the use of synchronous techniques, such as CCITT X.25 packet switching. A list of public data networks of this type is given in Table 15.2.15.

Table 15.2.14 Organizations Which Provide Communication Standards

CCITT	Comitè Consultatif Internationale de Telegraphie et Telephonie An international consultative committee that sets international communication usage standards
EIA	Electronic Industries Association A standards organization specializing in electrical and functional characteristics of interface equipment
ISO	International Organization for Standardization
ANSI	American National Standards Institute
IEEE	Institute of Electrical and Electronic Engineers

Table 15.2.15 Public Data Networks

Name	Location	Origination year
TELENET	United States	1975
EPSS	Britain	1977
DATAPAC	Canada	1977
TYMNET	United States	1977
TRANSPAC	France	1978
DX-2	Japan	1979
EURONET	Europe	1979

A fundamental relationship exists between digital and analog information called the **Nyquist criterion.** The required digital pulse rate depends upon the highest-frequency component contained in the analog information. Equation **(15.2.5)** shows this relationship. The minimum pulse rate (pulses per second) must be at least twice the highest frequency (hertz).

$$pr = 2f \tag{15.2.5}$$

This is a bilateral relationship. The frequency bandwidth of a transmission system must be equal to at least half the pulse rate.

$$BW = \frac{pr}{2} \tag{15.2.6}$$

These conditions are minimum requirements. A transmission system that has greater bandwidth will support slower pulse rates. A high pulse rate will approximate an analog signal more accurately than one that just meets the Nyquist criterion. Voice-grade telephone lines have a frequency bandwidth of about 5,000 Hz, so the maximum pulse rate for these lines is about 10,000 pulses per second. The rate normally stated is 9,600 baud (baud is equivalent to pulses per second).

There are several agencies which have written specifications or recommendations for data communications standards. ISO has established a layer standard for digital communications. The ISO layer model is shown in Fig. 15.2.27. The lowest layer in that model is the physical layer. This is essentially the function performed by the modem in a digital communication network. The link layer provides control of message routing through the communications system. The network layer provides the control specification for node addressing and packetizing of data. The top layer interfaces with the user; the bottom layer interfaces with network hardware.

The International Consultative Committee for Telephone and Telegraph (CCITT) has established a series of recommendations based on the layer approach to data communications. CCITT recommendation X.25, for packet-switched networks, has been gaining acceptance both in the United States and in Europe. This recommendation covers only layers 5, 6, and 7. The utility and the wide acceptance of X.25 are causing many manufacturers of computer communication equipment to design their equipment to meet this standard.

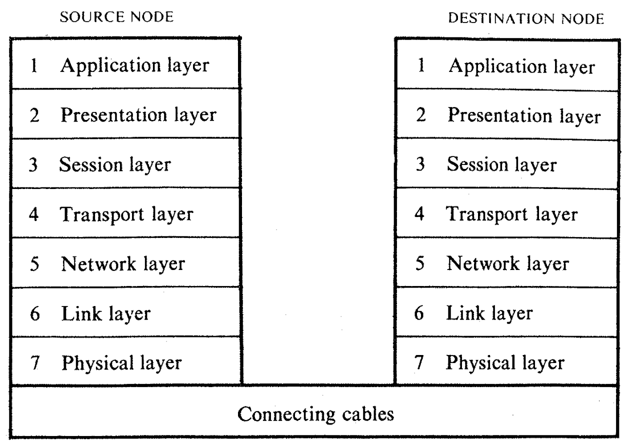


Fig 15.2.27 ISO-layered model for open system interconnection.

The Electronics Industries Association has established three interface standards which are frequently referenced for digital communications. These are RS-232C, RS-422, and RS-423. RS-232C is the oldest of these standards. This has been the primary standard for several years for low-speed-voltage-oriented digital communications. RS-232C and RS-423 use nonbalanced communication lines. Nonbalanced lines are more sensitive to noise. This limits the length of line and bandwidth that can be used satisfactorily. RS-232C is limited to a line length of about 250 ft at a bandwidth of 10 kHz. RS-423 is limited to a line length of about 400 ft at a bandwidth of 10 kHz. RS-422 uses a balanced line and can be used to a line length of 4,000 ft at a bandwidth of 100 kHz. It is expected that RS-423 and RS-422 standards will eventually replace the RS-232C standard.

Synchronous packetized data communication offers several advantages over asynchronous communications. For a given transmission medium, one can achieve higher data rates, better utilization of available bandwidth, and higher transmission accuracy. Each packet of data contains the source node address and the destination node address. This allows data packets to be routed through the network over alternate paths. In addition, each packet contains a **cyclic redundancy check (CRC).** At the source node, a CRC value is calculated based upon the data that are to be sent in that packet. At the receiving node, the CRC is recalculated based on the data that were received. The calculated CRC is compared with the transmitted CRC, and if there is an error, the data must be retransmitted. Data errors of less than one packet in 10^7 are easily achievable using CRC checking.

Many of the existing techniques and equipment that are used for digital communication by telephone lines are not suited to local data communication needs. Large-systems requirements place an overhead on the communication nodes that becomes burdensome. This reduces throughput and also introduces message setup delays that are unacceptable for many interactive computing situations. **Local area networks**

Table 15.2.16 Commercial Local Area Networks

Name	Sponsoring organization
Ethernet	Xerox/Digital Equipment/Intel
Net/One	Ungermann & Bass Co.
Z-Net	Zilog Corporation
Hyperbus	Networks System Corp.
Hyperchannel	Networks System Corp.
Ringnet	Prime Computer Corp.
Ring Token	Apollo Computer Corp.
Interactive System	3M
Data Exchange	Amdax
System 20	Sytek/NRC
Wangnet	Wang Computer Corp.

(LANs) have been devised to eliminate some of these problems. LANs can operate over distances of up to 1,000 ft or so with data rates of 10^6 baud. Table 15.2.16 lists several LANs.

Modern factories that make extensive use of computing equipment, both in design and in the shop, require very high communication rates and fast response. These requirements are not satisified by either LANs or telephone lines. Industrial dataways can be built using wideband cable-television coaxial cable and repeater amplifiers. These dataways have a bandwidth of 300 MHz. This will support data rates of up to 600 $\times 10^6$ baud.

For even greater communication bandwidths, light is being used to transmit information rather than electricity. Inexpensive fiber-optic devices can be used for distances of a few feet. For longer distances phase-locked lasers are being developed. Devices of this kind can be anticipated for use in both private and public data communication service.

Industrial Electronics

The power for dc motor armatures can be derived from thyristor circuits like those shown in Fig. 15.2.9. Single-phase bridge circuits are used for 5-hp drives and smaller. Three-phase bridge circuits are used for drives larger than 5 hp. A single set of six thyristors can supply power for about 300 hp. Above 300 hp, multiple sets of thyristors must be used in parallel. Mill drives have been built with more than 10,000 hp provided by thyristors.

The control of dc motors whether powered by thyristors or by dc generators is accomplished electronically. Control of individual drives can be accomplished by tachometer feedback or by armature voltage feedback. The speed-regulation accuracy for armature feedback is 5 percent; for tachometer feedback speed-regulation accuracy is from 0.1 to 1.0 percent. When two drives must be coordinated with each other, as in a continuous-web processing machine, they can be regulated to control torque, speed, position, draw, or a combination of these parameters. Torque controls can be achieved using dc motor armature current for a feedback signal. Speed-control signals are derived as for single motors. Position or draw control can be accomplished by using selsyn ties or dancer rolls. A **dancer roll** is a weight- or spring-loaded roll that rides on the web. It is free to move up and down, and as it does, a signal is taken from its position to serve as a feedback for the drive regulator before the dancer or after it.

Coordination of the motions of two or more drives requires tracking of the drives in both steady-state and transient conditions. Linearity of the control and feedback signals determine steady-state tracking. Provision must be made for both low-speed and high-speed matching signals. Transient matching requires that signals not only be the right magnitude but also arrive at the right time. An example will serve to illustrate this point. Suppose it is desired to have two drives with tachometer feedback which have a continuous web between them. One way to accomplish this would be to designate one drive as a master and the other as a slave. The tachometer on the master drive would serve as its own feedback signal and as the reference or command signal for the slave drive. The slave drive would have its own feedback from its own tachometer and so its regulator would try to minimize the difference between the two tachometer signals. On a transient basis the master drive will always start before the slave. An alternate and more common arrangement is to provide a common reference for both drives and let each drive receive its command signals at the same instant.

Digital computers are being used on-line in mills and continuous processing industries. DC motors can be controlled by either analog or digital regulators. With the greatly reduced cost of integrated circuits, digital regulators are being increasingly used.

DC motors have been widely used for variable-speed applications because of their excellent characteristics. AC motors have been used primarily for constant-speed applications. The control schemes described above are equally applicable to ac motors (except of course for armature voltage and armature current feedback). If power circuitry is properly handled, the control of an ac motor is just as flexible and versatile as that of a dc motor.

AC motors can be supplied either from **phase-controlled circuits** or from **inverter circuits.** Phase control is a simple electronic circuit, but its use results in high losses in the ac motor. This limits the application of this type of drive to either a very limited speed range or to loads in which the torque required decreases rapidly as the speed decreases. Large pump drives and fan drives have been built using this form of ac motor control. Inverters can be designed so that excessive motor losses are not encountered. Inverters are quite complex and require auxiliary power components to commutate the thyristors. Cost and complexity have prevented the widespread use of inverter-powered ac motor drives.

Phase-control circuits are extensively used to control power flow to process heaters. Most industrial heating is done by gas because it is cheaper than electrical energy. In many applications, electric heat is needed or is sufficiently more convenient. Phase-controlled thyristors modulate the power to these heaters and provide smoother control than simple on-off control by contactors.

High frequencies can be generated by **thyristor inverter** circuits. This permits the use of thyristors for **induction heating** and supersonic cleaning. Thyristor supplies have been built with frequency output from 100 to 50,000 Hz. These power supplies can be controlled in frequency much more easily and rapidly than motor-alternator sets and so have added new capability to induction-heating apparatus.

Dielectric heating requires frequencies from 100 kHz to 1 MHz. Large vacuum-tube oscillators are used to generate these frequencies.

Communications

The Federal Communications Commission (FCC) regulates the use of radio-frequency transmission in the United States.

Table 15.2.17 Partial Table of Frequency Allocations
(For a complete listing of frequency allocations, see "Reference Data for Radio Engineers," published by Howard Sams & Co.)

Frequency, MHz	Utilization
0.535–1.605	Commercial broadcast band
27.255	Citizens' personal radio
54–72	Television channels 2–4
76–88	Television channels 5–6
88–108	Frequency-modulation broadcasting
174–216	Television channels 7–13
460–470	Citizens' personal radio
470–890	Television channels 14–83

This regulation is necessary to prevent interfering transmissions of radio signals. Some of the frequency allocations are given in Table 15.2.17. The frequency bands are also classified as shown in Table 5.2.18. Very low frequencies are used for long-distance communications across the surface of the earth. Higher frequencies are limited to line-of-sight transmission. Because of bandwidth considerations, high frequencies are used for high-density communication links. Orbiting **satellites** allow the use of high-frequency transmission for long-distance high-density communications.

A **radio transmitter** is shown in Fig. 15.2.28. It consists of four basic parts: an rf oscillator tuned to the carrier frequency, an information-input device (microphone), a modulator to impress the input signal on the carrier, and an antenna to radiate the modulated carrier wave.

A **radio receiver** is shown in Fig. 15.2.29. This is called a **superheterodyne** receiver because it utilizes a frequency-mixing scheme. The tuned radio-frequency amplifier is tuned to receive the desired radio signal. The local oscillator is adjusted by the same tuning control to a lower frequency. The mixer produces an output frequency which is the difference between the incoming radio-signal frequency and the local-oscillator frequency. Since this difference frequency is constant for all tuning poistions, the intermediate-frequency amplifier always operates with a constant frequency. This allows optimum design of the intermediate-frequency (IF) amplifiers since they are constant-frequency amplifiers. The IF frequency signal is modulated in just the same way as the radio signal. The demodulator separates this audio signal, which is then amplified so that the loudspeaker can be driven.

The term **radar** is **derived** from the first letters of the words "*ra*dio *d*etection *a*nd *r*anging." It is essentially an echo system in which the location of an object is determined by sending out short pulses of radio waves and observing and measuring the time required for their reflections or echoes to return to the sending point. The time interval is a measure of the distance of the object from the transmitter. The velocity of radio waves

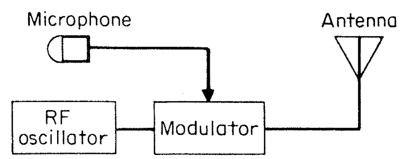
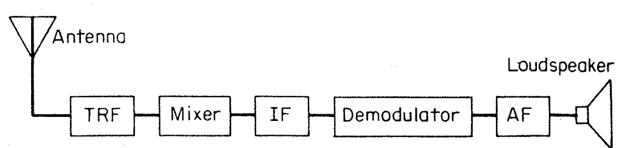


Fig. 15.2.28 Radio transmitter.

is the same as the velocity of light, or 984 ft/μs, so that each microsecond interval corresponds to a distance of 492 ft. The direction of an object can be determined by the position of the directional transmitting and receiving antenna. Radio waves penetrate darkness, fog, and clouds, and hence are able to detect objects that otherwise would remain concealed. Radar can be used for the automatic "tracking" of objects such as airplanes.



TRF Tuned radio frequency amplifier
IF Intermediate frequency amplifier
AF Audio frequency amplifier

Fig. 15.2.29 Radio receiver.

A block diagram of a radar system is shown in Fig. 15.2.30. The transmitting system consists of an rf oscillator which is controlled by a modulator, or pulser, so that it sends to the antenna intermittent trains of rf waves of relatively high power but of very short duration, corresponding to the pulses received by the modulator. The energy of the oscillator is transmitted through the duplexer and to the antenna through either coaxial cable or waveguides. The **receiver** is an ordinary heterodyne-type radio receiver which has high sensitivity in the band width corresponding to the frequency of the oscillator. For low frequencies the local oscillator is an ordinary oscillator for frequencies of 2,000 MHz; and higher a reflex **klystron** (hf cavity oscillator) is used. A common intermediate frequency is 30 MHz but 15 and 60 MHz are also frequently used.

In most radar systems the same antenna is used for receiving

Table 15.2.18 Frequency Bands

Designation	Frequency	Wavelength
VLF, very low frequency	3–30 kHz	100–10 km
LF, low frequency	30–300 kHz	10–1 km
MF, medium frequency	300–3,000 kHz	1,000–100 m
HF, high frequency	3–30 MHz	100–10 m
VHF, very high frequency	30–300 MHz	10–1 m
UHF, ultra-high frequency	300–3,000 MHz	100–10 cm
SHF, super-high frequency	3,000–30,000 MHz	10–1 cm
EHF, extremely high frequency	30,000–300,000 MHz	10–1 mm

NOTE: Wavelength in meters = $300/f$, where f is in megahertz.

as for transmitting. This requires the use of a **duplexer** which cuts off the receiver during the intervals when the oscillator is sending out pulses and disconnects the transmitter during the periods between these pulses when the echo is being received.

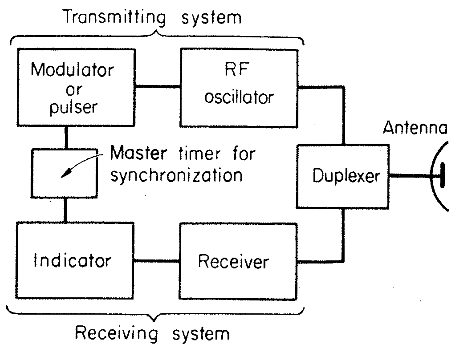


Fig. 15.2.30 Block diagram of radar system.

The antenna is highly directional. By noting its angular position, the direction of the object may be determined. In the PPI (plan position indicator), the angle of the sweep of the cathode-ray beam on the screen of the oscilloscope is made to correspond to the azimuth angle of the antenna.

The receiver output is delivered to the indicator which consists of a cathode-ray tube or oscilloscope. The pulses which are received, corresponding to echoes from the target, must be synchronized with the sending pulses in order that the distance to the target may be determined. This is accomplished by synchronization of the sweep circuit of the oscilloscope with the pulses by the master timer.

Displays Conversion of the received radar signals to usable display is accomplished by a cathode-ray oscilloscope. The simplest type, called the **A presentation,** is shown in Fig. 15.2.31*a*. When the pulser operates, a sawtoothed wave produces a linear sweep voltage (Fig. 15.2.31*b*) across the sweep plates of the cathode-ray tube; at the same time a transmitter pulse is impressed on the deflection plates and the return echoes appear as AM pulses, or "pips," on the screen, as shown in Fig. 15.2.31*a*. The distance on the screen between the transmitter pulse and the pip caused by the echo is proportional to the distance to the target, and the screen can be calibrated in distance such as miles. [The return of the spot to its initial starting position, produced by the sweep interval *cd* (Fig. 15.2.31*b*), is so rapid that it is not detectable by the eye.] The direction of the target may be determined by the angular position of the antenna, which can be transmitted to the operator by means of a selsyn. Different objects, such as airplanes, ships, islands, and land approaches, have characteristic pips, and operators become skilled in their interpretation. A bird in flight can be recognized on the screen. Also a portion of the scale such as *ab* can be segregated and amplified for close study of the characteristics of the pips.

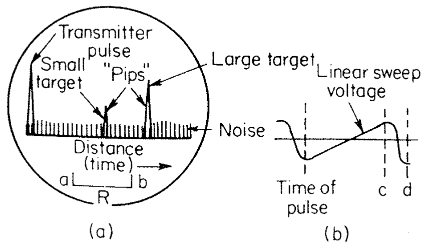


Fig. 15.2.31 Type A presentation.

Plan Position Indicator (PPI) In the PPI (Fig. 15.2.32) the direction of a radial sweep of the electron beam is synchronized with the azimuth sweep of the antenna. The sweep of the beam is rotated continuously in synchronism with the antenna, and the received signals intensity-modulate the electron beam as it sweeps from the center of the oscilloscope screen radially outward. In this way the direction and range position of an object can be determined from the pattern on the screen of the oscilloscope, as shown in Fig. 15.2.32.

There are two methods by which the angular direction of the cathode spot is made to correspond with the angular position of the antenna. In one method, used on board ship, two magnetic deflecting coils are rotated around the neck of the tube in synchronism with the antenna, by means of a selsyn. In the other method, used on aircraft, two fixed magnetic deflecting coils at right angles to each other and placed at the neck of the tube are supplied with current from a small two-phase synchronous generator whose rotor is driven by the antenna. Thus a rotating field, similar to that produced by the stator of an induction motor, is produced by the magnetic deflecting coils. These two rotating fields, although produced by different means, are equivalent and cause the cathode beam to sweep radially in synchronism with the antenna. Circular coordinates spaced radially corresponding to distance are obtained by impressing on the control electrode short positive pulses synchronized with the transmitted pulse but delayed by time values corresponding to the desired distances. These coordinates appear as circles on the screen. Since the time of rotation of the antenna is relatively slow, it is necessary that a persistent

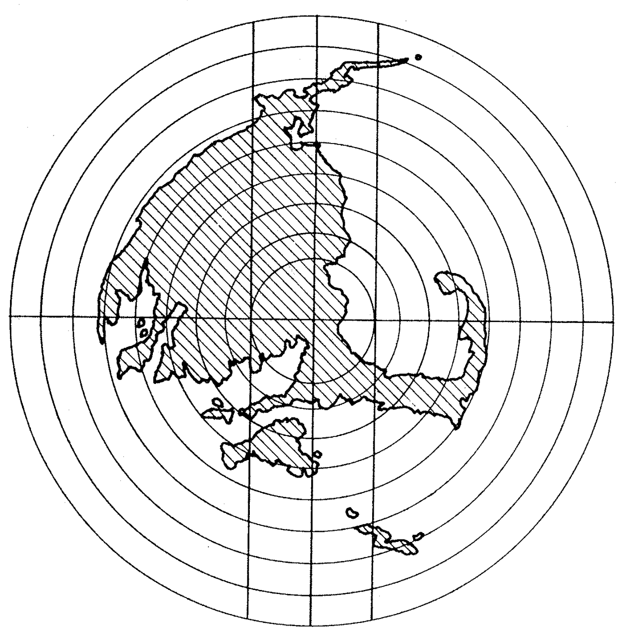


Fig. 15.2.32 Plan position indicator (PPI) of southeastern Massachusetts.

R.J. Reynolds Vapor Exhibit 1034-00373

screen be used in order that the operator may view the entire pattern. In Fig. 15.2.32 is shown a line drawing of a PPI presentation of Cape Cod, Mass., on a radar screen, taken from an airplane.

The applications of radar to war purposes are well known, such as detecting enemy ships and planes, aiming guns at them, and locating cities, rivers, mountains, and other landmarks in bombing operations. In peacetime, radar is used to navigate ships in darkness and poor visibility by locating navigational aids such as buoys and lighthouses, as well as protruding ledges, islands, and other landmarks. It can be similarly used in air navigation, as well as to operate altimeters for determining the height of the plane above ground. It is also used for aerial mapping.

There are also radio beacons, **shoran** (*sho*rt-*ran*ge navigation) and **loran** (*lo*ng-*ran*ge navigation) by which ships or planes can locate their positions. In the **ground-controlled approach** (GCA) for airplanes, the ground operator picks up the plane on a PPI presentation at distances up to 30 mi, using a general surveillance radar, and gives instructions to the pilot by radio course and procedure. As the plane approaches the landing field, it is brought into vision on the screen of a high-resolution short-range radar, and the pilot is given continual detailed instructions as to the glide path which the plane is to follow until the landing is made.

Television is accomplished by systematically scanning a scene or the image of a scene to be reproduced and transmitting at each instant a current or a voltage which is proportional to the light intensity of the elementary area of the scene which at the instant is being scanned. The varying voltage or current is amplified, modulated on a carrier wave, and then transmitted as a radio wave. At the receiver the radio wave enters the antenna, is amplified, and demodulated to give a voltage or a current wave similar to the original wave. This voltage or current wave is then used to control the intensity of a cathode-ray beam which is focused on a fluorescent screen in a cathode-ray reproducing tube. The cathode-ray beam is caused to move over the screen in the same pattern as the scanning beam at the transmitter and in synchronism with it. Thus each small area of the receiver screen is illuminated instantaneously with light intensity corresponding to that of a similarly placed area in the original scene. This process is conducted so rapidly that owing to persistence of vision of the eye, the reproduction of each instantaneous scene appears to be a complete picture and the effect with successive scenes is similar to that produced by the projection of successive frames of a motion picture.

Scanning and Blanking In the United States the ratio of width to height of a standard television picture is 4:3, and the picture is composed of 525 lines repeated 30 times a second, this last factor being one-half 60, the prevalent electric power frequency in the United States. The scanning sequence along the individual lines is from left to right and the sequence of the lines is from top to bottom. Also, interlacing is employed, the general method of which is shown in Fig. 15.2.33. The cathode-ray spot starts at 1 in the upper left-hand corner and is swept rapidly from left to right either by a sawtooth emf wave applied to the sweep plates or by the sawtooth current wave applied to the sweep coils of the tube. When the spot arrives at the right-hand side of the picture, the sawtooth wave of either emf or current in the sweep circuit acts to return the cathode-ray spot rapidly to point 3 at the left-hand side of the picture. However, during this period the cathode-ray is blanked, or entirely eliminated, by the application of a negative potential to the control grid of the tube. At the end of the return period, the blanking effect ceases and the spot appears at point 3, from which it again is swept across the picture and this process is

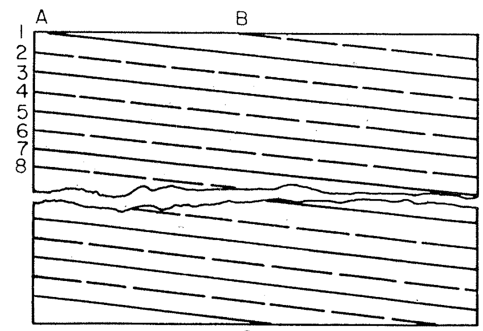


Fig. 15.2.33 Pattern of interlaced scanning.

repeated for 262.5 lines until the spot reaches a midpoint *C* at the bottom of the picture. It is then carried vertically and rapidly to *B*, the midpoint of the top of the picture, the beam also being blanked during this period. This process of scanning is then repeated, a second set of lines corresponding to the even numbers 2, 4, 6 being established between the lines designated by the odd numbers. These lines are shown dashed in Fig. 15.2.33. This method or pattern of scanning is called **interlacing.** The two sets of lines taken together produce a frame of 525 lines, which are repeated 30 times each second. However, owing to interlacing, the flicker frequency is 60 Hz which is not noticeable, 50 Hz having been determined as the threshold of flicker noticeable to the average eye. In Fig. 15.2.33, for the sake of clarity, the distances between horizontal lines are greatly exaggerated and no attempt is made to maintain proportions.

Frequency Band In order to obtain the necessary resolution of pictures, television frequencies must be high. In the United States, VHF frequencies from 54 to 88 MHz (omitting 72 to 76 MHz) and 174 to 216 MHz are assigned for television broadcasting. A UHF band of frequencies for commercial television use is also allocated and consists of the frequencies of from 470 to 890 MHz (see also Tables 15.2.17 and 15.2.18).

In order to obtain the 525 lines repeated 30 times per second, a band width of 6 MHz is necessary. The video, or picture, signal with the superimposed scanning and blanking pulses is amplitude-modulated, amplified, and transmitted. The carrier frequency associated with the sound transmitter is 4.5 MHz higher than the video carrier frequency and is frequency-modulated with a maximum frequency deviation of 25 kHz.

In scanning motion-picture films a complication arises because standard film rate is 24 frames per second, while the television rate is 30 frames per second. This difficulty is overcome by scanning the first of two successive film frames twice and the second frame three times at the 60-Hz rate, making the total time for the two frames $\frac{1}{12}$ ($\frac{2}{60}$ + $\frac{3}{60}$) or $\frac{1}{24}$ s average per frame.

Kinescope The kinescope (Fig. 15.2.34) is the terminal tube in which the televised picture is reproduced. It is relatively simple, being not unlike the cathode-ray oscilloscope tube. It has an electric gun operating at 8,000 to 20,000 V which produces an electron beam focused on a fluorescent surface within the front wall of the tube. The picture is viewed at the front wall. The horizontal and vertical deflections of the beam are

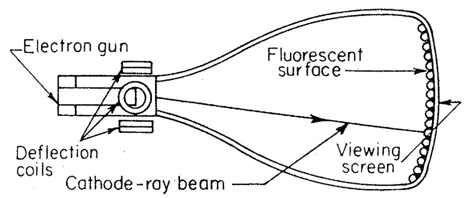


Fig. 15.2.34 Kinescope for television receiver.

normally controlled by deflection coils, as shown in Fig. 15.2.34.

Television Receivers A block diagram for a television receiver is given in Fig. 15.2.35. It is in reality a superheterodyne receiver with tuned rf amplification, the separating of the sound and video or picture channels taking place at the intermediate frequency in the mixer. The sound channel is then conventional, a discriminator being used to demodulate the FM wave (Fig. 15.2.21). The object of the dc restorer is to make the picture reproduction always positive, and it consists of applying a dc voltage at least equal in magnitude to the maximum values of the negative loops of the ac waves. The synchronizing pulses for both the vertical and the horizontal deflections are delivered by the dc restorer to an amplifier and the two pulses are then divided into the V and H components. The integrating and differentiating circuits are necessary to separate horizontal and vertical synchronizing signals.

As stated earlier, at any instant the magnitude of the current from the pickup tube varies in accordance with the light intensity of the part of the scene being scanned at that instant. This current is amplified and, together with the sound and synchronizing currents, is broadcast and received by the circuit shown in Fig. 15.2.35. The video current is detected by rectification, is amplified, and is then made to control the intensity of the kinescope electron beam. Tubes produce a scanning pattern, identical with that in the pickup tube, and these tubes are triggered by the synchronizing pulses which are transmitted in the broadcast wave. Hence, the original televised scene is reproduced on the fluorescent screen of the kinescope.

Color-television transmission is similar to black-and-white television, and the two signals must be compatible with each other. The kinescope for color TV has three electron guns, one for each primary color. The fluorescent screen has a matrix of three different colors of phosphor and a mask with many small holes in it. The intensity signals for each color are phase-shifted from each other so that the proper phosphors are excited by each electron stream at each mask point over the entire screen. A black-and-white signal does not have the same synchronizing signal as a color signal. The color receiver has circuits which recognize this state and switch it to black and white reception.

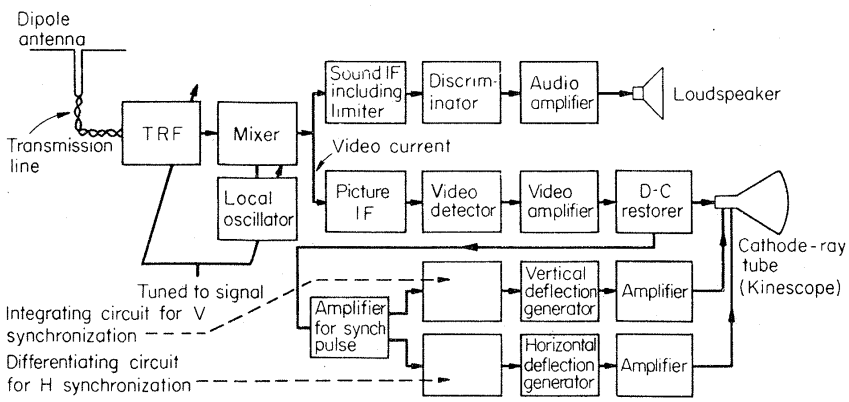


Fig. 15.2.35 Block diagram for televison receiver. TRF: tuned radio frequency; IF: intermediate frequency.